

제1권 품질관리 가이드라인 v3.0

CONTENTS

I 총론

제 1 장 추진 배경 및 목적	1
1. 추진 배경	1
2. 목적	2
제 2 장 품질관리 가이드라인 v3.0 구성	3
1. 제 1 권 품질관리 가이드라인 v3.0 구성	3
제 3 장 인공지능 학습용 데이터 품질관리 이해	5
1. 개요	5
2. 인공지능 학습용 데이터의 특성	6
3. 인공지능 학습용 데이터의 구축과정	7
4. 인공지능 학습용 데이터 품질관리 원칙	9
5. 인공지능 학습용 데이터 생애주기	10
6. 인공지능 학습용 데이터 생애주기 기반 품질관리 범위	12
7. 인공지능 학습용 데이터 품질관리 활동	13

II 품질관리 체계

제 1 장 품질관리 프레임워크	19
1. 품질관리 프레임워크 구성	19
2. 품질관리 이해관계자 정의	24
3. 품질관리 조직체계	26
제 2 장 품질관리 프로세스 및 산출물	29
1. 준비·계획단계(100)	29
2. 구축단계(200)	32
3. 운영·활용단계(300)	46
4. 단계별 산출물	52

제 3 장 품질 자가점검 및 품질검증	54
1. 개요	54
2. 품질 자가점검	55
3. 제 3 자 품질검증	62

III 부 록

부록 1. 품질관리 기준	75
1. 개요	75
2. 구축공정 품질관리 지표	78
3. 데이터 적합성 품질관리 지표	89
4. 데이터 정확성 품질관리 지표	100
5. 학습모델 품질관리 지표	105
6. 기타	138
부록 2. 품질지표 설정 가이드	140
1. 개요	140
2. 품질지표 선택가이드	146
[붙임 1] 품질지표·기준 선택 예시	181
[붙임 2] 데이터 라벨링 방법	204
부록 3. 개인정보보호 가이드	212
1. 개요	212
2. 생애주기별 개인정보보호 가이드라인	214

IV 참고자료

제 1 장 용어 정의	229
제 2 장 참고자료	236

인공지능 학습용 데이터

제 1 권 품질관리 가이드라인 v3.0 요약

- 인공지능 학습용 데이터 ‘제1권 품질관리 가이드라인’이란?
 - 과학기술정보통신부(이하 과기부)와 한국지능정보사회진흥원(이하 NIA)의 ‘인공지능 학습용 데이터 구축사업’의 사업수행기관(주관기관 및 참여기관), 사업관리기관, 제3자 품질검증기관 등 대상
 - ‘인공지능 학습용 데이터’를 구축함에 있어서 구축계획 수립단계부터 데이터 획득/수집, 정제, 가공 등에 대한 절차, 산출물 및 품질관리 활동을 제시한 기준서

특징

기존 ‘품질관리 가이드라인 v1.0’과 ‘품질관리 가이드라인 v2.0’을 품질관리 프레임워크 프로세스 기반으로 정립하여 체계화 및 고도화

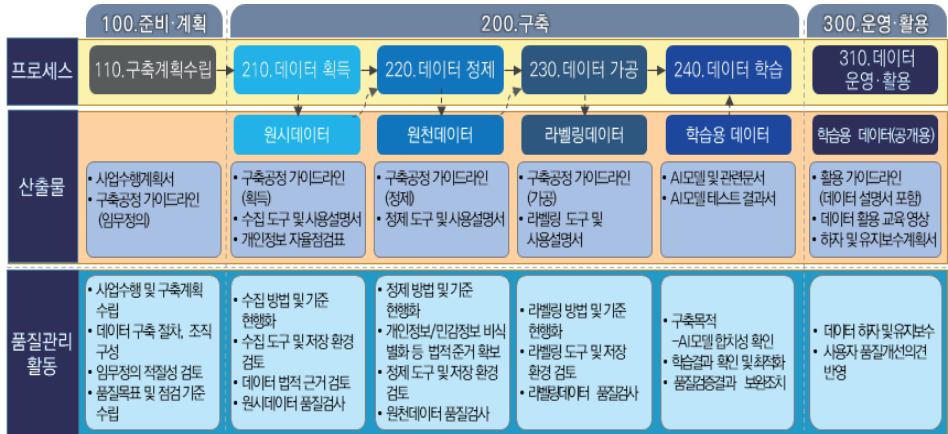
- 체계적 품질관리를 위한 ‘품질관리 프레임워크’ 정립
 - 데이터 품질관리 대상이 되는 구성 요소와 요소들 간의 관계 정의
 - 인공지능 학습용 데이터 구축사업 전반에 대한 이해와 품질관리에 대한 이해를 도울 수 있도록 ‘품질관리 프레임워크’를 제시



[그림 1] 인공지능 학습용 데이터 품질관리 프레임워크 구조

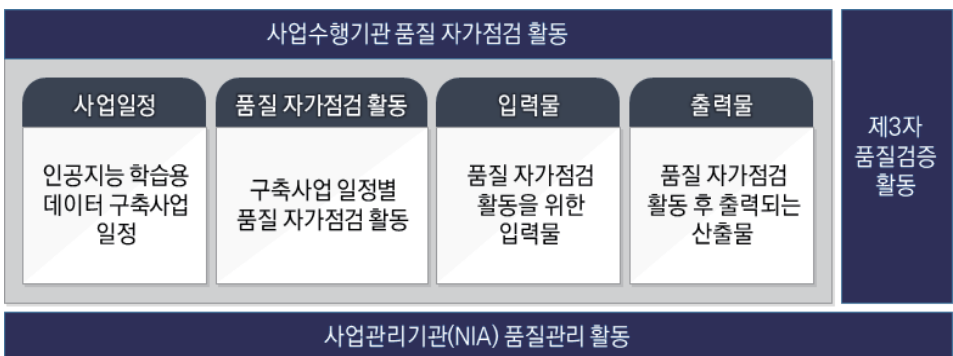
● ‘품질관리 프레임워크’의 프로세스 기반 산출물 및 품질관리 활동 정의

- 구축 프로세스에 따른 산출물과 품질관리 활동을 정의하여 공통된 공정에 따라 품질관리를 수행할 수 있도록 정립
- ‘품질관리 프레임워크’를 중심으로 단계별 프로세스와 품질관리 활동 및 산출물 등을 체계적으로 구성



[그림 2] 프로세스 기반 산출물 및 품질관리 활동

- 품질관리 활동 중 품질 자가점검 및 품질검증 활동은 인공지능 학습용 데이터 구축사업에서 변화되고 발전되는 품질관리 활동의 중요한 요소인 1-Cycle 자가점검에 대한 이해와 구축 시기별 품질관리 활동 제시



[그림 3] 품질 자가점검 및 품질검증 활동

- 인공지능 학습용 데이터 품질 수준 측정을 위한 품질관리 지표 및 사업수행기관의 데이터 유형별 (이미지, 비디오, 오디오, 텍스트 등) 품질지표 선정을 위한 방향성 제시



총론

제1장 추진 배경 및 목적

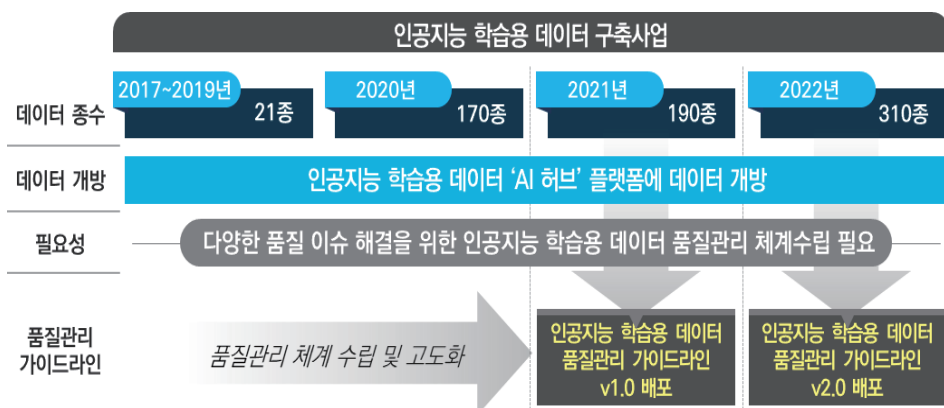
제2장 품질관리 가이드라인 v3.0 구성

제3장 인공지능 학습용 데이터 품질관리 이해

제1장 | 추진 배경 및 목적

1 추진 배경

- 정부의 인공지능 학습용 데이터 기반 정책 추진
 - 대규모의 인공지능 학습용 데이터 구축을 목표로 지능 데이터 기반 정책 추진



[그림 I-1] 인공지능 학습용 데이터 품질관리체계 수립 추진 배경

- 인공지능 학습용 데이터를 구축하고 민간에 개방함으로써 인공지능 기반의 산업 생태계 확산 추진
- 인공지능 학습용 데이터 품질관리 체계 수립 필요
 - 2017년부터 매년 인공지능 학습용 데이터 구축사업을 진행하여 데이터 개방
 - 개방된 인공지능 학습용 데이터의 데이터 품질관리 필요
 - 품질관리에 대한 체계 수립이 필요하여 사업수행기관¹⁾, 산업계 및 학계의 인공지능 및 품질 관리 전문가들의 의견을 반영한 '인공지능 학습용 데이터 품질관리 가이드라인' 배포

1) 사업수행기관 = 주관기관 + 참여기관

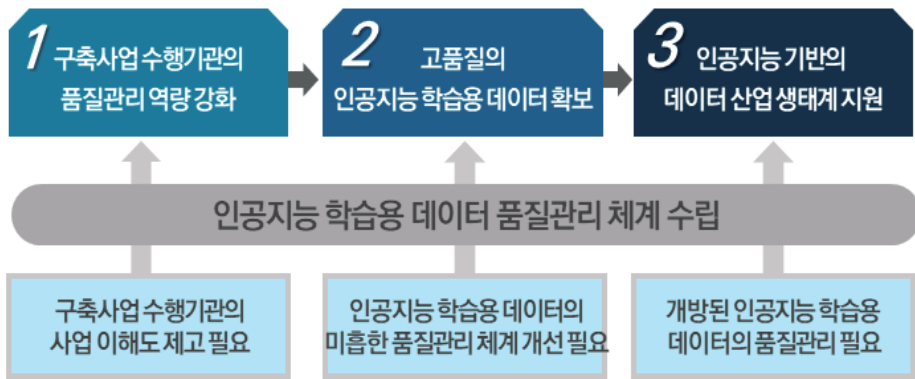
- (주관기관) 과제를 주관하여 수행하는 대표 기관(기업)

- (참여기관) 주관기관과 공동으로 해당 과제에 참여하여 사업을 수행하는 기관(기업)

- 인공지능 학습용 데이터 품질관리 가이드라인 v1.0 배포(2021.02.)
- 인공지능 학습용 데이터 품질관리 가이드라인 v2.0 배포(2022.02.)

2 목적

- 인공지능 학습용 데이터의 품질이 중요하다는 인식에 대해서는 공감대를 형성하고 있으나, 구체적인 품질관리 방법이나 절차 등에 대해서는 다양한 시각이 존재하므로 이에 대한 방향 제시 필요



[그림 1-2] 인공지능 학습용 데이터 품질관리 체계 수립 배경 및 목적

- ‘인공지능 학습용 데이터 구축사업’ 수행기관의 품질관리 역량 강화
 - 사업수행기관의 품질관리 역량을 강화하여 고품질의 학습용 데이터 확보
 - 인공지능 학습용 데이터 구축단계부터 품질관리 체계 확보
 - 사업수행기관의 담당자들에게 품질관리에 필요한 실무차원의 안내서(지침) 역할
- ‘인공지능 학습용 데이터 구축사업’의 구축 및 개방 데이터 품질 강화
 - 데이터의 품질관리에 대한 이해를 바탕으로 생애주기 전반의 품질관리에 기준 제공
 - 인공지능 학습용 데이터 구축에 대한 일관된 품질관리 기준 제시
- 인공지능 기반의 데이터 산업 생태계 지원
 - 인공지능 학습용 데이터 ‘AI 허브’ 플랫폼에 개방된 인공지능 학습용 데이터의 원활한 활용 지원

제2장 | 품질관리 가이드라인 v3.0 구성

- ‘품질관리 가이드라인 v3.0’ 구성
 - 『제1권. 품질관리 가이드라인 v3.0』, 『제2권. 데이터 구축안내서 v3.0』
 - 본 책자는 『제1권. 품질관리 가이드라인 v3.0』으로 해당 구성 내용 기술

1 제1권 품질관리 가이드라인 v3.0 구성

- Ⅰ. 총론 구성
 - 인공지능 학습용 데이터 구축사업 전반 및 품질관리에 대한 이해
- Ⅱ. 품질관리 체계 구성
 - ‘품질관리 프레임워크’에서 단계별 프로세스와 품질관리 활동 및 산출물 등을 체계적으로 구성
 - 인공지능 학습용 데이터 구축사업의 이해관계자 및 품질관리 조직의 역할 및 책임 정의
 - 품질관리 활동의 중요한 요소인 1-Cycle 자가점검에 대한 프로세스 및 품질관리 활동 제시
- Ⅲ. 부록 구성
 - 인공지능 학습용 데이터 품질 수준 측정을 위한 품질관리 지표
 - 사업수행기관의 데이터 유형별(이미지, 비디오, 오디오, 텍스트 등) 품질지표 선정을 위한 방향성 제시
 - 인공지능 학습용 데이터의 생애주기별 개인정보보호 가이드라인

목차	제목	소제목
Ⅰ. 총론	제1장 추진 배경 및 목적	1. 추진 배경
		2. 목적

목차	제목	소제목
	제2장 품질관리 가이드라인 v3.0 구성	1. 제1권 품질관리 가이드라인
	제3장 인공지능 학습용 데이터 품질관리 이해	1. 개요
		2. 인공지능 학습용 데이터의 특성
		3. 인공지능 학습용 데이터의 구축과정
		4. 인공지능 학습용 데이터 품질관리 원칙
		5. 인공지능 학습용 데이터 생애주기
		6. 인공지능 학습용 데이터 생애주기 기반 품질관리 범위
		7. 인공지능 학습용 데이터 품질관리 활동
II. 품질관리 체계	제1장 품질관리 프레임워크	1. 품질관리 프레임워크 구성
		2. 품질관리 이해관계자 정의
		3. 품질관리 조직체계
	제2장 품질관리 프로세스 및 산출물	1. 준비·계획단계
		2. 구축단계
		3. 운영·활용단계
		4. 단계별 산출물
	제3장 품질 자가점검 및 품질검증	1. 개요
		2. 품질 자가점검
		3. 제3자 품질검증
III. 부록	부록1. 품질관리 기준	1. 개요
		2. 구축공정 품질관리 지표
		3. 데이터 적합성 품질관리 지표
		4. 데이터 정확성 품질관리 지표
		5. 학습모델 품질관리 지표
		6. 기타
	부록 2. 품질지표 설정 가이드	1. 개요
		2. 품질지표 선택 가이드
	부록 3. 개인정보보호 가이드	1. 개요
		2. 생애주기별 개인정보보호 가이드라인
IV. 참고자료	제1장 용어 정의	
	제2장 참고자료	

〈표 I-2〉 품질관리 가이드라인 구성

제3장 | 인공지능 학습용 데이터 품질관리 이해

1 개요

- 인공지능 학습용 데이터의 구축과정과 이를 통해 생성되는 원시데이터, 원천데이터, 라벨링데이터 등 인공지능 학습용 데이터 특성 파악
- 데이터 구축과정을 기반으로 데이터의 생애주기를 운영 및 활용 영역까지 확대하여 정의하고, 생애주기 관점에서 데이터의 품질관리 영역 제시
- 구축한 데이터를 학습모델에 적용하여 인공지능 학습용 데이터로서 유효성을 확보했는지 자체 검사하도록 안내
- 인공지능 학습용 데이터(AI Dataset)란?
 - 머신러닝, 딥러닝 등 인공지능 모델 학습을 위해 활용되는 데이터를 총칭
 - 특히 ‘인공지능 학습용 데이터 구축’사업의 경우, 비정형 또는 반정형 데이터를 수집하고 참값(GT, Ground Truth) 어노테이션을 통해 라벨링하여, 지도학습(Supervised Learning)에 쓰이는 데이터를 구축하는데 초점
 - ※ AI Dataset = AI Training Data + Validation Data + Test Data
- ‘품질(Quality)’이란?
 - 사전적으로는 “물건의 성질과 바탕²⁾”이라고 정의하고 있으며, 데이터 관점에서는 일반적으로 ‘사용하기에 적절한 수준(Fit for use)’이라고 통용
 - 즉, 품질의 개념은 품질의 대상이 무엇인가에 따라 달라질 수 있음
 - 위 개념의 정의를 따랐을 경우, ‘인공지능 학습용 데이터 품질’이란, ‘데이터 품질 확보를 위한 품질관리 측면과 인공지능 학습에 필요한 데이터 자체의 품질을 확보하여 사용자에게 유용한 가치를 줄 수 있는 수준’이라고 정의

2) 표준국어대사전 인용

- ‘품질관리(Quality Management)’란?

- 품질을 사용자 요구에 맞춰 지속적으로 유지하거나 향상시키기 위한 체계적인 관리와 활동을 의미
- 이러한 품질관리 개념을 인공지능 학습용 데이터의 품질관리에 적용하면 ‘인공지능 학습용 데이터 품질을 확보하는 데 필요한 조직, 절차, 품질기준, 품질관리 방법이나 활동 등을 정의하여 점검하고 조치하는 일련의 활동’으로 정의

- 품질관리 활동에서의 품질검사, 품질검증, 품질관리에 대한 정의

구분	품질검사 (Quality Inspection)	품질검증 (Quality Verification)	품질관리 (Quality Management)
정의	인공지능 학습용 데이터 구축 및 모델 관점에서 주요 품질 문제를 식별하고 문제의 근원적 원인을 파악하여 개선 기회를 도출하는 기법을 의미	인공지능 학습용 데이터 구축사업의 품질보장을 위해 사업수행기관이 구축한 데이터의 품질을 제3자가 품질검증	인공지능 학습용 데이터의 구축 시 생산되는 원시데이터, 원천데이터, 라벨링데이터와 구축공정의 품질을 확보하기 위한 제반 활동
중점	데이터의 특성을 고려한 기준을 반영하여 지속가능한 품질확보에 중점	데이터의 결함을 예방하고 오류를 방지하는데 중점	데이터의 부적합, 결함을 감지하고 시정 조치 활동에 중점
대상	특정 주요 지표 및 절차	특정 주요 지표 및 절차	원칙, 대상, 기준, 조직, 절차 등
주체	사업수행기관	품질검증기관	사업관리기관, 품질검증기관, 사업수행기관

〈표 I-2〉 품질검사/품질검증/품질관리 정의

2 인공지능 학습용 데이터의 특성

- 인공지능 학습용 데이터의 품질관리를 위해서는 데이터가 갖는 특성을 파악하여 이를 기반으로 데이터의 품질확보에 적합한 품질관리 체계를 확보하는 것이 중요
- 인공지능 학습용 데이터의 품질확보를 위해서는 인공지능 학습용 데이터의 특성을 고려하여, 데이터베이스(Data Base)에 저장하는 구조화된 데이터(Structured Data)를 대상으로 하는 일반적인 빅데이터 품질관리와는 차별화된 품질관리 체계 확보 중요

구분	인공지능 학습용 데이터	빅데이터
주요 차이점	• 임무(Mission)로 정한 AI모델의 학습에 초점	• 대규모 데이터에서 숨겨진 가치를 찾는 것에 초점
구축목적	• 인공지능 모델의 학습 목적	• 인사이트 도출 및 비즈니스 개선
사용기술	• 예측, 분류, 군집화, 차원축소	• 통계적 분석, 텍스트 마이닝
데이터유형	• 이미지, 비디오, 오디오, 텍스트, 3D 등 비정형 데이터 중심	• 텍스트 데이터, 정형 데이터
데이터구조	• 원천데이터와 라벨링데이터가 쌍으로 구성	• 키와 값으로 구성
데이터 구축절차	• 구축계획 수립→ 획득/수집 → 정제 → 가공 → 검사 → 학습 → 저장	• 획득 → 정제 → 변환 → 저장
데이터 저장방식	• 파일형태로 저장 (원천데이터+JSON파일)	• 범용적인 NoSQL DBMS에 저장 (Key+Value 구조)

※ 해당 표의 ‘인공지능 학습용 데이터’의 경우, 지도학습(Supervised Learning)에 쓰이는 데이터로 한정함

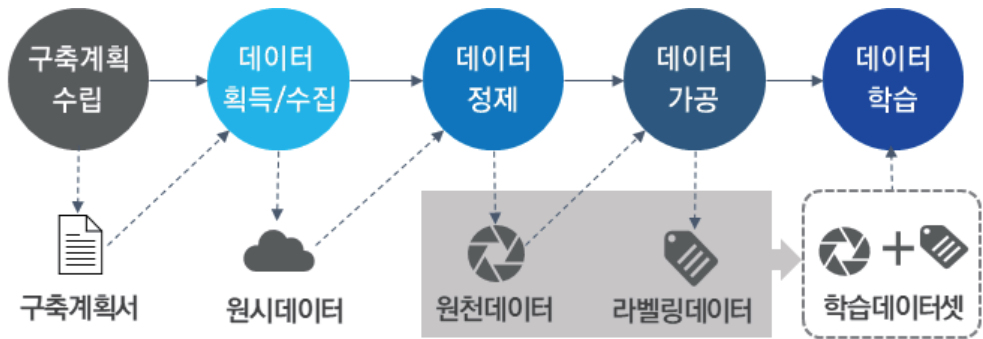
〈표 1-3〉 인공지능 학습용 데이터와 빅데이터의 비교

● 인공지능 학습용 데이터의 특성

- 첫째, 특정한 AI모델의 학습을 목적으로 생산되는 데이터로써 임무정의에 따라 구축되는 특성을 가짐
- 둘째, 지도학습이라는 특징에 따라 이미지, 비디오, 오디오, 텍스트 등 비정형 데이터를 대상으로 라벨링 작업을 통해 참값(GT)을 부여하며, 참값의 품질 여부에 따라 쓰임이 정해지는 등 품질이 엄격히 통제되어야 하는 특성을 가짐
- 셋째, AI 모델 학습용으로 구축되는 데이터는 활용자의 아이디어에 따라 무한한 활용 가능성을 가진다는 특성을 가짐

3 인공지능 학습용 데이터의 구축과정

- 인공지능 학습용 데이터 구축과정 과정에서 확보된 품질이 학습데이터 전체의 품질을 결정하기 때문에, 인공지능 학습용 데이터의 품질관리를 위해서는 데이터 구축 시 수행되는 구축공정의 이해 필요
- 데이터 구축과정 과정은 다양한 유형의 데이터가 사전에 정의된 목적에 따라 구축되기 때문에 구축과정 과정이 세부적으로는 상이할 수 있으나, 구축계획 수립, 데이터 획득/수집, 데이터 정제, 데이터 가공, 데이터 학습의 순서로 진행



[그림 1-3] 인공지능 학습용 데이터 구축 과정

● 구축계획 수립단계

- 인공지능이 기계학습을 통해 해결하고자 하는 문제를 명확하게 정의하고, 문제 해결에 필요한 인공지능 학습용 데이터를 구체적으로 정의하고 설계하는 활동
- 구축계획서에는 ‘구축 개요’, ‘데이터 구축’, ‘품질 자가점검 계획’ 등을 작성하며, 작성 요령은 『제2권. 데이터 구축안내서 v3.0』의 구축 계획서 자료 참고

● 데이터 획득/수집단계

- 인공지능의 기계학습에 필요한 데이터를 현실 세계에서 직접 생성하거나, 이미 보유하고 있는 조직이나 시스템 등으로부터 인공지능 학습에 필요한 데이터를 확보하는 등, ‘원시데이터’를 확보하는 활동 수행
- ‘원시데이터’는 기계학습을 목적으로 획득/수집단계에서 획득 또는 생성한 ‘이미지’, ‘비디오’, ‘오디오’, ‘텍스트’, ‘3D’ 등의 데이터를 의미

● 데이터 정제단계

- 획득한 원시데이터를 기계학습에 필요한 형식이나 크기로 맞추고, 데이터 중복 제거
- 원시데이터 획득 시 포함된 개인정보를 비식별화하여 처리하는 등의 과정을 통해 ‘원천데이터’ 확보
- 이 단계에서 확보된 ‘원천데이터’는 라벨링이 부여되지 않은 상태의 데이터를 의미(원천데이터는 원시데이터를 라벨링 작업에 투입하기 위해 필요한 정제 작업을 수행한 데이터)

- 데이터 가공단계
 - 인공지능이 기계학습에 활용할 수 있도록 기능이나 목적에 부합하는 라벨을 원천데이터에 부착하는 ‘라벨링’ 작업 수행
 - 여기서 생성된 ‘라벨링데이터’는 원천데이터에 부여한 ‘참값’, 파일 형식, 해상도 등의 데이터 속성과 설명, 주석 등이 포함된 ‘어노테이션’의 집합을 의미
- 데이터 학습단계
 - ‘학습데이터셋’을 이용하여 사전에 정의된 인공지능 알고리즘을 학습시키고, 학습된 인공지능 모델의 성능을 향상시키거나 보정하는 활동 수행
 - ‘학습데이터셋’은 ‘원천데이터’와 ‘라벨링데이터’의 묶음을 의미

4 인공지능 학습용 데이터 품질관리 원칙

- 고품질의 인공지능 학습용 데이터 확보를 위해서는 유형별 인공지능 학습용 데이터인 원시데이터, 원천데이터, 라벨링데이터 자체의 품질확보와 데이터 품질을 사용자 요구에 맞춰 지속적으로 유지시키고 향상시키기 위한 품질관리 체계 확보 필요
- 인공지능 학습용 데이터의 품질관리가 지향해야 하는 방향성을 제시하는 상위 수준의 ‘품질관리 원칙’이 필요
- 인공지능 학습용 데이터의 품질관리 원칙을 크게 데이터 품질 확보를 위한 품질관리 측면과 데이터 측면으로 구분하여 제시

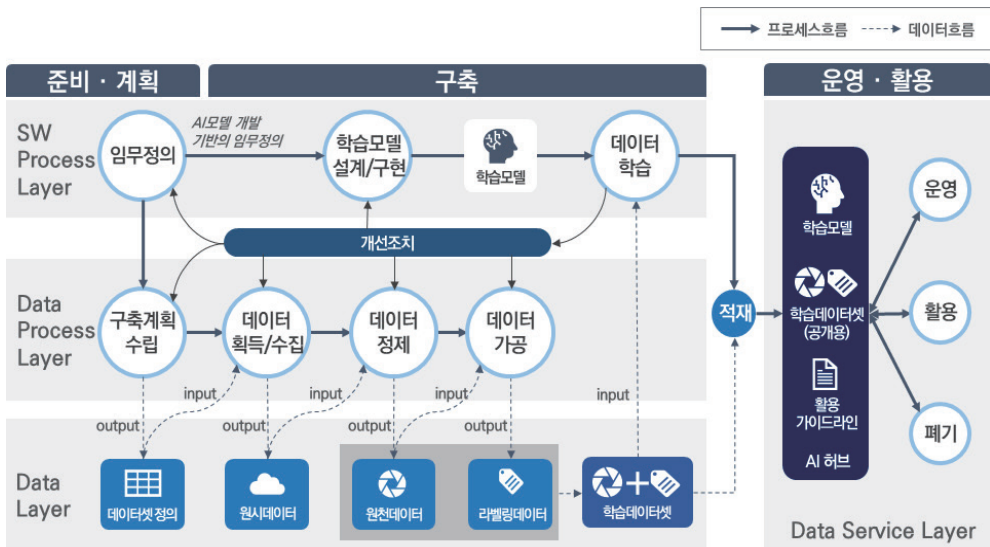
구분	원칙
품질관리 측면	<ul style="list-style-type: none"> • 인공지능 학습용 데이터의 전 생애주기의 품질을 보장해야 함 • 상시적이고 지속적인 인공지능 학습용 데이터 품질 개선이 가능해야 함 • 데이터 품질관리를 위한 조직을 구성하고, 정해진 역할과 책임에 따라 수행해야 함 • 조직의 품질관리 역량을 확보하도록 품질관리 교육 등 지원체계를 확보해야 함
데이터 측면	<ul style="list-style-type: none"> • 인공지능 학습용 데이터가 학습하는 데 필요한 요구사항을 충족해야 함 • 법률적 제약 없이 누구나 활용 가능해야 함 • 학습 목적에 부합하도록 획득/수집되어야 하며, 획득/수집한 데이터는 중복없이 원하는 목적에 따라 정제되어야 함 • 인공지능 학습용 데이터에 부여하는 참값(Ground Truth) 등의 라벨링 정보의 정확성이 확보되어야 함 • 인공지능 학습 모델을 통해 목표로 하는 유효성이 확보되어야 함

〈표 1-4〉 품질관리 원칙

5 인공지능 학습용 데이터 생애주기

5.1 인공지능 학습용 데이터 생애주기

- 인공지능 학습용 데이터의 품질관리 활동을 식별하기 위해서는 데이터의 생애주기를 분석하고, 생애주기를 기반으로 품질관리 범주 식별 필요
- 인공지능 학습용 데이터를 구축하여 민간에 개방하기까지의 생애주기는 다음과 같음

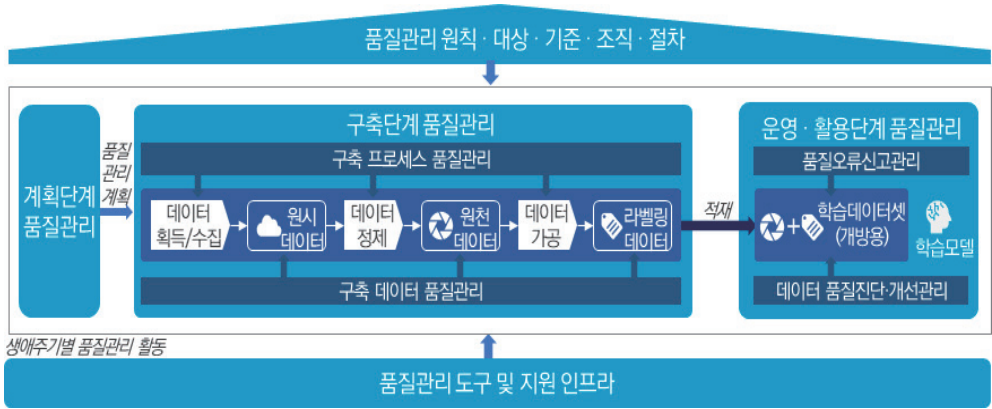


[그림 I-4] 인공지능 학습용 데이터 생애주기

- 인공지능 학습용 데이터의 생애주기는 크게 '준비·계획', '구축', '운영·활용' 영역으로 구분되며, 각 영역에서는 품질관리의 주체를 식별하고, 세부 활동 정의
- 업무를 수행하는 과정에서 생기는 의사소통의 어려움이나 혼란 방지

5.2 생애주기 기반 품질관리체계

- 고품질의 인공지능 학습용 데이터를 확보하고, 지속적인 품질을 유지하거나 품질향상을 위한 인공지능 학습용 데이터의 품질관리체계를 생애주기 기반으로 제시하면 다음과 같음

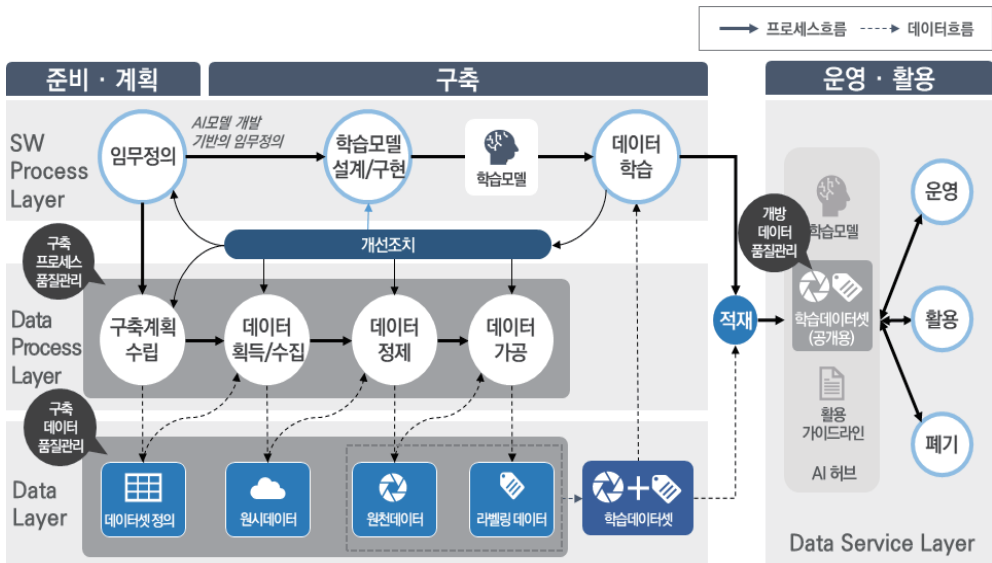


[그림 1-5] 생애주기 기반 품질관리체계

- (품질관리 원칙) 품질관리 측면(4개 원칙)과 데이터 측면(5개 원칙)에서의 9개 원칙
- (품질관리 대상) 인공지능 학습용 데이터 구축사업 및 AI Hub를 통해 민간에 개방하는 인공지능 학습용 데이터셋 대상
- (품질관리 기준) 품질관리 원칙의 준수에 대한 객관적인 검증지표를 말하며, 구축공정 품질기준과 데이터 자체의 품질기준으로 나눔
- (품질관리 조직) 인공지능 학습용 데이터의 품질확보 및 품질관리 활동을 수행하는 조직을 의미하며, 역할 및 책임을 부여하고 품질관리 활동을 수행
- (품질관리 절차) 인공지능 학습용 데이터의 품질을 검사하고, 저품질이나 오류 등의 원인을 분석해서 개선 조치를 하는 일련의 활동을 의미
- (생애주기별 품질관리 활동) 계획단계, 구축단계, 운영·활용단계의 각 영역에서 수행해야 할 품질관리 활동 정의
- (품질관리 도구 및 지원 인프라) 품질검사나 품질관리 활동을 수행하는 데 사용하는 도구(Tool)나 기술(Technology), 플랫폼(Platform) 등을 의미

6 인공지능 학습용 데이터 생애주기 기반 품질관리 범위

- 인공지능 학습용 데이터의 생애주기를 기반으로 품질관리 대상 정의
- 구축 프로세스 품질관리
 - 데이터 획득/수집, 데이터 정제, 데이터 가공 등 구축과정에서 원하는 수준의 데이터 품질 보장을 위하여 품질관리 관점에서의 활동을 수행하는지 모니터링하고, 발견된 문제점 보완
- 구축 데이터 품질관리
 - 구축사업을 통해 생성되는 원시데이터, 원천데이터, 라벨링데이터 등 데이터 자체의 품질을 검사하고, 발견된 오류를 개선하는 활동
- 개방 데이터 품질관리
 - 구축사업을 통해 AI Hub에 적재된 데이터를 대상으로 운영·활용 단계에서 학습용 데이터셋의 품질을 향상시키는 지속적 활동 의미
 - 실제 민간에 개방한 학습데이터의 품질에 개선의견이 제기되는 경우, 이에 적극적으로 대응 및 조치하는 품질관리 활동



[그림 I-6] 인공지능 학습용 데이터 품질관리 범위

단계	품질관리 영역	수행 내용
준비·계획단계 구축단계	구축 프로세스	<ul style="list-style-type: none"> 구축사업에 참여하는 사업수행기관이 자체적으로 품질관리 수행 민간에 개방된 학습데이터셋의 품질 개선의견이 제기된 경우, 사업수행기관이 품질개선 활동 직접 수행
	구축 데이터	
운영·활용 단계	개방 데이터	<ul style="list-style-type: none"> AI Hub를 통해 개방된 데이터의 개선사항을 도출하여 고품질의 데이터를 지속적으로 유지 및 운영하기 위해 구축사업에 참여한 사업수행기관과 사업관리기관이 함께 품질관리 활동 수행 사업관리기관은 사용자로부터의 데이터 품질에 대한 개선의견을 검토하고, 해당 데이터를 구축한 사업수행기관에 품질 개선을 요청하며, 사업수행기관은 이를 반영하여 품질 개선 수행

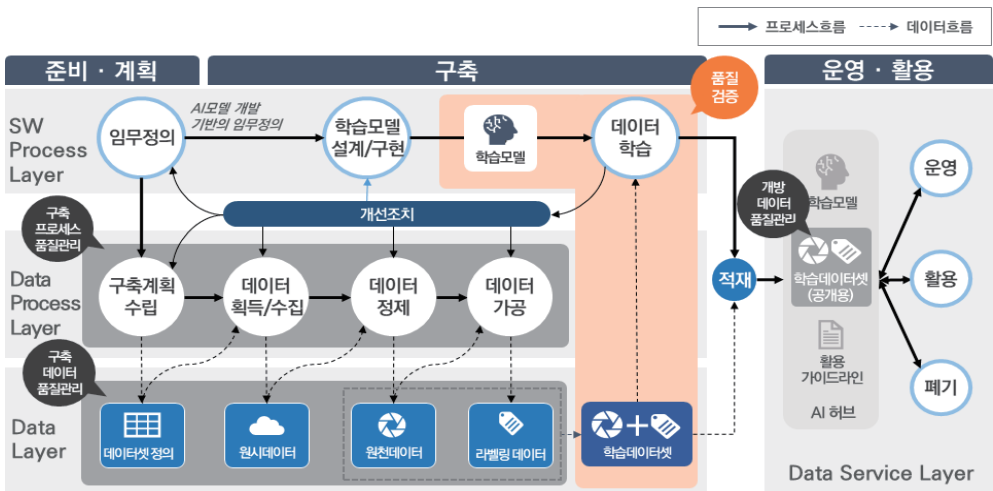
〈표 Ⅰ-5〉 품질관리 영역별 수행 내용

- 사업관리기관에서는 인공지능 학습용 데이터 구축사업의 품질 보장을 위해, 사업수행기관이 실시한 품질관리 활동을 객관적으로 확인하는 제3자 품질검증 활동을 수행하고 있으며, 해당 범위는 다음 그림과 같음

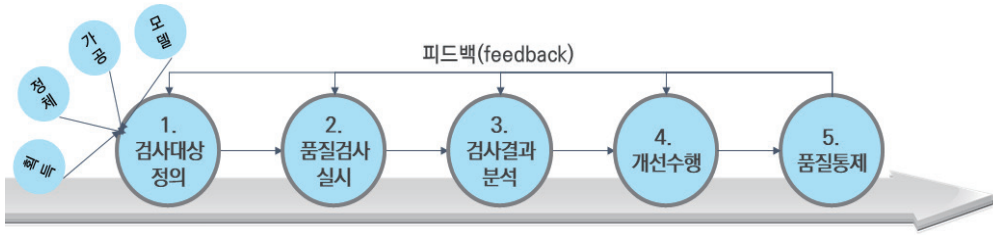
7 인공지능 학습용 데이터 품질관리 활동

7.1 품질관리 수행 절차

- 사업수행기관 및 참여기관의 품질관리자가 구축공정별 품질관리를 위해 수행해야 하는 품질검사 방법을 아래의 절차로 정의함



〔그림 Ⅰ-7〕 인공지능 학습용 데이터 품질검증 범위



[그림 I-8] 인공지능 학습용 데이터 품질검사 및 개선절차

번호	품질관리 수행 절차	수행 내용
1	검사대상 정의	<ul style="list-style-type: none"> 원시데이터, 원천데이터, 라벨링데이터 등의 검사대상을 선정하고 해당 데이터가 요구되는 품질수준에 부합하는 상태인지를 판단하기 위한 품질검사 계획 수립
2	품질검사 실시	<ul style="list-style-type: none"> 정의된 검사대상에 대해 적합성, 정확성, 유효성, 준비성, 완전성, 유용성 등의 품질관리 지표에 대해 체크리스트와 같은 검사기법을 적용하여 품질검사 실시
3	검사결과 분석	<ul style="list-style-type: none"> 품질검사 결과를 바탕으로 주요 품질문제를 식별하고 문제의 근본적 원인을 파악하여 품질문제를 해결하기 위한 개선 기회를 도출하는 단계
4	개선수행	<ul style="list-style-type: none"> 품질문제 해결을 위해 개선계획 및 방안을 정의하고 우선순위를 결정하며, 결정된 우선순위에 따라 데이터 보정, 추가 작업 등 개선 영역별 품질개선 활동 수행
5	품질통제	<ul style="list-style-type: none"> 개선수행 결과의 확인 및 점검을 통해 품질 목표를 재설정하여 품질문제 재발 방지 및 고품질 데이터를 유지하기 위한 품질관리 활동 수행 단계

〈표 I-6〉 품질관리 영역별 수행 내용

7.2 품질검사 수행 범위

- 인공지능 학습용 데이터 구축사업은 구축계획 수립, 데이터 획득/수집, 데이터 정제, 데이터 가공 및 데이터 개방의 생애주기(Life Cycle)를 가지며, 이에 따라 품질지표와 생애주기 관점에서 각 품질검사 활동을 수행하며, 이에 따른 각 절차의 수행 범위는 다음과 같음

지표		생애주기(Life Cycle)				
		구축계획 수립	데이터 획득/수집	데이터 정제	데이터 가공	데이터 학습
구축 공정	준비성	●	●	●	●	
	완전성		●	●	●	
	유용성	●				●
데이터 적합성	기준 적합성	●	●			
	기술 적합성	●	●	●	●	
	통계적 다양성	●	●			
데이터 정확성	의미 정확성				●	
	구문 정확성				●	
유효성	알고리즘 적정성					●
	학습 모델 유효성					●

※ 생애주기별 품질검사 적용지표를 활용하여 인공지능 학습용 데이터 구축사업과 사업수행기관 실정에 적합한 지표 선정

〈표 I -7〉 생애주기별 품질검사 적용지표

- 각 품질관리 지표의 적용은 본 『제1권. 품질관리 가이드라인 v3.0』 부록에 제시된 검사지표 및 방법에 해당 사항을 선택하여 적용
- 인공지능 학습용 데이터 구축의 목적 및 데이터 특성 등에 따라 필요한 검사기준을 추가할 수 있음

7.3 품질검사 방법

- (육안 검사) 사람의 육안으로 데이터의 오류, 손상 또는 데이터 자체의 이상 유무를 검사하는 품질검사의 기본적인 방법
- (도구 검사) 정제 도구, 품질검사 도구, 체크리스트 등 다양한 방법으로 검사



품질관리 체계

제1장 품질관리 프레임워크

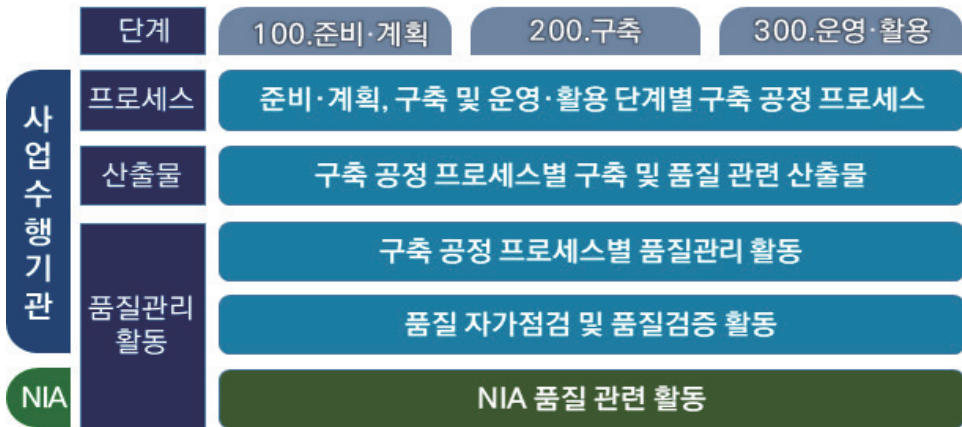
제2장 품질관리 프로세스 및 산출물

제3장 품질 자가점검 및 품질검증

제1장 | 품질관리 프레임워크

1 품질관리 프레임워크 구성

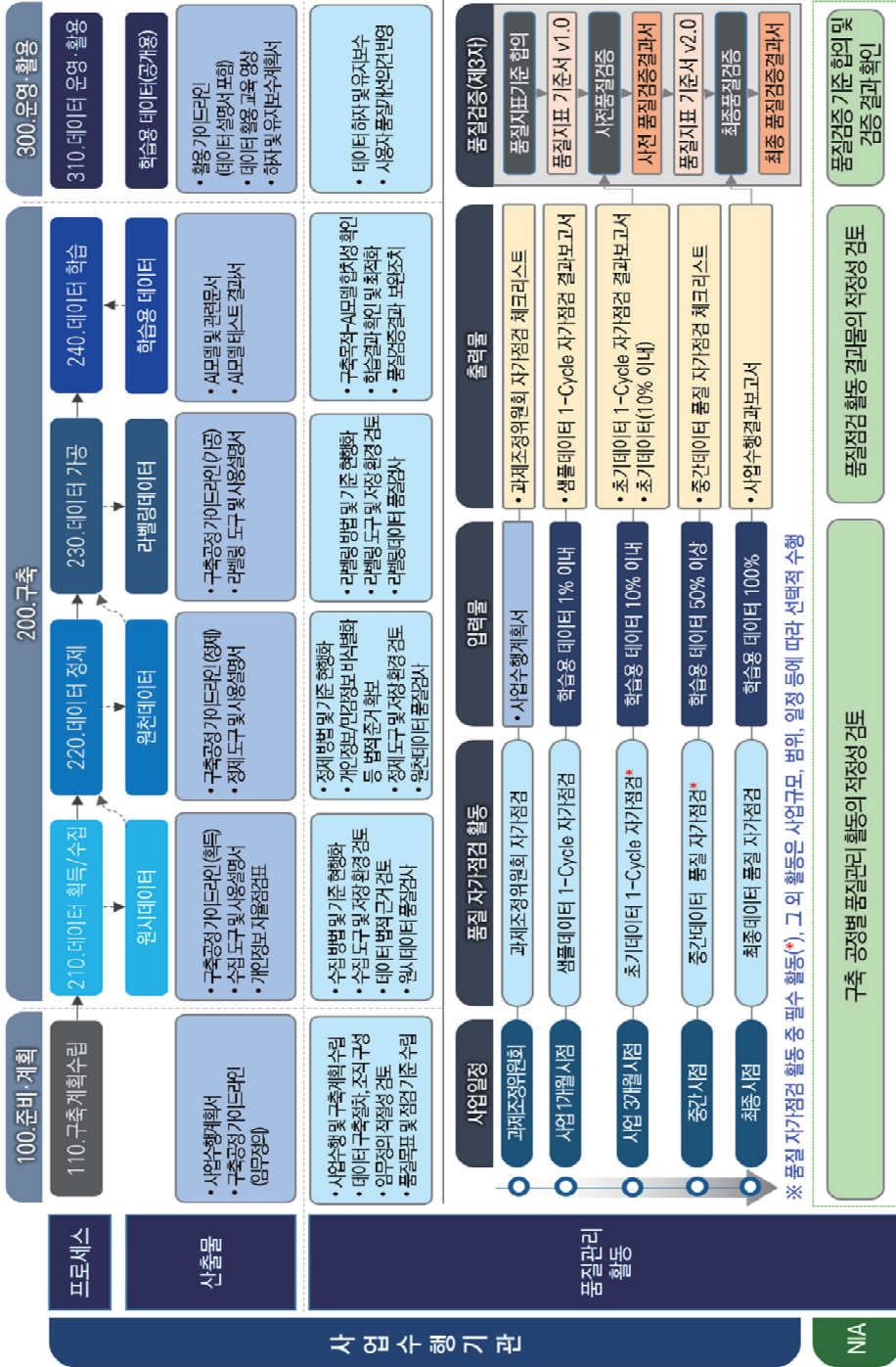
- 인공지능 학습용 데이터의 품질확보를 위한 품질관리 프레임워크의 구성 요소는 다음과 같음



[그림 II-1] 품질관리 프레임워크 구성 요소

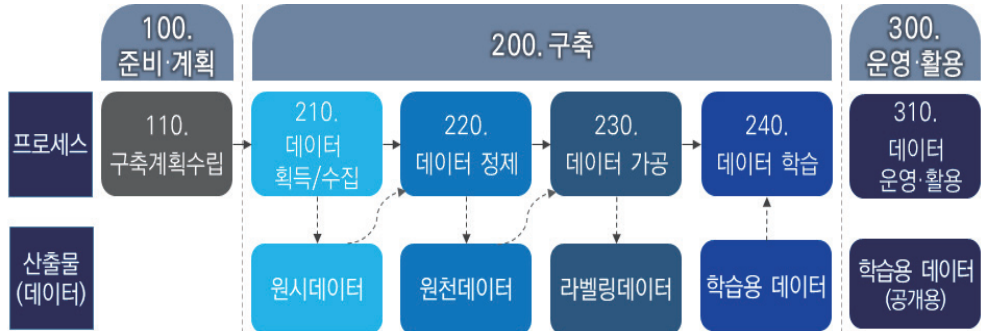
조직	구분	내용
-	단계	• 준비·계획, 구축, 운영·활용 3단계 사업 단계
사업수행기관	프로세스	• 준비·계획, 데이터 획득/수집, 데이터 정제, 데이터 가공, 데이터 학습 및 운영·활용 등 인공지능 학습용 데이터 계획, 구축, 운영의 사업 모든 단계와 관련된 프로세스
	산출물	• 사업수행 프로세스별 활동을 통해 생산되는 데이터 및 문서 등 산출물
	품질관리 활동	• 사업수행 단계별 품질관리 활동에 대한 정의 • 사업수행 일정별로 수행하는 자가점검 활동 • 제3자 품질검증 활동 정의
사업관리기관	품질관리 활동	• 사업발주 및 사업관리기관의 품질 관련 활동 정의

〈표 II-1〉 품질관리 프레임워크 구성 항목



[그림 II-2] 인공지능 학습용 데이터 품질관리 프레임워크

- 인공지능 학습용 데이터의 구축 및 운영단계는 준비·계획, 구축, 운영·활용 3단계로 구성되며, 각 단계별 프로세스 정의는 다음과 같음



[그림 II-3] 품질관리 프레임워크 프로세스 및 데이터 산출물

단계	프로세스	프로세스 정의
100. 준비·계획	110. 구축계획 수립	<ul style="list-style-type: none"> 구축목적의 일치성과 일관성 확보를 위한 구체적 계획 수립 사업수행기관의 컨소시엄 간 업무와 역할 정의 품질지표 및 목표, 품질점검 기준 수립 단계
200. 구축	210. 데이터 획득/수집	<ul style="list-style-type: none"> 기계학습에 필요한 데이터를 획득/수집하는 단계로 원시데이터 생성 데이터 획득/수집에 대한 프로세스 및 원시데이터에 대한 품질관리
	220. 데이터 정제	<ul style="list-style-type: none"> 데이터 획득/수집 프로세스로 생성된 원시데이터 중복제거, 비식별화 등 정제 작업을 진행하는 단계로 정제 작업을 수행한 원천데이터 생성 데이터 정제에 대한 프로세스 및 원천데이터 품질관리
	230. 데이터 가공	<ul style="list-style-type: none"> 정제 프로세스에서 생성된 원천데이터에 기능이나 목적에 부합되게 라벨링데이터를 부여하여 인공지능 학습할 수 있는 형태로 가공하는 단계 데이터 가공을 거쳐 라벨링데이터가 생성 데이터 가공에 대한 프로세스 및 라벨링데이터 품질관리
	240. 데이터 학습	<ul style="list-style-type: none"> 데이터 정제단계의 원천데이터와 데이터 가공단계의 라벨링데이터 묶음이 학습데이터셋으로, 학습데이터셋을 이용하여 학습시키고 학습된 인공지능 모델의 성능을 보장하는 단계 데이터 학습에 대한 프로세스 및 학습용 데이터 품질관리
300. 운영·활용	310. 데이터 운영·활용	<ul style="list-style-type: none"> AI-Hub에 적재되어 개방되는 학습용 데이터에 대한 품질관리 개방용 데이터 운영, 활용 지원을 위한 하자 및 유지보수 단계

〈표 II-2〉 품질관리 프레임워크 프로세스 정의

- 인공지능 학습용 데이터의 구축 및 운영단계별 프로세스 수행을 위해 필요한 품질관리 활동은 다음과 같음



[그림 II-4] 품질관리 프레임워크 프로세스별 품질관리 활동

단계	프로세스	품질관리 활동
100. 준비·계획	110. 구축계획 수립	1101. 사업수행 및 구축계획 수립
		1102. 데이터 구축 절차, 조직 구성
		1103. 임무정의에 대한 적절성 검토
		1104. 품질목표 및 점검기준 수립
200. 구축	210. 데이터 획득/수집	2101. 획득/수집 방법 및 기준 현행화
		2102. 데이터 법적 근거 검토
		2103. 획득/수집 도구 및 저장환경 검토
		2104. 원시데이터 품질검사
	220. 데이터 정제	2201. 정제 방법 및 기준 현행화
		2202. 개인정보/민감정보 비식별화 등 법적 준거 확보
		2203. 정제 도구 및 저장환경 검토
		2204. 원천데이터 품질검사
	230. 데이터 가공	2301. 라벨링 방법 및 기준 현행화
		2302. 라벨링 도구 및 저장환경 검토
		2303. 라벨링데이터 품질검사
	240. 데이터 학습	2401. 구축목적-SI모델 합치성 확인
		2402. 학습결과 확인 및 최적화
		2403. 품질검증결과 보완조치
300. 운영·활용	310. 데이터 운영·활용	3101. 데이터 하자 및 유지보수
		3102. 사용자 품질개선의견 반영

〈표 II-3〉 사업수행기관의 단계별 품질관리 활동

- 인공지능 학습용 데이터 구축사업 수행 일정별로 자가점검을 위한 활동에 대한 정의와 제3자 품질검증 활동은 다음과 같음



[그림 II-5] 품질 자가점검 및 품질검증 활동

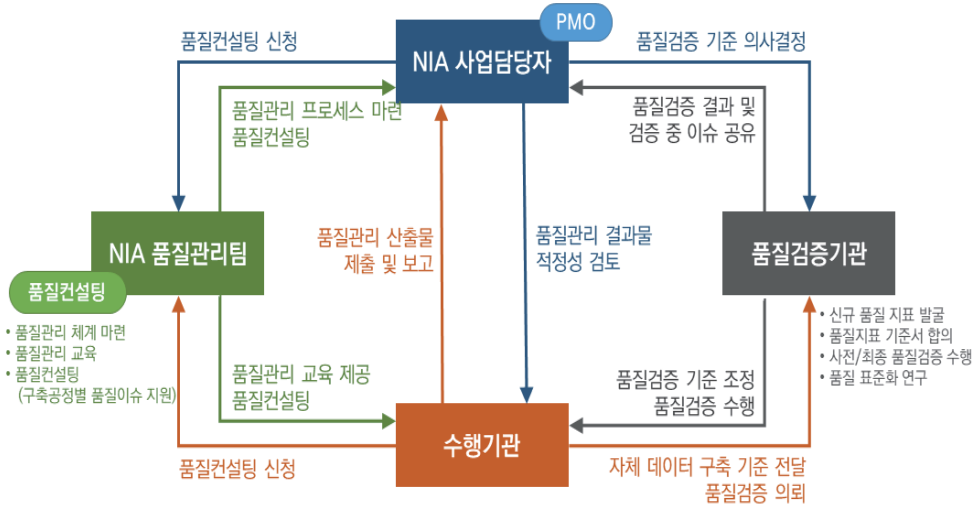
- 품질 자가점검 활동 중 초기데이터 1-Cycle 자가점검, 중간데이터 품질 자가점검은 필수적으로 수행하고, 그 외 품질 자가점검 활동은 구축사업의 규모나 범위, 일정 등에 따라 선택적으로 수행 가능

사업일정	자가점검 활동	활동 내용	품질검증(제3자)
과제조정위원회	과제조정위원회 자가점검	<ul style="list-style-type: none"> 과제조정위원회 개최 전 자가점검 체크리스트를 이용하여 사업수행계획서에 대한 자가점검 활동 우선협상대상자를 대상으로 과제조정위원회 개최 시 자가점검 체크리스트 확인 	사업수행계획서 및 과제조정위원회 결과를 기반으로 사업수행 기관과 품질지표에 대한 기준 합의
사업 1개월 시점	샘플데이터 1-Cycle 자가점검	<ul style="list-style-type: none"> 샘플데이터(1% 이내)를 대상으로 데이터 획득/수집, 정제, 가공, 학습까지의 전체 공정에 대한 1-Cycle 자가점검 활동 	초기데이터(10% 이상)를 대상으로 사전 품질검증 진행
사업 3개월 시점	초기데이터 1-Cycle 자가점검	<ul style="list-style-type: none"> 초기데이터(10% 이내)를 대상으로 데이터 획득/수집, 정제, 가공, 학습까지의 전체 공정에 대한 1-Cycle 자가점검 활동 	-
중간 시점	중간데이터 품질 자가점검	<ul style="list-style-type: none"> 학습용 데이터 50% 이상을 대상으로 데이터 획득/수집, 정제, 가공, 학습까지의 전체 공정에 대한 품질점검 활동 	-
최종 시점	최종데이터 품질 자가점검	<ul style="list-style-type: none"> 사업 최종 시점에 전체 학습용 데이터를 대상으로 전체 공정에 대한 품질점검 활동 	최종 데이터에 대한 제3자 품질검증 진행

〈표 II-4〉 사업일정에 따른 자가점검 활동 및 품질검증

2 품질관리 이해관계자 정의

- 인공지능 학습용 데이터 구축사업의 조직구조는 사업발주기관인 사업관리기관, 사업수행을 위한 사업수행기관과 제3자 품질검증을 위한 품질검증기관으로 구성되어 있음



[그림 II-6] 품질관련 이해관계자 관계

이해관계자	조직	내용
사업관리기관	사업담당자 (PMO)	<ul style="list-style-type: none"> 사업담당자 하부에 PMO 조직을 두어 사업자 선정, 과제조정, 협약체결, 사업수행관리, 최종 산출물 관리 등 사업 총괄 관리
	품질관리팀 (품질컨설팅)	<ul style="list-style-type: none"> 품질관리팀에서는 품질컨설팅 조직을 하부에 두어 품질컨설팅과 품질교육 등을 담당하여 사업수행기관의 품질관리 역량강화 및 품질관련 이슈사항을 해결하고, 사업관리기관의 사업담당자의 품질관리 지원
사업수행기관	품질관리조직	<ul style="list-style-type: none"> 준비·계획 단계에서 품질조직을 구성하여 구축단계별 품질관리 활동을 수행하여 관련 산출물 제출
품질검증기관 (제3자 품질검증)	-	<ul style="list-style-type: none"> 사업수행기관과 품질지표 기준을 합의하고, 합의된 지표를 기준으로 품질검증을 실시하여 품질검증 결과 확인

〈표 II-5〉 품질관련 이해관계자 조직 정의

- 사업관리기관, 사업수행기관, 품질검증기관(제3자)의 품질 관련 상세 활동 및 역할은 다음과 같음

품질검증기관(제3자)	사업수행기관	사업관리기관	
		사업담당자	품질관리팀
<p>〈품질지표 기준마련 및 품질검증〉</p> <ul style="list-style-type: none"> • 품질지표 기준 마련 및 기준서 합의 • 초기데이터 사전검증/최종 데이터 검증 수행 • 사전 및 최종 품질검증결과서 발급 • 미달성 과제 대상 보완조치 결과 확인 <p>〈품질 표준화 연구〉</p> <ul style="list-style-type: none"> • 신규 데이터 유형의 품질지표 선제 연구 • 다양한 산업분야에 따른 데이터 규격 마련 • 데이터 품질관리 표준 제정 	<p>〈준비·계획〉</p> <ul style="list-style-type: none"> • 데이터 구축계획설계 • 품질관리 조직 구성 • 데이터 법적 근거 및 비식별화 방안 수립 • 임무정의에 따른 품질목표 및 지표 설정 • AI모델 설계 <p>〈구축〉</p> <ul style="list-style-type: none"> • 구축공정 세부사항 수립 및 적용 • 데이터 법적 근거 및 비식별화 조치 • 구축공정별 저작도구 및 저장환경 확보 • 구축공정별 산출물 작성 및 현행화 • 구축목적에 따른 AI모델 최적화 <p>〈운영·활용〉</p> <ul style="list-style-type: none"> • 개방 전/후 보완조치 - 품질검증 결과, 사용자 품질개선요건 등 <p>〈자가점검 계획〉</p> <ul style="list-style-type: none"> • 과제조정위원회 자가점검 • 샘플데이터 1-Cycle 자가점검 <p>〈자가점검 실행〉</p> <ul style="list-style-type: none"> • 초기데이터 1-Cycle 자가점검 • 중간데이터 품질 자가점검 • 최종데이터 품질 자가점검 	<p>인공지능 학습용 데이터 생애주기별 품질관리 활동의 적정성 검토</p> <p>PMO</p> <p>각 품질점검 활동 결과물의 적정성 검토</p>	<p>〈품질관리체계 마련〉</p> <ul style="list-style-type: none"> • 품질관리 프로세스/프레임워크 정립 • 품질관리교육 시행 - 품질관리 가이드라인 - 품질지표 설정 가이드 - 산출물 작성 가이드(구축안내서) <p>〈품질관리 교육 및 품질컨설팅〉</p> <ul style="list-style-type: none"> • 구축공정 및 자가점검 과정의 품질자문 지원 - 품질자문지원단 구성 및 운영 - 중간점검 등 현장점검 지원 - 품질관리 헬프데스크 운영 <p>품질컨설팅</p>

〈표 II-6〉 품질관련 이해관계자 역할 및 책임

3 품질관리 조직체계

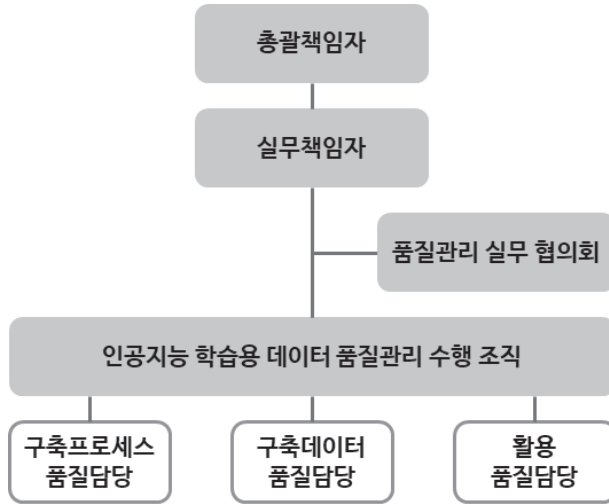
- 사업수행기관은 구축되는 인공지능 학습용 데이터의 실질적 품질관리 실무를 담당하는 ‘품질관리 수행조직’ 구성
- 구성되는 품질관리 수행조직은 작업자, 모델설계자, 외부기관 또는 전문가 등으로 구성하며 각 수행조직의 역할과 책임을 공식화하고, 품질관리 업무가 원활히 수행되는지 ‘총괄책임자’ 또는 ‘실무책임자’ 주도하에 수시로 확인 및 관리

3.1 품질관리 조직체계 구성 시 고려사항

- ① 사업수행기관(주관기관, 참여기관)의 유기적인 관계를 지닌 조직 구성
- ② 명확한 업무분장을 통한 책임 소재를 파악할 수 있는 업무수행조직 구성
- ③ 사업수행기관(주관기관, 참여기관) 간의 보고체계 구축
- ④ 데이터 획득/수집 기관, 정제 및 가공(라벨링) 기관, 데이터 학습모델 기관 등이 상이할 경우, 각 공정별로 참여 조직 구성
- ⑤ 특히 인공지능 학습용 데이터 학습모델을 담당하는 기관의 참여는 필수
- ⑥ 구축 후 데이터 운영·활용 체제로의 원활한 전환
- ⑦ 인공지능 학습용 데이터의 품질 확보

3.2 품질관리 조직의 주요 업무 내용

- ① 품질관리 계획의 수립 및 보완
- ② 품질관리를 위한 협의체 구성 및 운영
- ③ 구축 프로세스의 품질검사 및 개선
- ④ 인공지능 학습용 데이터의 품질검사 및 개선
- ⑤ 사업수행기관의 품질관리 역량 강화를 위한 품질관리 교육 실시
- ⑥ 데이터 운영·활용단계에서 발견된 품질 개선의견 조치



[그림 II-7] 구축사업 사업수행기관의 품질관리 조직 구성(안)

● 조직별 역할

조직	역할
총괄책임자	<ul style="list-style-type: none"> 해당 인공지능 학습용 데이터 구축사업의 공모과제 책임자가 되며, 구축사업 착수 이전에 데이터 품질관리 실무를 총괄하는 '실무책임자' 지정
실무책임자	<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축사업의 품질관리 실무를 총괄하며, 타 업무와 겹치지 않고, 품질관리 업무만을 전담하여 수행 구축하는 인공지능 학습용 데이터의 품질관리 계획을 수립하고, 품질관리 실무 수행조직을 운영하여 품질관리 활동을 수행하고, 품질관리 실무와 관련된 협의체 구성 및 운영 담당
품질관리 실무 협의회	<ul style="list-style-type: none"> 인공지능 학습용 데이터의 품질관리 관련 실무차원의 주요사항을 검토하고, 조직간 협의사항 조정, 조율 등을 위한 '품질관리 실무 협의회'를 설치·운영 사업수행기관의 품질관리 계획의 적정성 등을 평가하여 보완하고, 구축과정에서 발견된 데이터 품질 이슈를 논의하여 해결방안을 제시하는 등 상호 간의 의사소통 및 품질관리 실무를 협의하는 역할 담당 구축공정별 실무 담당자, 인공지능 모델설계자, 외부 품질전문가 등 참여
구축 프로세스 품질담당	<ul style="list-style-type: none"> 데이터 획득/수집, 정제, 가공 등 구축과정별로 구분하여 담당자를 지정할 수 있으며, 인공지능 학습용 데이터의 구축과정이나 구축 규모, 구축 난이도 등을 고려해서 담당자 지정 품질관리계획에 명시된 품질관리 활동의 준수 여부를 확인하고, 품질 이슈 발생 시 이를 기록하고 개선하는 활동을 수행하며, 즉각적인 조치가 불가능하거나 협의가 필요한 사항은 실무책임자에게 보고하는 등 역할 수행

조직	역할
구축 데이터 품질담당	<ul style="list-style-type: none"> 구축과정에서 생성되는 '원시데이터', '원천데이터', '라벨링데이터'의 품질을 확보하기 위해 품질계획에 수립된 검사기준과 절차 등에 따라 품질을 검사하고, 개선하는 활동 수행
활용 품질담당	<ul style="list-style-type: none"> 구축 이후에 AI Hub를 통해 민간에 개방한 인공지능 학습용 데이터의 품질오류가 발견된 경우, 이를 전달받아 조치하는 역할 수행 '활용 품질담당'은 구축과정에서 '구축 데이터 품질담당'의 역할을 지원하며, 실제 구축과정에서의 품질관리 경험을 통해 활용 단계에서도 품질 활동을 지원할 수 있도록 '품질관리계획서' 및 사용자에게 제공되는 '활용 가이드라인'에 담당자 명시

〈표 II-7〉 조직별 역할 정의

● 품질 이슈 해결 방안

- ① 품질관리 수행조직(구축공정별 실무 담당자, 인공지능 모델설계자, 외부 품질전문가 등)을 구성하여 검사 실시
- ② 객관성 확보를 위해 외부품질기관에 병행 처리하여 검사 실시
- ③ 품질검사 관련 총괄책임자, 실무책임자 주도하에 참여기관에서의 품질검사 수시 확인 및 검사
- ④ 활용(AI-Hub를 통해 개방)을 수행하는 인공지능 학습용 데이터의 품질을 담당하는 조직을 통한 최종 결과 보고 및 반영 등

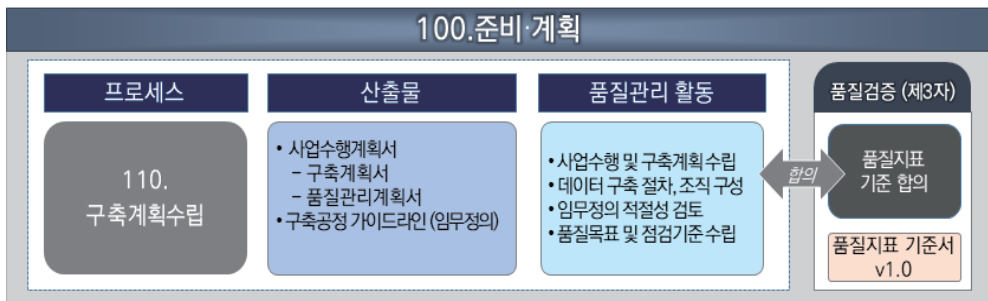
제2장 | 품질관리 프로세스 및 산출물

1 준비·계획단계(100)

1.1 구축계획 수립(110)

- 인공지능 학습용 데이터 유형별로 데이터 구축에 필요한 절차 및 구성 요소를 제시하여 데이터 구축과정에서의 시행착오를 줄이고 체계적인 계획을 수립하는 단계
- 구축계획 수립단계에서는 자료 준비 및 사업수행계획서 작성 등 사업 수행을 위한 계획과 학습용 데이터를 목적 등에 맞게 구체적으로 설계하는 단계
- 구축목적의 일치성과 일관성 확보를 위한 구축계획문서를 작성하여 사업 종료 시까지 변경 및 현행화 필요

1.1.1 구축계획 수립(110) 프로세스



[그림 II-8] 구축계획 수립 프로세스

1.1.2 구축계획 수립(110) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 내용	구축 및 품질 산출물
110. 구축계획 수립	1101. 사업수행 및 구축계획 수립	<ul style="list-style-type: none"> 구축목적 및 범위, 수행계획 수립 검토 품질관리방안 검토 품질목표수준의 적정성 및 품질관리기준 정의 검토 	사업수행계획서 - 구축계획서 - 품질관리계획서
	1102. 데이터 구축 절차, 조직 구성	<ul style="list-style-type: none"> 데이터 명세 정보 및 데이터 구성 정의 데이터 구축을 위한 절차, 조직, 도구 등 역할과 책임 수립 검토 데이터셋 개요, 구성내역, 설계기준 및 분포 현황, 구조, 구축 과정, 검수 및 품질활동 내역 검토 	
	1103. 임무정의에 대한 적절성 검토	<ul style="list-style-type: none"> 구축목적의 일치성과 일관성 확보를 위한 계획 검토 	구축공정 가이드라인 (임무정의)
	1104. 품질목표 및 점검기준 수립	<ul style="list-style-type: none"> 사업수행기관이 제출한 사업수행계획서 등 검증 관련 자료를 토대로 품질지표기준을 수립하고, 품질검증 지표 및 목표에 대해 합의 (사업수행기관 ↔ 제3자 품질검증기관, 사업관리기관) 	품질지표 기준서

〈표 II-8〉 구축계획 수립 단계 품질관리 활동 및 산출물

1.1.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

산출물명	산출물 정의	제출 여부	유형	시기
사업수행계획서 - 구축계획서 - 품질관리계획서	<ul style="list-style-type: none"> 사업수행계획서 : 사업 공모 시 제출한 사업수행계획에 대한 공식 산출물 붙임 1) 구축계획서 : 데이터 구축공정, 구축 데이터 정의, 데이터 검사 등 붙임 2) 품질관리계획서 : 품질지표 및 기준, 검사절차, 품질교육 등 과제조정위원회 검토의견, 품질지표 합의 등에 따라 구축계획서 및 품질관리계획서 수정 	Y	문서	과조위 전
구축공정 가이드라인 (임무정의)	<ul style="list-style-type: none"> 구축목적의 일치성과 일관성 확보를 위해 구축계획 단계에서 임무정의 및 공정별 가이드라인 초안 작성 구축공정별로 산출물 현행화 및 변경 관리 	Y	문서	협약 후

산출물명	산출물 정의	제출 여부	유형	시기
품질지표 기준서	<ul style="list-style-type: none"> 품질검증을 위한 필수 산출물로 협약 후 품질검증기관과의 합의에 의한 산출물 최종 제출 시는 사업수행계획서의 품질관리계획서에 붙임으로 제출 	Y	문서	협약 후

〈표 Ⅱ-9〉 구축계획 수립 단계 산출물 정의

1.1.4 고려사항

● 구축계획 수립 단계의 고려사항

항목	고려사항
목적성 확보	<ul style="list-style-type: none"> 데이터 구축 시 목적에 대한 명확한 정의를 위한 절차
법·제도 준수	<ul style="list-style-type: none"> 원시데이터 획득/수집 시 관련 법·제도적 규정 등을 반드시 준수하여야 함
구축량 확인	<ul style="list-style-type: none"> 데이터 구축목표 수량과 방법에 대한 검토
사실적 획득/수집 환경 구성	<ul style="list-style-type: none"> 원시데이터를 인위적인 환경이나 조건에서 획득/수집해야 하는 경우, 사실적인 획득/수집 환경을 현실적 가능성에 맞게 계획 수립

〈표 Ⅱ-10〉 구축계획 수립 단계의 고려사항

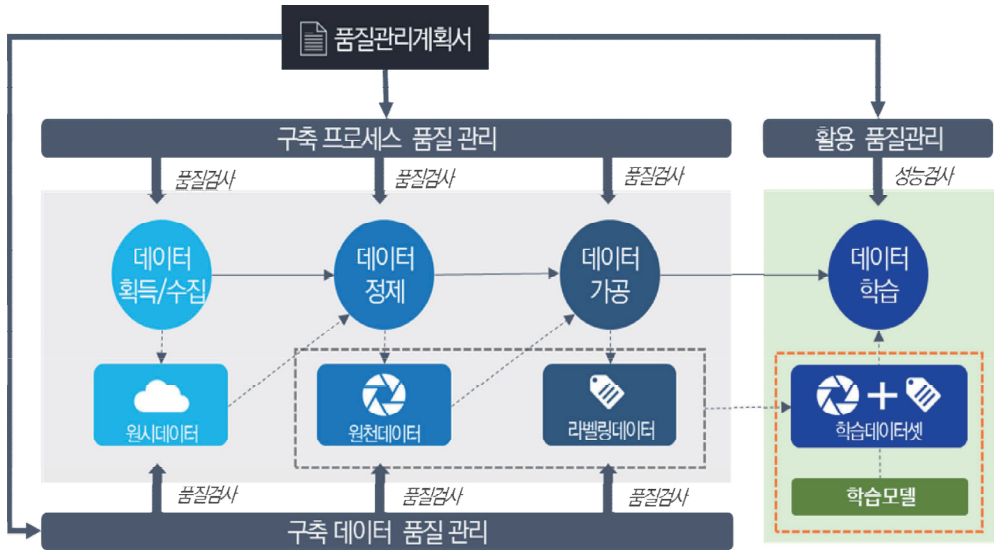
● 품질관리계획서 작성 시 고려사항

품질관리계획서 작성 시 고려사항
<ul style="list-style-type: none"> 데이터의 품질 목표 및 근거 품질관리 조직 구성 및 역할과 책임 데이터의 품질검사 절차 데이터 구축단계별 품질검사의 기준과 방법 데이터 품질관리를 위한 협의체 구성 및 운영 데이터 품질관리를 위한 참여기관의 품질관리 교육계획 수립 및 실시 데이터 품질검사에 따른 개선 조치 및 사후관리 방안 그 밖의 데이터 품질관리에 필요하다고 요청한 사항

〈표 Ⅱ-11〉 품질관리계획서 작성 시 고려사항

- 수립된 품질관리 계획은 구축사업을 추진하는 과정에서 데이터의 품질에 영향을 미치는 상황이 발생하였거나, 사업관리기관이 ‘품질관리계획서’의 보완을 요청한 경우, 품질관리협의회 등을 통해 ‘품질관리계획서’의 적정성 등을 종합적으로 검토하여 수정

- 계획단계에서 작성한 ‘품질관리계획서’를 기반으로 구축단계 품질관리에서 사업수행기관의 품질 관리 조직이 자체적인 품질검사 활동을 구축 프로세스 측면, 구축 데이터 측면, 활용 측면에서 수행하는 체계로 되어 있음



[그림 II-9] 인공지능 학습용 데이터 구축사업 품질관리 체계

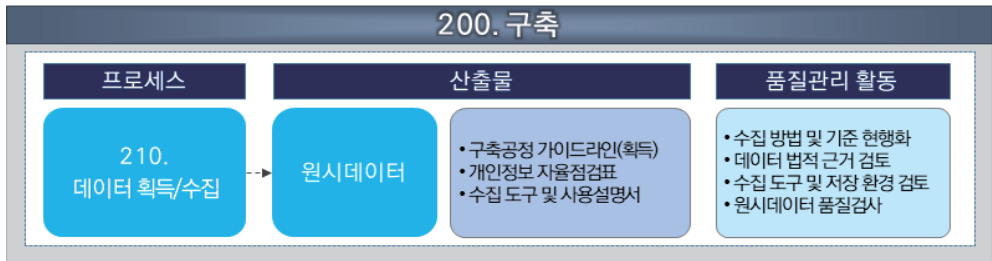
2 구축단계(200)

2.1 데이터 획득/수집(210)

- 인공지능의 기계학습에 필요한 데이터를 현실 세계에서 직접 획득/수집 또는 생성하거나, 이미 보유하고 있는 조직이나 시스템 등으로부터 법률적 제약이 없도록 목적에 따른 원시데이터를 확보하는 활동
- 인공지능 학습용 데이터의 전체 품질수준은 구축단계에서 80~90%가 좌우될 만큼 중요
 - 이유 1 : 인공지능 학습용 데이터가 인공지능 학습에 필요한 데이터 요구사항을 기반으로 구축되는 목적성을 갖기 때문
 - 이유 2 : 한번 구축된 인공지능 학습용 데이터는 자체적인 품질 이슈가 발견되지 않는 한 민간에 데이터가 그대로 개방되기 때문

- 이 단계에서는 정제 및 가공단계에서의 요구사항을 우선 파악 후 라벨링의 목적에 맞는 데이터 환경 구축 중요
- 구축단계의 품질관리는 인공지능 학습용 데이터 자체의 품질확보와 함께 인공지능 학습용 데이터를 생성하는 구축과정 단계별 프로세스 측면의 품질관리 필요

2.1.1 데이터 획득/수집(210) 프로세스



[그림 II-11] 데이터 획득/수집 프로세스

2.1.2 데이터 획득/수집(210) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 검토 내용	구축 및 품질 산출물
210. 데이터 획득/수집	2101. 획득/수집 방법 및 기준 현행화	<ul style="list-style-type: none"> 데이터 획득 및 확보 방법, 절차, 법 제도적 적절성 획득항목 및 구축량 획득/수집 단계 품질 기준 및 조치, 절차, 도구 데이터 획득 방법 및 기준, 교육, 검사에 대한 체계 수립 데이터 획득 기준 변경 관리 	구축과정 가이드라인 (획득/수집)
	2102. 데이터 법적 근거 검토	<ul style="list-style-type: none"> 법 제도 준수(개인정보, 초상권, 저작권) 근거자료 확보 데이터 획득/수집의 편향성 방지 방안 	개인정보 자율점검표
	2103. 수집 도구 및 저장환경 검토	<ul style="list-style-type: none"> 획득/수집에 사용될 도구 확보 및 사용 방법 획득/수집 데이터 저장환경 구축 	수집 도구 수집 도구 사용설명서
	2104. 원시데이터 품질검사	<ul style="list-style-type: none"> 원시데이터가 인공지능을 활용하여 문제를 해결하는데 요구되는 다양성, 신뢰성, 충분성, 균일성, 사실성, 공정성 등 기준에 적합하게 구성되어 있는지 확인 파일포맷, 해상도, 프레임 레이트 등 기술적합성 확인 데이터에 대한 클래스 분포, 인스턴스 분포, 분석 가능 항목의 통계적 분포 확인 	원시데이터 품질관리결과서 품질검사 도구

〈표 II-12〉 데이터 획득/수집단계 품질관리 활동 및 산출물

2.1.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

산출물명	산출물 정의	제출 여부	유형
구축공정 가이드라인 (획득/수집)	• 데이터 획득/수집의 대상, 절차, 기준, 조직 및 획득 방법과 도구, 검사 등 정의하는 가이드라인(지침) 제시	Y	문서
개인정보 자율점검표	• 개인정보보호법, 저작권, 지식재산권, 초상권 등을 고려한 데이터 획득/수집이 이루어지는지에 대한 점검표	Y	문서
수집 도구 및 사용설명서	• 데이터 획득/수집을 위해 필요한 장비 또는 수집 도구, 예를 들어 카메라, 비디오, 마이크, 크롤링 도구, 센서, 수집 앱 등 • 수집 도구에 대한 사용설명서 및 저장환경 등	C	문서/ 기타
원시데이터	• 획득/수집단계에서 생성된 데이터	C	데이터
품질관리결과서	• 품질검사 기준, 조직, 절차, 검사결과 등 작성	Y	문서
품질검사 도구	• 획득된 데이터의 적합성(기준 적합성, 기술 적합성, 통계적 다양성)을 중점으로 검사하여 결과를 확인 할 수 있는 도구 정의 • 품질검사 도구는 Application 또는 점검표(육안 검사결과 표 등) 등을 정의 할 수 있음	C	Application

〈표 II-13〉 데이터 획득/수집단계 산출물 정의

2.1.4 고려사항

- 데이터 획득/수집단계 품질확보를 위한 고려사항

항목	고려사항
법·제도 준수	<ul style="list-style-type: none"> • 원시데이터 획득 시 관련 법·제도적 규정 등을 반드시 준수하여야 함 • 개인정보 및 사생활 보호가 필요한 항목 획득 시, 개인정보보호법 등에 따라 적절한 법적, 기술적 절차를 거친 데이터를 활용하며, 그렇지 않은 데이터는 정제 과정에서 처리될 수 있도록 함 • 의료 데이터와 같은 경우는 IRB(의학연구윤리심의위원회)와 데이터 공개에 대한 해당 기관의 동의를 사전에 득해야 함 • 지적재산권 이슈가 있는 경우 반드시 해결방안을 마련하여 획득해야 함
데이터 다양성 확보	<ul style="list-style-type: none"> • 인공지능 학습모델이 현실을 잘 반영하고 본래의 구축목적에 달성할 수 있도록, 획득하는 데이터가 일부 범주에만 치우치지 않고 가능한 다양한 시간, 공간, 집단 수준 등이 포함될 수 있도록 함

항목	고려사항
데이터 편향 방지 및 윤리 준수	• 인공지능 학습모델이 인간의 비윤리 또는 편견을 학습하지 않고 사회적 윤리를 준수할 수 있도록 비윤리적 내용, 편견, 편향된 데이터의 획득/수집은 지양
사업수행계획서 및 데이터 구축요건 일치	• 사업수행계획서에 정의한 데이터 구축 기준에 맞추어 데이터를 획득/수집하도록 모니터링 및 검사
사실적 획득/수집 환경 구성	• 원시데이터를 인위적인 환경이나 조건에서 획득해야 하는 경우, 실제 환경이나 상황적 특성을 최대한 반영한 획득/수집 환경 구성
데이터 동기화	• 다중 속성의 데이터 소스 간 정교한 동기화를 위한 절차 마련

〈표 II-14〉 데이터 획득/수집단계 고려사항

- 원시데이터를 수집하거나 직접 원시데이터를 생성하는 경우, 품질 특성을 고려해서 원시데이터의 품질을 확보해야 함

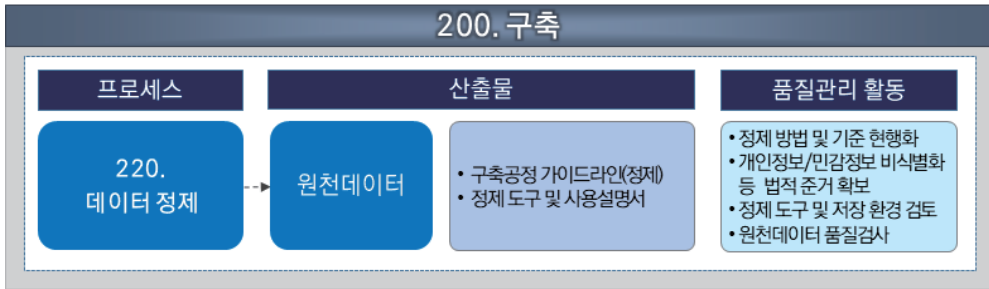
특 성	내 용
다양성	<ul style="list-style-type: none"> • 학습목적에 부합하도록 실제 세상의 데이터와 유사한 특성을 가진 데이터를 확보해야 함 <ul style="list-style-type: none"> - 사물, 사람, 장소, 시간, 환경, 언어 특성 등 학습에 유용한 모든 특성 정보를 포함할 수 있도록 고려 - 사물, 사람, 장소, 시간, 환경, 언어 특성 등의 특성 정보가 학습에 유용한 범위에서 다양하게 수집
신뢰성	• 데이터는 반드시 신뢰할 수 있는 출처로부터 수집해야 함
충분성	• 데이터에 포함된 카테고리화 인스턴스 등 특성정보는 학습에 유용한 수량이어야 함
균일성	• 분류/탐지/인식/이해/예측 등의 카테고리별 인스턴스 수량의 균일성과 적정 비율을 확보해야 함
사실성	• 원시데이터를 인위적인 환경과 조건하에 수집해야 하는 경우, 반드시 실제 환경과 상황 특성을 반영해야 함
공평성	• 원시데이터는 지역, 사회 및 인종적 편견 등 활용 의도와 무관한 편향적인 특성이 제거되고, 윤리적으로 공평해야 함

〈표 II-15〉 원시데이터의 품질 특성

2.2 데이터 정제(220)

- 획득한 원시데이터를 기계학습에 필요한 형식으로 맞추거나 불필요한 중복을 제거하며, 개인정보를 비식별화하여 처리하는 등 일련의 전처리 과정을 통해 ‘원천데이터’를 확보하는 활동
- 원시데이터 획득 시 정제기준을 명확하게 정의하고, 구축목적에 따라 적절한 데이터 선별을 위한 명확한 정제기준을 수립하며, 기준 미달 또는 활용 불가능한 데이터의 효과적 제거 방법 수립

2.2.1 데이터 정제(220) 프로세스



[그림 II-12] 데이터 정제 프로세스

2.2.2 데이터 정제(220) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 검토 내용	구축 및 품질 산출물
220. 데이터 정제	2201. 정제방법 및 기준 현행화	<ul style="list-style-type: none"> 구축목적에 적절한 데이터 선별을 위한 정제기준 수립 정제단계 품질 기준 및 조직, 절차, 도구 데이터 정제 방법 및 정제 결과에 대한 교육 및 훈련 데이터 정제 결과에 대한 검사 절차와 기준 수립 데이터 정제 기준 변경 관리 	구축과정 가이드라인 (정제)
	2202. 개인정보/민감정보 비식별화 등 법적 준거 확보	<ul style="list-style-type: none"> 개인정보보호 등 비식별화 기준 및 방안 등 법적 준거 확보 	
	2203. 정제 도구 및 저장환경 검토	<ul style="list-style-type: none"> 정제단계에 사용될 저작도구 확보 및 사용방법 정제 데이터 저장환경 구축 	저작도구(정제) 사용설명서
	2204. 원천데이터 품질검사	<ul style="list-style-type: none"> 사용하기 적합한 데이터 여부 품질검사 원천데이터 품질검사를 위한 검사 도구 및 검사결과 품질검사결과에 따른 보완조치 	원천데이터 품질관리결과서 품질검사 도구

〈표 II-16〉 데이터 정제단계 품질관리 활동 및 산출물

2.2.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

산출물명	산출물 정의	제출 여부	유형
구축과정 가이드라인(정제)	<ul style="list-style-type: none"> 데이터 정제단계의 원천데이터 포맷, 데이터 정제 규모, 절차, 기준, 조직, 방법, 도구 등 정의하고 가이드라인 제시 	Y	문서
저작도구(정제) 및 사용설명서	<ul style="list-style-type: none"> 원시데이터 획득 시 중복제거, 개인정보 비식별화 등 정제 작업을 위한 도구 정제 도구에 대한 사용설명서 및 저장환경 등 	Y	문서/ Application

산출물명	산출물 정의	제출 여부	유형
원천데이터	<ul style="list-style-type: none"> 원시데이터를 라벨링 공정에 투입하기 위해 필요한 전처리 등 정제 작업을 수행한 데이터로 라벨링이 부여되지 않은 상태의 데이터 	Y	데이터
품질관리결과서	<ul style="list-style-type: none"> 품질검사 기준, 조직, 절차, 검사결과 등 작성 	Y	문서
품질검사 도구	<ul style="list-style-type: none"> 획득된 데이터의 적합성과 중복성을 검사하여 결과를 확인할 수 있는 도구 정의 품질검사 도구는 Application 또는 점검표(육안 검사결과 표 등) 등을 정의 할 수 있음 	C	Application

〈표 II-17〉 데이터 정제단계 산출물 정의

2.2.4 고려사항

- 정제단계 품질관리를 위한 고려사항

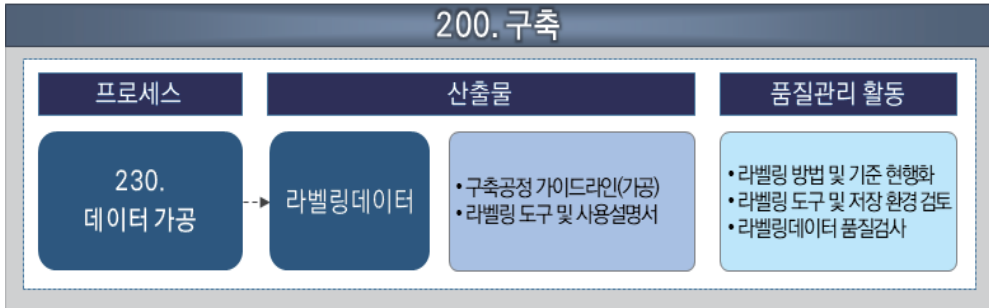
항목	고려사항
정제기준의 명확성	<ul style="list-style-type: none"> 데이터 사용 목적에 적합한 정제기준 수립 여부 검사
데이터 중복성 방지	<ul style="list-style-type: none"> 데이터 정제 후 정보 비교 후 중복도 여부 검사
정제 가이드라인	<ul style="list-style-type: none"> 정제작업을 위한 정제 가이드라인 작성 및 관리 여부 검사
정제 도구	<ul style="list-style-type: none"> 정제작업에 사용될 소프트웨어(S/W) 도구 확보 및 사용 방법 숙지
정제 작업 방식	<ul style="list-style-type: none"> 데이터 특성 및 활용목적에 맞는 적절한 정제 방식 선정 여부 및 선정 기준 타당성 여부 검사

〈표 II-18〉 데이터 정제단계 고려사항

2.3 데이터 가공(230)

- 인공지능이 기계학습에 활용할 수 있도록 기능이나 목적에 부합하는 정보를 원천데이터에 부착하는 활동
- 학습에 필요한 어노테이션 정보 설정을 데이터 확장을 고려하여 항목명, 타입, 필수 구분, 항목 설명을 포함해야 하고, 라벨링 용어 및 분류체계를 표준 준수하여 라벨링 방법에 따른 도구 사용 및 작업 방식에 의한 모든 산출 데이터는 표준화된 포맷을 지원할 수 있도록 CSV, JSON 등으로 묶어서 제공하며 이에 대한 품질검사 진행
- 가공단계의 품질관리를 위해서는 데이터 가공을 위한 작업 매뉴얼을 확보하고, 라벨링 도구를 확보하며, 해당 데이터의 특성을 고려해서 라벨링 작업 방식을 정해야 함

2.3.1 데이터 가공(230) 프로세스



[그림 II-13] 데이터 가공 프로세스

2.3.2 데이터 가공(230) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 검토 내용	구축 및 품질 산출물
230. 데이터 가공	2301. 라벨링 방법 및 기준 현행화	<ul style="list-style-type: none"> • 라벨링 방법 및 기준 정의 • 학습에 필요한 라벨링 작업 및 어노테이션 설정 • 데이터 형식 및 입력값 범위 기준 마련 • 데이터 유형별 객체 구분 작업 • 데이터 가공 방법 및 기준, 교육, 검사체계 수립 • 데이터 가공 기준변경 관리 	구축과정 가이드라인 (가공)
	2302. 라벨링 도구 및 저장환경 검토	<ul style="list-style-type: none"> • 가공에 사용될 저작도구 확보 및 사용 방법 • 데이터 가공 기준변경 관리 • 라벨링데이터 저장환경 구축 	저작도구(가공) 사용설명서
	2303. 라벨링데이터 품질검사	<ul style="list-style-type: none"> • 정확성(의미 정확성, 구문 정확성)을 중점으로 검사 • 라벨링데이터 품질검사를 위한 검사 도구 및 검사 결과 • 품질검사결과에 따른 보완 조치 	라벨링데이터 품질관리결과서 품질검사 도구

〈표 II-19〉 데이터 가공단계 품질관리 활동 및 산출물

2.3.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

산출물명	산출물 정의	제출 여부	유형
구축과정 가이드라인(가공)	<ul style="list-style-type: none"> • 데이터 가공단계의 품질지표, 어노테이션 포맷, 가공 대상, 절차, 기준, 방법, 도구, 검사 등 정의하는 가이드라인 제시 	Y	문서
저작도구(가공) 및 사용설명서	<ul style="list-style-type: none"> • 데이터 가공 저작도구는 데이터 학습 전 원천데이터에 특정 값을 부여해주는 도구 • 라벨링 도구에 대한 사용설명서 및 저장환경 등 	Y	문서/ Application

산출물명	산출물 정의	제출 여부	유형
라벨링데이터	• 원천데이터에 부여한 '참값', 파일형식이나 해상도 등의 속성, 설명이나 주석 등이 포함된 '어노테이션'의 집합	Y	데이터
품질관리결과서	• 품질검사 기준, 조직, 절차, 검사결과 등 작성	Y	문서
품질검사 도구	• 라벨링데이터의 정확성(구문 정확성, 의미 정확성)을 검사할 수 있는 도구 정의 • 품질검사 도구는 Application 또는 점검표(육안 검사결과 표 등) 등을 정의 할 수 있음 ※ 자동검사 후 육안검사 실시 시 사용하는 도구 모두 정의	C	Application

〈표 II-20〉 데이터 가공단계 산출물 정의

2.3.4 고려사항

- 가공단계의 품질관리를 위해서는 라벨링을 위한 작업 매뉴얼을 확보하고, 라벨링 도구를 확보하여, 해당 데이터의 특성을 고려해서 라벨링 작업 방식을 정해야 함
- 데이터 가공 작업에 필요한 라벨링 규격은 다음의 사항을 반드시 포함하도록 고려해야 함

항목	고려사항
라벨 정의	• 라벨 유형, 라벨 구성 항목, 속성값, 타입 등을 정의해야 함
구축공정 가이드라인 (가공)	<ul style="list-style-type: none"> • 목적에 맞게 작성된 구축공정 가이드라인(가공)에 대한 타당성 고려 • 참값(Ground Truth) 라벨링, 오류 라벨링, 모호한 라벨링 처리, 예외처리 라벨링 등을 사전에 정의하여 통과 및 반려 기준을 마련해야 함 • 목적에 맞는 어노테이션 항목 구성 여부 검사 후 작업자 전달 • 작업자가 효율적으로 작업할 수 있는 환경 구성
라벨링 도구	• 라벨링 작업에 사용하는 플랫폼 및 SW 도구의 기능 및 사용 방법을 기술해야 하며 플랫폼 또는 SW 도구 선택 시 데이터 보안 준수 여부를 확인해야 함

〈표 II-21〉 데이터 가공단계 고려사항

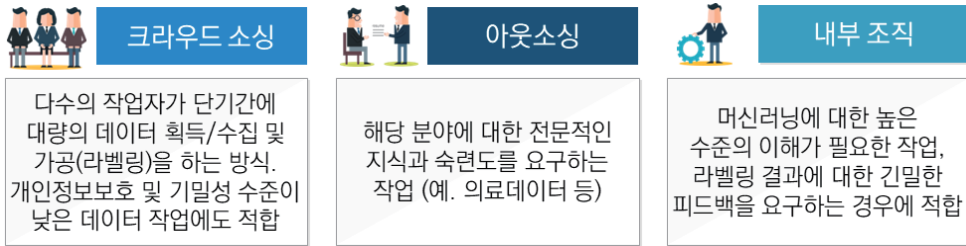
- 데이터 특성별 가공 시 품질 확보 측면 고려사항

특성	고려사항
이미지/ 비디오 데이터 가공	<ul style="list-style-type: none"> • 카테고리 라벨은 표준화된 분류체계(또는 MECE* 원칙을 준용하여 자체 정의한 분류체계)를 사용해야 함 * MECE(Mutually Exclusive Collectively Exhaustive)란 각 항목들이 상호 배타적이면서, 모였을 때는 전체를 이루는 상태 • 바운딩박스는 이미지 내에서 특정객체의 위치정보를 추출하는 것으로 대상 객체 전체를 감싸는 형태로 공백을 최소화해야 함 • 시멘틱 세그멘테이션(semantic segmentation)은 객체 영역을 픽셀 단위로 표시하는 작업으로 동일 카테고리의 객체가 중첩되어 있는 경우를 구별하지 않음

특성	고려사항
	<ul style="list-style-type: none"> 인스턴스 세그멘테이션(instance segmentation)은 각 객체 영역을 픽셀 단위로 구분 표시함
텍스트/ 오디오 데이터 가공	<ul style="list-style-type: none"> 오디오 전사의 경우 표준발성에서 벗어나거나 같은 전사에 대하여 두 가지 이상 발음이 가능한 경우 발음전사와 철자전사를 병행토록 하는 등 'ETRI 전사규칙'을 준수해야 함

〈표 II-22〉 라벨링데이터의 고려사항 (예시)

- 라벨링 작업은 데이터 특성 및 활용목적을 고려하여 클라우드소싱, 아웃소싱, 내부조직 활용 등 데이터 라벨링 작업자 구성 방식 선택
- 데이터 라벨링 작업자 구성 방식에 따른 특성은 다음과 같음



[그림 II-14] 데이터 특성 및 활용 목적을 고려한 데이터 라벨링 작업자 구성 방식

2.4 데이터 학습(240)

- 데이터 학습단계에서는 원천데이터와 라벨링데이터의 묶음인 ‘학습데이터셋’을 이용하여 사전에 정의된 인공지능 알고리즘을 학습시키고, 학습된 인공지능 모델의 성능을 향상시키거나 보정하는 활동 수행

2.4.1 데이터 학습(240) 프로세스



[그림 II-15] 데이터 학습 프로세스

2.4.2 데이터 학습(240) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 검토 내용	구축 및 품질 산출물
240. 데이터 학습	2401. 구축목적-AI 모델 합치성 확인	<ul style="list-style-type: none"> 데이터를 인공지능 알고리즘으로 훈련시켰을 때 영역별 품질검사 기준 확인 구축 목적에 부합하는 AI 모델 선정 확인 AI 모델 구축/운영 및 테스트를 위한 환경 구축 정보 AI 모델 정보/라이선스 가이드 	AI모델 설명서
	2402. 학습결과 확인 및 최적화	<ul style="list-style-type: none"> 학습용 데이터가 AI 모델로 훈련시켰을 때 목표로 했던 수준의 성능 달성이 가능한지 검사 알고리즘 성능 검사 모델 단계 유효성 검사 학습모델, 활용 서비스 개발에 가능한 학습데이터셋 품질오류 및 이슈 내용을 분석하여 보완조치 반영 	AI모델 소스코드 학습모델 파일 모델 구축 도커 이미지
	2403. 품질검증결과 보완조치	<ul style="list-style-type: none"> 최종데이터 제3자 품질검증결과가 미달성인 경우, 보완조치하여 결과 보고 ※ 3. 제3자 품질검증 > 3.2 제3자 품질검증 세부 활동 정의 > 3.2.3 최종품질검증 프로세스 부분 내용 참고	AI모델 테스트 결과서
			품질검사 도구
			보완조치결과서

〈표 II-23〉 데이터 학습단계 품질관리 활동 및 산출물

2.4.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

산출물명	산출물 정의	제출 여부	유형
AI모델 설명서	<ul style="list-style-type: none"> AI모델 구축/운영 및 테스트를 위한 환경 구축 내역 AI모델 소스코드 활용을 위한 환경 구축 및 실행 매뉴얼 모델 관련 정보 및 라이선스 정보 라이브러리, 패키지 버전 및 구축환경 정보 	Y	문서
AI모델 소스코드	<ul style="list-style-type: none"> 모델 학습 코드 모델 성능 테스트 코드 기타 	Y	Code
학습모델 파일	<ul style="list-style-type: none"> 특정 유형의 패턴을 인식하도록 학습된 파일로 학습가중치, 아키텍처 정보, 기타 학습요소 등 포함 	Y	binary
모델 구축 도커 이미지	<ul style="list-style-type: none"> AI모델 구축환경정보가 내장된 도커 파일을 모델 산출물 파일과 같이 제출 	Y	tar

산출물명	산출물 정의	제출 여부	유형
시모델 테스트 결과서	• 제3자 품질검증 전 사업수행기관의 자체적인 모델 테스트 수행하여 미흡한 경우 보완조치 필요	C	문서
품질검사 도구	• 학습용 데이터셋의 유효성 품질검사 도구	C	Application
보완조치결과서	• 제3자 최종검증 미달성 기관만 해당 산출물 제출 • 최종 품질검증 결과서의 미흡 내용을 대상으로 공정별 개선 후 품질검사결과 작성	Y	문서

〈표 II-24〉 데이터 학습단계 산출물 정의

2.4.4 고려사항

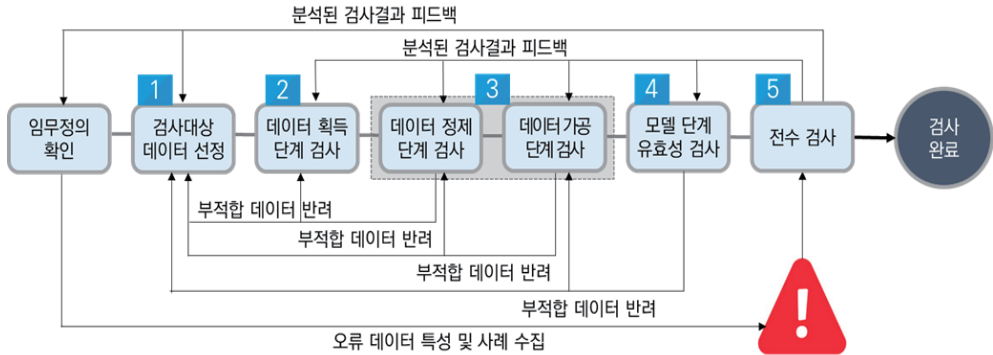
항목	고려사항
데이터 학습 유효성	<ul style="list-style-type: none"> • ‘학습데이터셋’을 이용하여 사전에 정의된 인공지능 알고리즘을 학습시키고, 학습된 인공지능 모델의 성능을 향상시키거나 보정하는 검사를 수행 • 구축한 데이터셋을 인공지능 학습모델로 훈련시킨 후 목표로 했던 수준의 성능 달성이 가능한지 확인 • 학습모델 평가를 위한 정확도, 정밀도, 재현율 등 성능지표 확인
시모델 성능	<ul style="list-style-type: none"> • 알고리즘 성능 검사
품질오류 분석	<ul style="list-style-type: none"> • 학습모델, 활용 서비스 개발에 가능한 학습데이터셋 품질오류 및 이슈 내용을 분석

〈표 II-25〉 데이터 학습단계 시 고려사항

2.5 구축단계 품질검사

2.5.1 검사 절차

- 구축단계 품질검사는 품질검사 절차 중 품질통제를 제외한 모든 영역을 구축공정별로 적용
- 현 구축공정의 품질검사 결과가 앞서 수행한 구축공정에도 영향을 미칠 수 있음
- 예를 들어, 데이터 정제단계에서의 다양성 관련 지표의 품질오류가 발견되면 데이터 획득/수집단계에서 적합성이 보장받지 못할 수 있음
- 이러한 경우 데이터 획득/수집단계부터 재수행하거나 보완이 필요하며, 각 공정의 품질검사 시 재작업을 최소화하기 위해 각 공정 진행 중에 지속적이고 반복적인 품질검사 수행 필요



[그림 II-16] 공정별 구축 데이터 검사 절차

① 검사대상 데이터 선정

구 분	내 용
검사절차명	• 검사대상 데이터 선정
검사지표명	• N/A
검사내용	<ul style="list-style-type: none"> • 각 공정의 품질기준(지표)에 따른 적절한 기준과 수량 선정 • 구축대상 데이터의 분류체계(카테고리)가 모두 포함되게 선정 • 초기데이터 구축단계에서부터 품질검사를 고려하여 각 분류체계의 데이터를 적절한 비율로 모두 획득/수집할 수 있도록 계획 수립
검사예시	• 약초의 식별 학습을 위해 인삼, 홍삼, 당귀, 참당귀, 느릅나무를 선정하여 데이터를 구축하는 경우 검사대상 데이터 선정 시 해당 데이터가 모두 포함

② 데이터 획득/수집 단계 검사

구 분	내 용
검사절차명	• 데이터 획득/수집 단계 검사
검사지표명	• 적합성(기준 적합성, 기술 적합성, 통계적 다양성)
검사내용	<ul style="list-style-type: none"> • 적합성(기준 적합성, 기술 적합성, 통계적 다양성)을 중점으로 검사 • 획득/수집단계에서의 검사결과에 따른 개선을 수행하기 위해서는 임무정의, 학습모델에 따른 데이터 구축 요구사항을 재정리하여 구축계획을 수정하고 데이터 획득/수집 재실시 • 검출한 오류사항은 그 특징과 사례를 정리하여 추후 검사 시 우선 체크하여 오류 최소화
검사예시	<ul style="list-style-type: none"> • 이미지 촬영 또는 오픈 이미지 사용에 따른 적합성 여부 확인(이미지 촬영, 다양성, 법적 문제 방안) • 품목 적합성 확인, 이미지 적합성 확인, 개인정보 침해 여부 확인

③ 데이터 정제, 가공 단계 검사

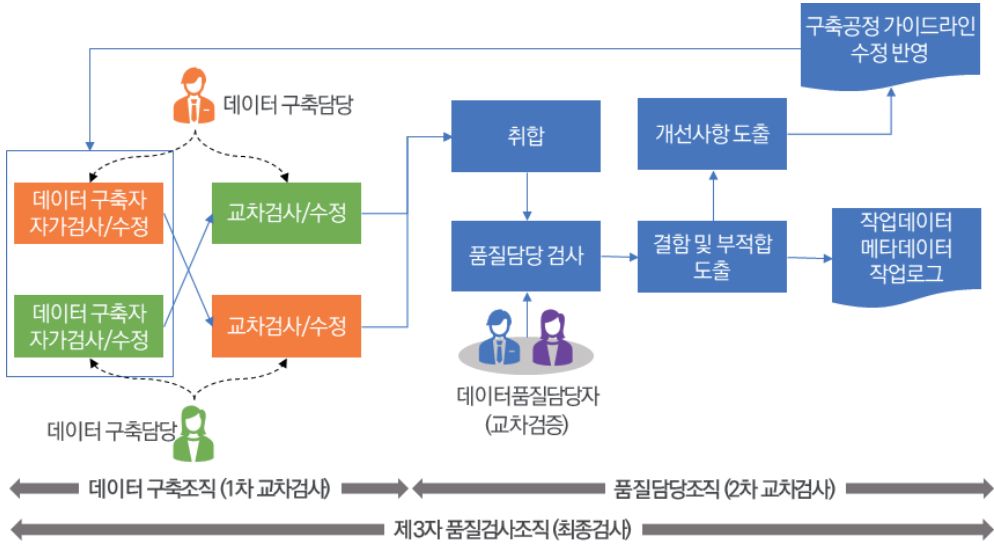
구 분	내 용
검사절차명	• 데이터 정제, 가공단계 검사
검사지표명	• 정확성(의미 정확성, 구문 정확성)
검사내용	<ul style="list-style-type: none"> 정확성(의미 정확성, 구문 정확성)을 중점으로 검사 인공지능 학습모델이 학습을 정확하게 수행하기 위한 필수적 요소이며, 구축공정의 특성상 작업자간 품질차이가 큼 품질관리담당의 정기 또는 수시 품질검사가 반복 수행되어야 함 본 단계의 검사에서 발생하는 오류는 현(現) 공정상의 오류일 수도 있으나, 전(前) 공정에서 발생한 오류가 현(現) 공정에 영향을 미치는 경우도 있음
검사예시	• 전(前) 공정에서의 영향을 분석하여 품질관리 또는 프로젝트 관리조직의 주관하에 공정 수행 계획 재수립 및 재실시를 통해 오류를 최소화할 수 있도록 해야 함

④ 모델 단계 검사

구 분	내 용
검사절차명	• 모델 단계 유효성 검사
검사지표명	• 유효성(전체 데이터를 사용하여 훈련/검증/데이터셋으로 분할하고 알고리즘 훈련 완료 후 테스트셋으로 성능 측정)
검사내용	<ul style="list-style-type: none"> 데이터를 인공지능 알고리즘으로 훈련시켰을 때 분류성능, 탐지성능, 인식성능, 질의응답, 기계번역, 문서요약, 이미지, 영상, 텍스트, 음성, 센서 등 영역별 품질검사 기준 확인
검사예시	<ul style="list-style-type: none"> 학습용 데이터가 AI 알고리즘으로 훈련시켰을 때 목표로 했던 수준의 성능 달성이 가능한지 검사 → 유효성 <ul style="list-style-type: none"> 평균 IoU, AUC, 평균 정밀도, 평균 재현율, Fβ-점수, 픽셀 정확도, 평균 정확도, 일치도, 문장 유사도 등 분류·탐지·인식 성능은 정밀도, 재현율, Fβ-점수 및 기타 변형 지표로 측정 ※ IoU(Intersection Over Union), AUC(Area Under the Curve), Fβ-점수는 Precision(정밀도)과 Recall(재현율)의 가중치 조화 평균

⑤ 전수 검사

구 분	내 용
검사절차명	• 데이터 전수 검사
검사지표명	• 데이터 정제 검사 기준을 준수하여 검수 진행
검사내용	<ul style="list-style-type: none"> 부적합 판정 데이터 분포 확인 외부 검수자, 도메인 전문가, 데이터 요청자에 의한 검사 지표 및 기준 통일성 검사
검사예시	<ul style="list-style-type: none"> 클라우드소싱 검사 공정 진행으로 데이터 정제 검사 기준을 준수하여 검사 진행 정확성 시험 : 어노테이션 포맷 정확도 정합률 99% 이상 <ul style="list-style-type: none"> 참값 정확도 <ol style="list-style-type: none"> 미태깅율 3% 이하, 라벨링 대상 객체에 대한 class 정보입력 및 바운딩/폴리곤 작업 누락 오태깅율 1% 이하, 라벨링 대상 객체에 대한 class 정보입력 오류



[그림 II-17] 공정단위 검사 과정

- (1차 교차검사) 각 공정에 관계없이 학습모델 설계 조직, 데이터 획득 조직, 데이터 정제 조직, 데이터 가공 등 내부 수행조직 내의 실제 작업자와 작업자 관리자 1차 검사 진행
- (2차 교차검사) 작업자 자가 검사, 작업자 간 교차 검수 및 품질담당자의 2차 교차 검사
- (최종 검사) 제3자 품질검사조직 등이 최종적인 검사 수행

2.5.2 검사기준

- 인공지능 학습용 데이터 구축사업의 특성에 따라 지표를 선택적으로 적용하거나 추가할 수 있으며, 아래 지표 및 기준은 인공지능 학습용 데이터 구축사업에 필수적으로 적용해야 하는 지표임

구분	검사기준 지표	내용
임무정의, 구축계획 수립 및 데이터 획득 단계	법·제도 준수	• 원시데이터 획득 시 관련 법·제도적 규정 등 반드시 준수
	사실적 획득 환경 구성	• 원시데이터를 인위적인 환경과 조건하에 획득해야 하는 경우 사실적인 획득 환경 구성
	데이터 동기화	• 다중 속성 데이터 소스 간 정교한 동기화를 위한 절차 마련
	편향성 방지	• 데이터 편향을 방지하기 위한 절차 마련
데이터	정제기준의 명확성	• 데이터 사용 목적에 적합한 정제기준 수립 여부 검수

구분	검사기준 지표	내용
정제단계	중복성 방지	• 데이터 정제 후 정보 비교 등을 통한 중복도 여부 검수
	정제 작업 매뉴얼	• 정제 작업을 위한 매뉴얼 작성 및 관리 여부 검수
	정제 도구	• 정제 작업에 사용될 소프트웨어(S/W) 도구를 확보 및 사용 방법 숙지
	정제 작업 방법	• 데이터 특성 및 활용 목적에 맞는 적절한 정제 방식 선정 여부 및 선정 기준 타당성 여부 검수
데이터 가공단계	구축공정 가이드라인(가공)	• 목적에 맞게 작성된 구축공정 가이드라인(가공)에 대한 타당성 여부 검사 후 데이터 라벨링 작업자들에게 해당 내용의 가이드라인 전달
	어노테이션 항목	• 목적에 맞는 어노테이션 구성 여부 검수 후 확인된 내용을 포함하도록 작업자들에게 전달
	품질검사 도구	• 자동화 도구를 통해 검사 후 검사자가 육안으로 부적합 데이터 여부에 대해 2차 확인 및 조건 오류 전수 검사
전수 검사	부적합 판정 데이터 분포 확인	• 데이터의 오류율, 특성 분포 확인을 통한 데이터 획득, 정제, 가공, 부문 최적화
	외부 검사자	• 제3자 품질검증기관 등 외부 검사자, 해당 분야 전문가, 데이터 요청자에 의한 검사를 위해 검사지표 및 기준이 통일되게 전달되었고, 검사 방법의 동의를 이루어졌는지 확인

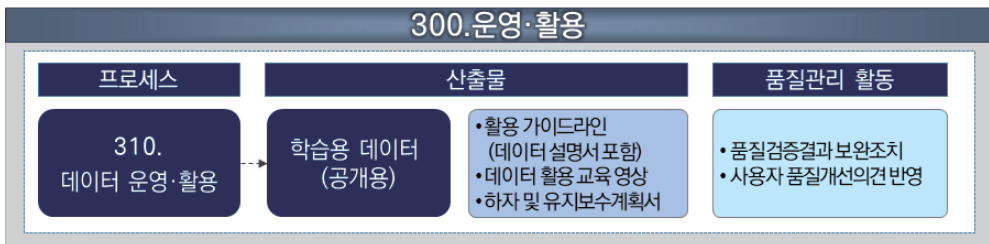
〈표 II-26〉 품질검사 기준지표

3 운영·활용단계(300)

3.1 데이터 운영·활용(310)

- 인공지능 학습용 데이터의 운영·활용 단계 품질관리는 데이터를 민간에 개방하기 위해 AI Hub에 적재된 학습데이터셋의 품질을 향상시키기 위한 제반 활동을 의미

3.1.1 데이터 운영·활용(310) 프로세스



[그림 II-19] 데이터 운영·활용 프로세스

3.1.2 데이터 운영·활용(310) 프로세스 품질관리 활동 및 산출물

프로세스	품질관리 활동	주요 검토 내용	구축 및 품질 산출물
310. 데이터 운영·활용	3101. 데이터 하자 및 유지보수	<ul style="list-style-type: none"> 시허브에 공개된 데이터에 대해 하자보수 및 유지보수 	하자 및 유지보수계획서
	3102. 사용자 품질개선의견 반영	<ul style="list-style-type: none"> 사용자들의 개발 및 활용 지원을 위한 산출물 공개 데이터 품질에 대한 하자 및 유지보수 계획 사용자들의 품질에 대한 개선 의견이 있는 경우 반영 등 유지보수 	활용 가이드라인 (데이터 설명서 포함) 데이터 활용 교육 영상

〈표 II-27〉 데이터 운영·활용 단계 품질관리 활동 및 산출물

3.1.3 산출물 정의

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

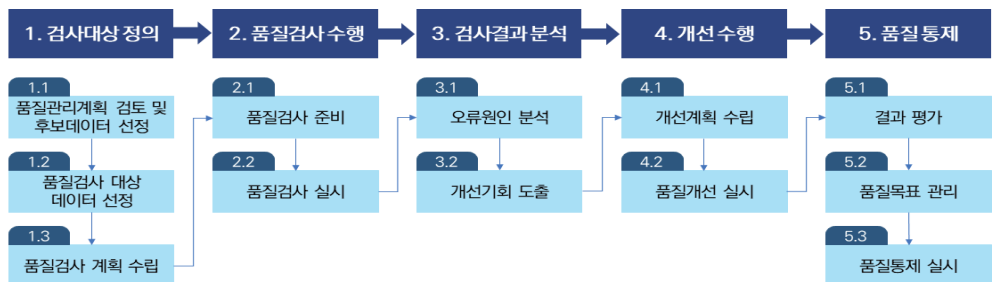
산출물명	산출물 정의	제출 여부	유형
활용 가이드라인 (데이터 설명서 포함)	<ul style="list-style-type: none"> 데이터 설명서 및 활용 가이드로 구성하여 누구나 사용가능하도록 AI-Hub에 개방 	Y	문서
데이터 활용 교육 영상	<ul style="list-style-type: none"> 데이터 1종당 1개의 데이터 활용 교육 영상 가이드 제작하여 AI-Hub에 개방 	Y	동영상
하자 및 유지보수계획서	<ul style="list-style-type: none"> 공개용 학습데이터셋에 대한 하자 및 유지보수계획 	Y	문서

〈표 II-28〉 데이터 운영·활용 단계 산출물 정의

3.2 운영·활용 단계 품질검사


3.2.1 품질검사 절차

- 인공지능 학습데이터 품질검사는 검사대상 정의(3), 품질검사 수행(2), 검사결과 분석(2), 개선수행(2), 품질 통제(3)의 5단계, 12개 상세 수행 절차로 구성



〔그림 II-20〕 운영·활용 단계 품질검사 상세 수행 절차

● 1. 검사대상 정의

수행 절차	1.1 품질관리계획 검토 및 후보데이터 선정
주요 활동	<ul style="list-style-type: none"> 품질관리 담당자는 구축사업의 품질관리계획과 관련된 자료를 수집하고 데이터 구축 및 인공지능 학습 모델 담당자와 협의를 통해 변경된 요구사항, 업무절차 및 작업방법 등의 부가정보 획득 <div>후보데이터 선정 시 고려사항</div> <ul style="list-style-type: none"> 라벨링 작업 그룹, 클라우드워크 플랫폼 등의 데이터 구축환경에서 접수되는 품질오류 및 개선의견 접수 내용 분석 학습 모델, 활용 서비스 개발 그룹으로부터 접수되는 품질 오류 및 개선의견 분석 인공지능 학습용 데이터 구축사업 관련 전문가 검토에 따라 정비가 필요한 데이터에 대한 분석 데이터의 특성상 품질 이슈 발생 확률이 높은 데이터에 대한 개선 수요 분석
수행 절차	1.2 품질검사 대상 데이터 선정
주요 활동	<ul style="list-style-type: none"> 품질관리 계획 검토 및 후보 데이터로 식별한 데이터 중에서 품질검사의 중요성, 시급성 관점으로 검사대상 데이터를 선정 <div>데이터 선정 기준</div> <div> <div>중요성 측면</div> <ul style="list-style-type: none"> 구축 공정별 데이터의 적합성, 정확성, 유효성이 확보되어야 하는 데이터 수집된 품질이슈를 해결하기 위해 재작업 또는 추가작업을 실시한 데이터 인공지능 학습모델이 동작하는데 핵심적인 기능을 담당하는 데이터 </div> <div> <div>시급성 측면</div> <ul style="list-style-type: none"> 품질 오류 및 이슈가 빈번하게 발생하는 데이터 시급히 해결해야 하는 품질 현안을 보유한 데이터 </div>
수행 절차	1.3 품질검사 계획 수립
주요 활동	<ul style="list-style-type: none"> 품질검사 대상 데이터 선정이 완료되면 선정한 데이터를 대상으로 품질검사계획서를 작성하는 단계 <div>품질검사계획서 작성 시 포함사항</div> <div>  <p>품질검사계획서</p> <ul style="list-style-type: none"> 품질검사 목적 검사방법 및 절차 검사일정 검사조직 (역할과 책임 포함) 소요예산 </div>

● 2. 품질검사 수행

수행 절차	2.1 품질검사 준비 / 2.2 품질검사 실시
주요 활동	<ul style="list-style-type: none"> 품질검사 준비는 본격적인 품질검사 수행에 앞서 품질검사를 효율적으로 수행하기 위해 사업수행기관에서 사전에 준비해야 할 검사환경을 준비하는 단계
	품질검사 환경 준비
	<div>품질검사 전담인력 지정</div> <ul style="list-style-type: none"> 검사 데이터 규모, 기간, 투입 인력 및 역할 등의 제반 여건에 부합되도록 전담인력을 구성
	<div>협조체계 구성</div> <ul style="list-style-type: none"> 수행기관 및 참여기관에서 학습모델 개발 조직 및 담당자와 협조체계를 구성, 다양한 이해관계자 협조 필요
	<div>검사대상 데이터 준비</div> <ul style="list-style-type: none"> 유관 부서의 사전 협조를 통해 검사대상 데이터가 누락되지 않도록 유의
	<div>검사 도구 준비</div> <ul style="list-style-type: none"> 검사 수행에 필요한 품질관리 지표들을 식별하고 현행화 적용 가능한 자동화 도구가 있다면 사전에 준비
	<div>수행 장소 마련</div> <ul style="list-style-type: none"> 수행 조직의 규모 및 인원, 기간 등을 식별 검사 데이터의 접근성, 보안, 원활한 업무 협조 등 고려

● 3. 검사결과 분석

수행 절차	3.1 오류원인 분석
주요 활동	<ul style="list-style-type: none"> 오류 원인을 추적하기 위해서는 선행적으로 품질검사 수행단계에서 조사된 추정 오류를 대상으로 실제로 오류 여부를 분류하고, 설사 오류라고 할지라도 임무와 데이터 특성을 고려하여 분석 대상 포함 여부 확인 ※ 오류 원인분석은 품질 수행체계-데이터구조-학습 모델 관점 등에서 추적하여 구체적인 근본 원인(Root Cause) 도출
	오류 원인분석 방향
	<div>구축 방법 측면의 오류 원인</div> <p>모델 데이터 대상 선정, 수집, 정제, 가공 등의 통제 미흡으로 인하여 구축 절차, 구조, 학습모델 측면의 다양한 오류 데이터 발생</p>
	<div>가이드라인 측면의 오류 원인</div> <p>사업수행기관이 정의한 구축 가이드라인의 불완전성, 미준수로 인하여 작업자 간 서로 상이하게 작업을 수행하거나 데이터 간 일관성 위배</p>
	<div>데이터셋 측면의 오류 원인</div> <p>데이터셋 설계의 부족, 구문 정확성 위배, 데이터 구축 중복 등으로 인하여 사업수행기관이 목표로 하는 품질목표 위배</p>
	<div>학습모델 측면의 오류 원인</div> <p>학습모델에 적합한 데이터 구축이 수행되지 않았거나, 잘못된 학습모델 선정으로 데이터 구축 방향이 잘못되어 품질목표를 맞출 수 없는 상황</p>

수행 절차	3.2 개선기회 도출			
주요 활동	<ul style="list-style-type: none"> 도출된 구체적인 개선기회는 향후 “개선 수행(Implement)” 단계의 실제적인 활동 내용으로 연결되어야 함 			
	<div>개선기회 유형</div> <table> <tr> <td> 구축 방법 측면 학습모델, 데이터 구조 및 구축 방법 변경·통제 프로세스 수립 </td><td> 가이드라인 측면 용어 개정 필요, 수집 및 획득 방법 개선, 도메인 추가 제정, 도메인의 중복 제거, 라벨링 수행방법, 검사방법 개선 등 </td></tr> <tr> <td> 데이터 구조 측면 라벨링 및 어노테이션 정규화 필요, 메타데이터 식별자 항목 조정, 원시/원천 데이터와 라벨링 데이터 간 1대1 매핑 및 관리 구조 통합 등 </td><td> 학습모델 측면 학습알고리즘 변경 또는 학습프로세스 보완, 학습모델의 원천데이터 및 라벨링데이터 간의 선·후행 활용 관계 프로세스 보완 등 </td></tr> </table>	구축 방법 측면 학습모델, 데이터 구조 및 구축 방법 변경·통제 프로세스 수립	가이드라인 측면 용어 개정 필요, 수집 및 획득 방법 개선, 도메인 추가 제정, 도메인의 중복 제거, 라벨링 수행방법, 검사방법 개선 등	데이터 구조 측면 라벨링 및 어노테이션 정규화 필요, 메타데이터 식별자 항목 조정, 원시/원천 데이터와 라벨링 데이터 간 1대1 매핑 및 관리 구조 통합 등
구축 방법 측면 학습모델, 데이터 구조 및 구축 방법 변경·통제 프로세스 수립	가이드라인 측면 용어 개정 필요, 수집 및 획득 방법 개선, 도메인 추가 제정, 도메인의 중복 제거, 라벨링 수행방법, 검사방법 개선 등			
데이터 구조 측면 라벨링 및 어노테이션 정규화 필요, 메타데이터 식별자 항목 조정, 원시/원천 데이터와 라벨링 데이터 간 1대1 매핑 및 관리 구조 통합 등	학습모델 측면 학습알고리즘 변경 또는 학습프로세스 보완, 학습모델의 원천데이터 및 라벨링데이터 간의 선·후행 활용 관계 프로세스 보완 등			

4. 개선 수행

수행 절차	4.1 개선계획 수립					
주요 활동	<ul style="list-style-type: none">• 검사결과 분석 단계에서 도출된 개선 기회를 종합하여 개선과제를 도출하고, 분류된 개선과제별로 상세 추진계획 수립					
	<div>품질개선계획 수립 내용</div> <table><tr><td><div>품질개선 목적</div><ul style="list-style-type: none">• 품질개선 필요성 기술</td><td><div>주요 과업 범위</div><ul style="list-style-type: none">• 개선 유형과 수행 절차 기술</td><td><div>일정계획 수립</div><ul style="list-style-type: none">• 품질개선 일정 계획 수립</td></tr><tr><td><div>추진체계</div><ul style="list-style-type: none">• 수행 조직 및 이해관계자 역할과 책임 정의</td><td><div>소요예산</div><ul style="list-style-type: none">• 투입 인력 규모, 필요 H/W, S/W 규모 및 소요 비용 등</td><td><div>기대효과</div><ul style="list-style-type: none">• 개선효과를 정량적, 정성적으로 평가</td></tr></table>	<div>품질개선 목적</div> <ul style="list-style-type: none">• 품질개선 필요성 기술	<div>주요 과업 범위</div> <ul style="list-style-type: none">• 개선 유형과 수행 절차 기술	<div>일정계획 수립</div> <ul style="list-style-type: none">• 품질개선 일정 계획 수립	<div>추진체계</div> <ul style="list-style-type: none">• 수행 조직 및 이해관계자 역할과 책임 정의	<div>소요예산</div> <ul style="list-style-type: none">• 투입 인력 규모, 필요 H/W, S/W 규모 및 소요 비용 등
<div>품질개선 목적</div> <ul style="list-style-type: none">• 품질개선 필요성 기술	<div>주요 과업 범위</div> <ul style="list-style-type: none">• 개선 유형과 수행 절차 기술	<div>일정계획 수립</div> <ul style="list-style-type: none">• 품질개선 일정 계획 수립				
<div>추진체계</div> <ul style="list-style-type: none">• 수행 조직 및 이해관계자 역할과 책임 정의	<div>소요예산</div> <ul style="list-style-type: none">• 투입 인력 규모, 필요 H/W, S/W 규모 및 소요 비용 등	<div>기대효과</div> <ul style="list-style-type: none">• 개선효과를 정량적, 정성적으로 평가				

수행 절차	4.2 품질개선 실시
주요 활동	<ul style="list-style-type: none"> 수립된 개선계획 중 단기 과제를 우선으로, 개선 수행을 위한 환경을 준비하고, 실제 개선을 수행
	<p>품질개선 준비 및 실시</p>

5. 품질통제

수행 절차	5.1 결과 평가
주요 활동	<ul style="list-style-type: none"> 품질 관리체계 수립, 데이터 운영계획 수립, 데이터 보정으로 구분 구축과정 가이드라인 개정 영역과 실제 개선 영역을 종합적으로 표현, 실제 개선한 영역의 개선 효과를 정량적으로 평가
수행 절차	5.2 품질목표 관리
주요 활동	<ul style="list-style-type: none"> 품질통제 대상에 대한 품질목표 설정 품질을 유지하거나 더 높은 품질 확보방안 수립
수행 절차	5.3 품질통제 실시
주요 활동	<ul style="list-style-type: none"> 품질목표 관리 대상에 대해 수시로 모니터링 수행 사업수행기관 구성원들 교육, 홍보 활동으로 품질목표 관리 활동에 대한 공감대 형성 적극적인 참여 유도과 소통 필요

4 단계별 산출물

4.1 구축 및 품질 산출물

※ 제출여부 - Y : 필수 제출 산출물, C : 선택적 제출 산출물

구분	산출물명	유형	필수 여부	AI 허브 공개산출물
구축계획 수립	사업수행계획서 (※ 구축계획서, 품질관리계획서 포함)	문서	Y	-
데이터 획득/수집, 정제, 가공	개인정보 자율점검표	문서	Y	-
	구축공정 가이드라인	문서	Y	-
	저작도구(Viewer) 사용설명서	문서	Y	Y
	저작도구/Viewer	어플리케이션	Y	Y
	원시데이터(획득/수집데이터)	데이터	C	-
	원천데이터(이미지, 영상, 오디오 등)	데이터	Y	-
데이터 학습	라벨링데이터(JSON, XML, CSV 등)	데이터	Y	-
	AI모델 설명서	문서	Y	Y
	AI모델 소스코드	Code	Y	Y
	학습모델 파일	binary	Y	Y
	모델 구축 도커이미지	tar	Y	Y
데이터 품질	AI모델 테스트 결과서	문서	C	-
	품질관리결과서	문서	Y	-
	품질검증결과서	문서	Y	-
	품질지표 기준서(사업수행계획서 내 포함)	문서	Y	-
데이터 운영·활용	활용 가이드라인	문서	Y	Y
	데이터 활용 교육 영상	동영상	Y	Y
	하자 및 유지보수 계획서	문서	Y	-

〈표 II-29〉 구축 및 품질 관련 산출물 목록

4.2 구축단계별 산출물

※ 범례 - * 필수 제출 산출물, ** 필수(공개) 제출 산출물

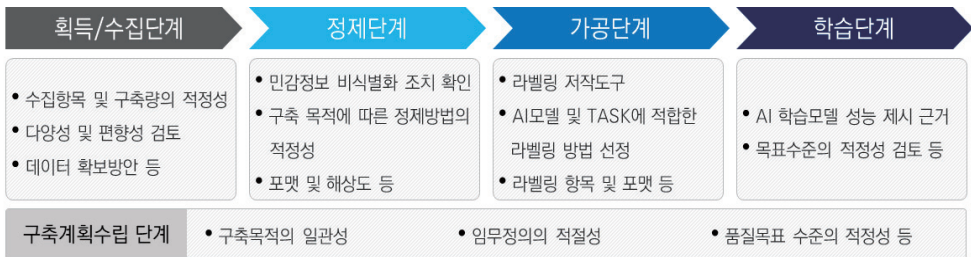
구분	구축계획 수립	데이터 획득/수집	데이터 정제	데이터 가공	데이터 학습	데이터 운영·활용
문서	사업수행계획서 (별첨포함)*					사업수행계획서 (별첨포함)*
	구축공정 가이드라인 (임무정의)*	구축공정 가이드라인 (획득)*	구축공정 가이드라인 (정제)*	구축공정 가이드라인 (가공)*		
						활용 가이드라인**
		획득 저작도구 사용설명서	정제 저작도구 사용설명서**	라벨링 저작도구 사용설명서**		
		품질관리결과서*	품질관리결과서*	품질관리결과서*	품질관리결과서*	
		개인정보 자율점검표*				
					시모델 설명서**	
					시모델 테스트 결과서	
						하자 및 유지보수 계획서*
동영상						데이터 활용 교육 영상**
데이터		원시데이터 (획득/수집 데이터)	원천데이터 (이미지, 영상, 오디오 등)**	라벨링데이터 (JSON, XML, CSV 등)**		
어플리케이션		저작도구	저작도구**	저작도구**		
기타					시모델 소스코드**	
					학습모델 파일**	
					모델 구축 도커이미지**	

〈표 Ⅱ-30〉 구축단계별 산출물

제3장 | 품질 자가점검 및 품질검증

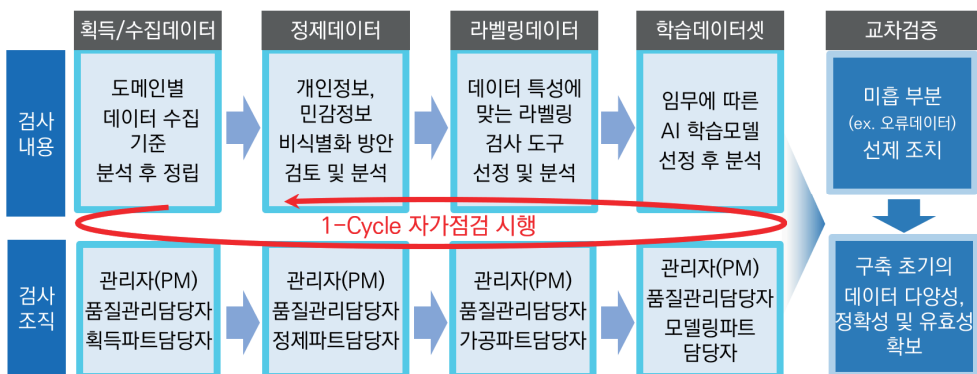
1 개요

- 사업수행기관의 품질관리 역량, 데이터 유형별 구축 난이도, 해당 분야 전문가 및 인공지능 전문가 확보 유무 등에 따라 품질수준 편차 발생으로 품질 자가점검 필요성 증대



[그림 II-21] 구축단계별 자가점검 체크리스트 활용

- 1-Cycle 자가점검은 인공지능 학습용 구축 초기의 데이터를 이용하여 「획득/수집 → 정제 → 가공 → 학습」까지 전체 공정에 대한 1-Cycle 데이터(적합성, 정확성, 유효성) 검사 실시 후, 수행 방법이 목표에 부합하지 않는다면 데이터의 획득/수집, 정제, 가공 절차를 빠르게 변경하여 재작업을 하기 위한 필수 절차



[그림 II-22] 1-Cycle 자가점검 검사 내용 및 조직

2 품질 자가점검

2.1 품질 자가점검

- 인공지능 학습용 데이터 구축사업 진행 시 구축공정별 품질 이슈를 최소화하기 위해 사업수행기관의 초기데이터 1-Cycle 자가점검, 중간데이터 품질 자가점검은 필수적으로 수행하고, 그 외 품질 자가점검은 사업규모, 사업범위, 일정 등에 따라 선택적으로 수행

2.1.1 품질 자가점검 프로세스



[그림 II-23] 품질 자가점검 프로세스

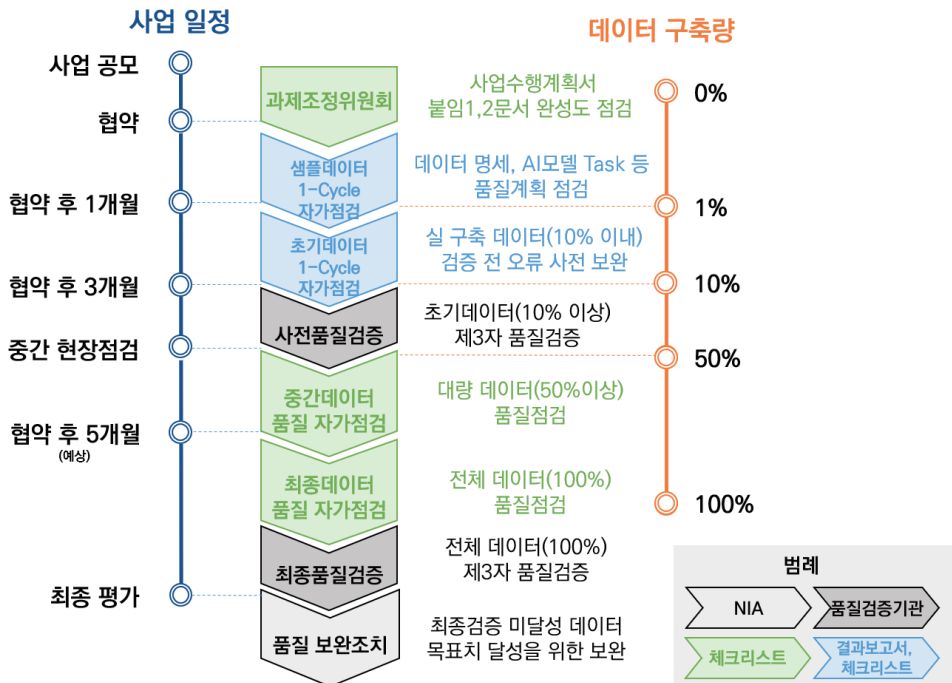
2.1.2 품질 자가점검 프로세스 주요 내용

사업일정	품질 자가점검 활동	주요 검토내용
과제조정위원회	과제조정위원회 자가점검	• 사업수행계획서 내 구축계획서, 품질관리계획서의 적정성 검토
사업 1개월 시점	샘플데이터 1-Cycle 자가점검	• 데이터 설계의 적정성 점검 및 데이터 구축 목적과 AI 학습모델 후보군의 합치성 및 준비도 점검
사업 3개월 시점	초기데이터 1-Cycle 자가점검	• 측정지표의 적정성 및 정량목표의 달성도 확인 • 데이터 형태별 오류유형에 따른 보완조치 방안 마련
중간시점	중간데이터 품질 자가점검	• 초기데이터 1-Cycle 결과 반영 여부 검토 • 대량 데이터(50% 이상) 기반의 유효성 점검
최종시점	최종데이터 품질 자가점검	• 최종 데이터 품질 전반 점검

〈표 II-31〉 품질 자가점검 활동의 주요 검토내용

품질 자가점검 활동	입력물	출력물	유형	품질검증
과제조정위원회 자가점검	사업수행계획서	<ul style="list-style-type: none"> 과제조정위원회 자가점검 체크리스트 	문서	-
샘플데이터 1-Cycle 자가점검	학습용 데이터 1% 이내	<ul style="list-style-type: none"> 샘플데이터 1-Cycle 자가점검 결과보고서 	문서	-
초기데이터 1-Cycle 자가점검	학습용 데이터 10% 이내	<ul style="list-style-type: none"> 초기데이터 1-Cycle 자가점검 결과보고서 	문서	-
		<ul style="list-style-type: none"> 초기데이터(10% 이내) 	데이터	대상
중간데이터 품질 자가점검	학습용 데이터 50% 이상	<ul style="list-style-type: none"> 중간데이터 품질 자가점검 체크리스트 	문서	-
최종데이터 품질 자가점검	학습용 데이터 100%	<ul style="list-style-type: none"> 사업수행결과보고서 	문서	-
		<ul style="list-style-type: none"> 전체 데이터(100%) 	데이터	대상

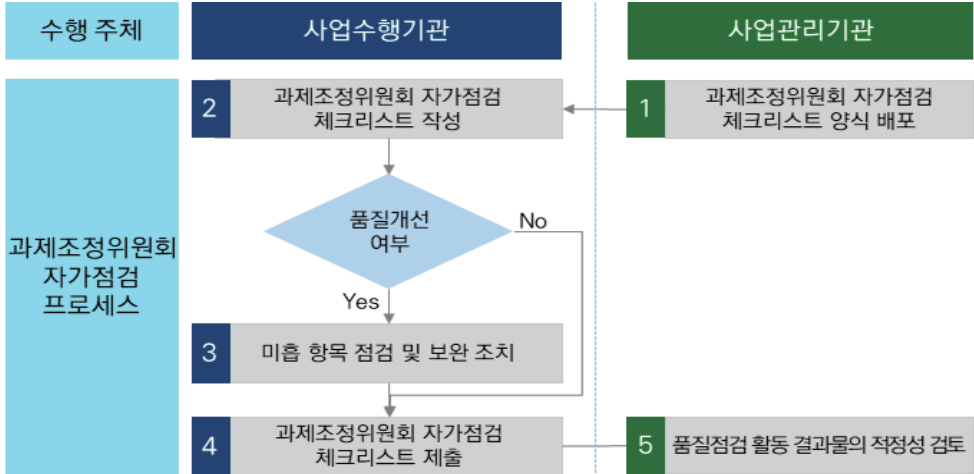
〈표 II-32〉 품질 자가점검 활동의 입력물 및 출력물



〔그림 II-24〕 사업 일정에 따른 품질 자가점검 활동

2.2 품질 자가점검 세부 활동 정의

2.2.1 과제조정위원회(과조위) 자가점검 프로세스

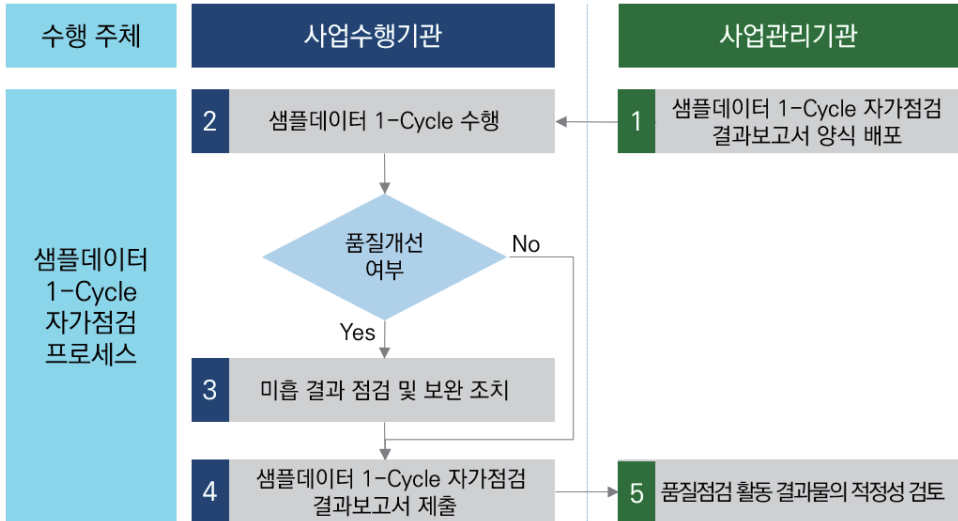


[그림 II-25] 과제조정위원회(과조위) 자가점검 프로세스

번호	활동	주체	주요 내용	산출물
1	과제조정위원회 자가점검 체크리스트 양식 배포	사업관리기관	<ul style="list-style-type: none"> 사업수행계획서 내 구축계획서 및 품질계획서 자가점검 체크리스트 양식 배포 우선협상대상자에게 전달하여 작성요청 	과제조정위원회 자가점검 체크리스트 양식
2	과제조정위원회 체크리스트 작성	사업수행기관	<ul style="list-style-type: none"> 사업수행기관이 자가점검 체크리스트 작성 	과제조정위원회 체크리스트
3	미흡항목 점검 및 보완	사업수행기관	<ul style="list-style-type: none"> 자가점검 체크리스트 점검항목 중 미흡 항목에 대해 점검 및 보완 	과제조정위원회 체크리스트, 사업수행계획서
4	과제조정위원회 체크리스트 제출	사업수행기관	<ul style="list-style-type: none"> 최종 작성된 자가점검 체크리스트 제출 	과제조정위원회 체크리스트
5	품질점검 활동 결과물의 적정성 검토	사업관리기관	<ul style="list-style-type: none"> 제출된 자가점검 체크리스트는 과조위 위원들과 공유, 과조위 개최 시 점검 및 활용 최종 과제조정위원회 종료 후, 유닛별 품질담당자에게 최종본 회신 	과제조정위원회 검토의견서, 컨설팅보고서

〈표 II-33〉 과제조정위원회 자가점검 품질관리 활동 및 산출물

2.2.2 샘플데이터 1-Cycle 자가점검 프로세스

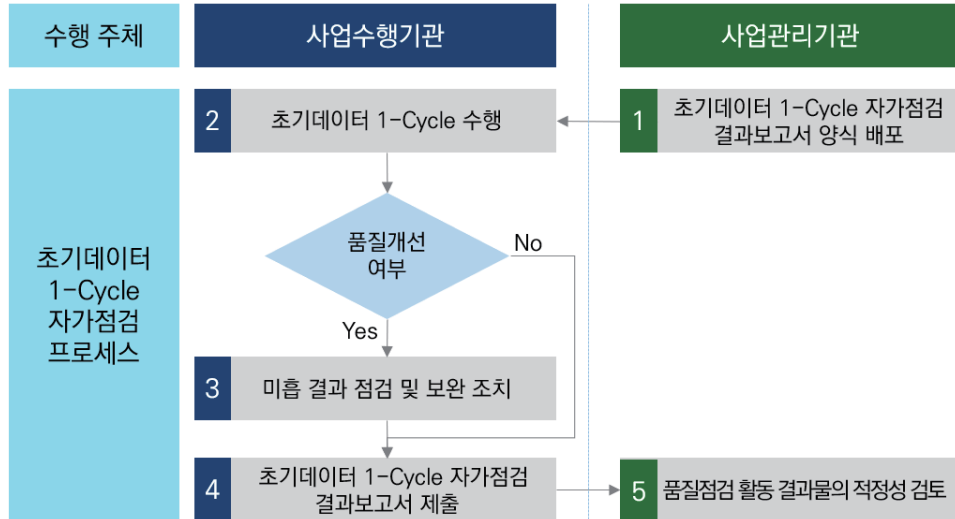


[그림 II-26] 샘플데이터 1-Cycle 자가점검 프로세스

번호	활동	주체	주요 내용	산출물
1	샘플데이터 1-Cycle 자가점검 결과보고서 양식 배포	사업관리기관	<ul style="list-style-type: none"> 최소 목표 수량(1% 이내)으로 샘플데이터의 획득/수집, 정제, 가공, AI 모델 적용방안 검토 및 데이터 명세 확정을 위한 자가점검 결과보고서 양식 배포 	샘플데이터 1-Cycle 자가점검 결과보고서 양식
2	샘플데이터 1-Cycle 수행	사업수행기관	<ul style="list-style-type: none"> 샘플데이터 1-Cycle 수행 자가점검 결과보고서 작성 	샘플데이터 1-Cycle 결과보고서, 체크리스트
3	미흡 결과 점검 및 보완	사업수행기관	<ul style="list-style-type: none"> 샘플데이터 1-Cycle 결과에 따라 보완 단계별 구축공정 가이드라인에 반영 	보완된 결과보고서 등 산출물
4	샘플데이터 1-Cycle 자가점검 결과보고서 제출	사업수행기관	<ul style="list-style-type: none"> 최종 작성된 자가점검 결과보고서 제출 	샘플데이터 1-Cycle 결과보고서, 체크리스트
5	품질점검 활동 결과물의 적정성 검토	사업관리기관	<ul style="list-style-type: none"> 사업담당자, 유닛별 품질담당자 등 공유 품질점검 활동 결과 분석 및 적정성 검토 	컨설팅보고서

〈표 II-34〉 샘플데이터 1-Cycle 자가점검 품질관리 활동 및 산출물

2.2.3 초기데이터 1-Cycle 자가점검 프로세스

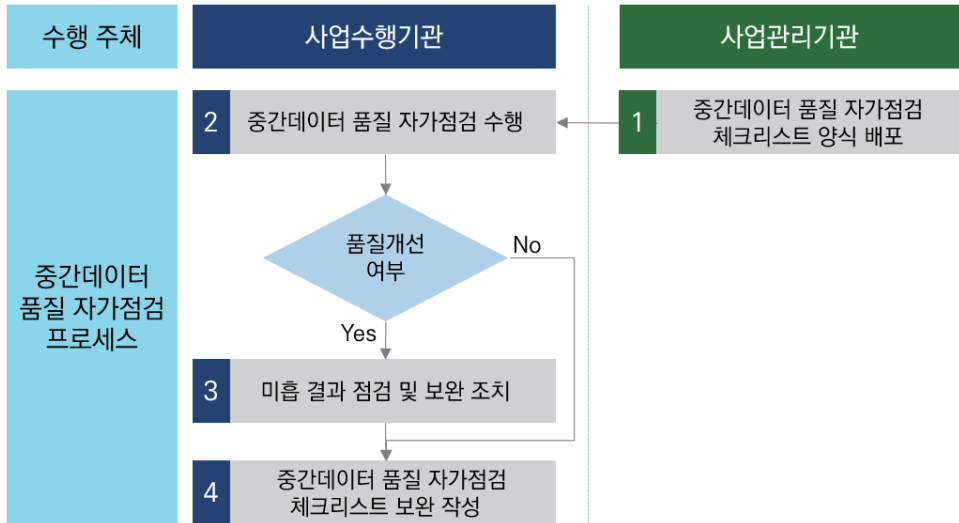


[그림 II-27] 초기데이터 1-Cycle 자가점검 프로세스

번호	활동	주체	주요 내용	산출물
1	초기데이터 1-Cycle 자가점검 결과보고서 양식 배포	사업관리기관	<ul style="list-style-type: none"> 초기데이터 수량(10% 이상)으로 품질 정량 목표에 부합하는지에 대한 자가점검 결과보고서 양식 배포 	초기데이터 1-Cycle 자가점검 결과보고서 양식
2	초기데이터 1-Cycle 수행	사업수행기관	<ul style="list-style-type: none"> 초기데이터 1-Cycle 수행 측정지표의 적정성 및 정량 목표의 달성도 점검 자가점검 결과보고서 작성 	초기데이터 1-Cycle 결과보고서, 체크리스트
3	미흡 결과 점검 및 보완	사업수행기관	<ul style="list-style-type: none"> 샘플데이터 1-Cycle 결과 반영여부 점검 데이터 오류 유형에 따른 보완 방안 마련 단계별 구축공정 가이드라인에 반영 	초기데이터, 보완된 결과보고서 등 산출물
4	초기데이터 1-Cycle 자가점검 결과보고서 제출	사업수행기관	<ul style="list-style-type: none"> 최종 작성된 자가점검 결과보고서 제출 	초기데이터 1-Cycle 결과보고서, 체크리스트
5	품질점검 활동 결과물의 적정성 검토	사업관리기관	<ul style="list-style-type: none"> 사업담당자, 유닛별 품질담당자 등 공유 품질점검 활동 결과 분석 및 적정성 검토 	컨설팅보고서

〈표 II-35〉 초기데이터 1-Cycle 자가점검 품질관리 활동 및 산출물

2.2.4 중간데이터 품질 자가점검 프로세스

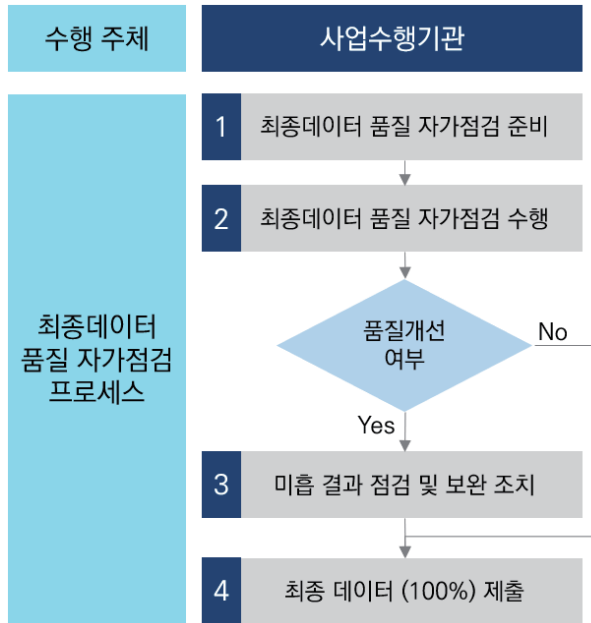


[그림 II-28] 중간데이터 품질 자가점검 프로세스

번호	활동	주체	주요 내용	산출물
1	중간데이터 품질 자가점검 체크리스트 양식 배포	사업관리기관	<ul style="list-style-type: none"> 대량 데이터(50% 이상)에 대한 품질 자가점검을 위한 체크리스트 양식 배포 	중간데이터 품질 자가점검 체크리스트 양식
2	중간데이터 품질 자가점검 수행	사업수행기관	<ul style="list-style-type: none"> 초기데이터 1-Cycle 결과 반영 여부 점검 대량 데이터(50% 이상) 기반의 정확성 및 유효성 점검(대량 데이터에서 다양성 기반 균등 샘플링 후 점검) 	중간데이터 품질 자가점검 체크리스트
3	미흡 결과 점검 및 보완	사업수행기관	<ul style="list-style-type: none"> 중간데이터 품질점검결과와 보완 조치 단계별 구축공정 가이드라인에 반영 	산출물 보완
4	중간데이터 품질 자가점검 체크리스트 보완 작성	사업수행기관	<ul style="list-style-type: none"> 자가점검 체크리스트 재검토 및 보완 	중간데이터 품질 자가점검 체크리스트

〈표 II-36〉 중간데이터 품질 자가점검 품질관리 활동 및 산출물

2.2.5 최종데이터 품질 자가점검 프로세스



[그림 II-29] 최종데이터 품질 자가점검 프로세스

번호	활동	주체	주요 내용	산출물
1	최종데이터 품질 자가점검 준비	사업수행기관	<ul style="list-style-type: none"> 전체 데이터(100%) 구축 사업수행계획서, 품질지표 기준서 v2.0, 품질검사 도구 등 준비 	전체 데이터, 최종 산출물 등
2	최종데이터 품질 자가점검 실행	사업수행기관	<ul style="list-style-type: none"> 구축완료 데이터(100%)로 사업수행계획서 전 공정 결과값이 품질기준에 부합하는지 여부 점검 	체크리스트
3	미흡 결과 점검 및 보완	사업수행기관	<ul style="list-style-type: none"> 최종 품질점검결과와 보완 조치 오류 원인 및 유형 분석 후 개선 조치 	산출물 보완
4	최종 데이터 제출	사업수행기관	<ul style="list-style-type: none"> 최종 데이터를 사업관리기관에 제출 제3자 품질검증을 위한 데이터 제출 	학습용 데이터 100%

〈표 II-37〉 최종데이터 품질 자가점검 품질관리 활동 및 산출물

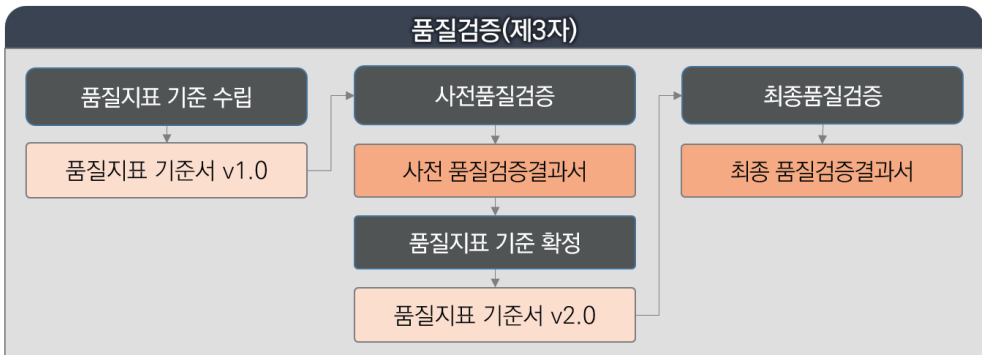
3 제3자 품질검증

3.1 제3자 품질검증

- 인공지능 학습용 데이터 개별 특성에 따른 객관적 품질지표 기준을 마련하고, 지표 기준에 따라 제3자 품질검증을 수행하여 구축 데이터의 품질목표 달성여부 측정

※ 제3자 품질검증 : 인공지능 학습용 데이터의 품질관리 기준, 절차 및 방법을 이해하고 있는 기관 또는 업체를 통해 사업수행기관이 구축한 인공지능 학습용 데이터 품질을 검증하는 활동

3.1.1 제3자 품질검증 프로세스



[그림 II-30] 제3자 품질검증 프로세스

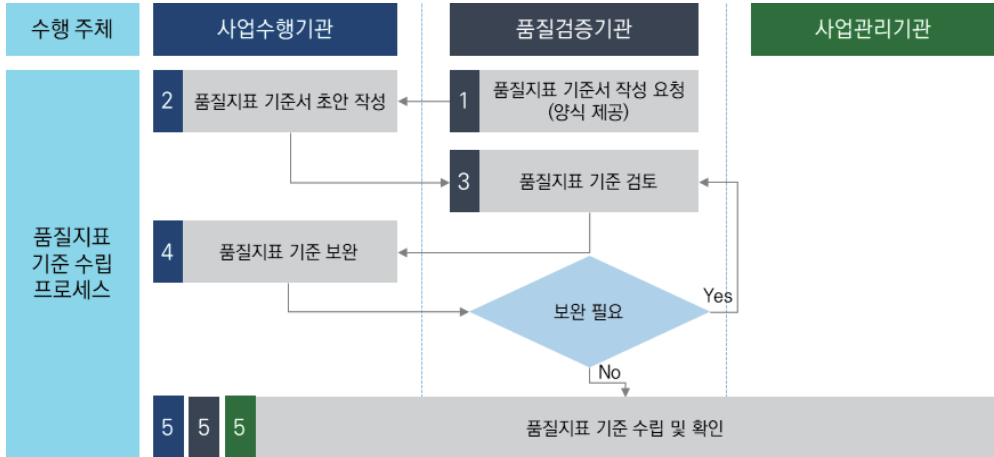
3.1.2 제3자 품질검증 프로세스 주요 내용

활동	주요 내용	주요 산출물
품질지표 기준 수립	<ul style="list-style-type: none"> 사업수행기관이 제출한 사업수행계획서를 토대로 사전 품질검증을 위한 품질지표 기준 수립 	품질지표 기준서 v1.0
사전품질검증	<ul style="list-style-type: none"> 구축 중인 학습용 데이터셋의 오류를 사전에 검출하여 수정하고 예방 검증 대상 : 다양성, 구문 정확성, 의미 정확성 	사전 품질검증 결과서
품질지표 기준 확정	<ul style="list-style-type: none"> 사전 품질검증 결과를 토대로 이슈 및 변경 사항 발생 시 협의하여 품질지표 기준서에 반영 최종품질검증을 위한 품질지표 기준 최종 확정 	품질지표 기준서 v2.0
최종품질검증	<ul style="list-style-type: none"> 최종 구축된 학습용 데이터셋을 대상으로 품질목표 달성여부 측정 검증 대상 : 다양성, 구문 정확성, 의미 정확성, 유효성 	품질검증 결과서

〈표 II-38〉 제3자 품질검증 프로세스 주요 내용

3.2 제3자 품질검증 세부 활동 정의

3.2.1 품질지표 기준 수립 프로세스



[그림 II-31] 품질지표 기준 수립 프로세스

번호	활동	주체	주요 내용	산출물
1	품질지표 기준서 작성 요청	품질검증기관	• 품질지표 기준서 양식 및 안내자료 배포	• 품질지표 기준서 양식 • 품질지표 기준서 작성 안내자료
2	품질지표 기준서 초안 작성	사업수행기관	• 사업수행계획서를 토대로 품질지표 기준서 초안 작성 • 항목별 측정조건 작성 • 구문 정확성 검사규칙 작성 • 의미 정확성 검사기준서 작성	• 품질지표 기준서 (초안) • 항목별 측정조건 • 구문 정확성 검사규칙 • 의미 정확성 검사기준서
3	품질지표 기준 검토	품질검증기관	• 품질지표 기준 제반문서 검토	-
4	품질지표 기준 보완	사업수행기관	• 검토내용을 토대로 품질지표 기준 수정 및 보완	• 품질지표 기준서 (수정) • 항목별 측정조건 • 구문 정확성 검사규칙 • 의미 정확성 검사기준서
5	품질지표 기준 수립 및 확인	사업수행기관 품질검증기관 사업관리기관	• 최종 보완된 품질지표 및 목표에 대한 합의 • 합의된 품질지표 기준 확인	• 품질지표 기준서 v1.0

〈표 II-39〉 품질지표 기준 수립 활동 및 산출물

3.2.2 사전품질검증 프로세스

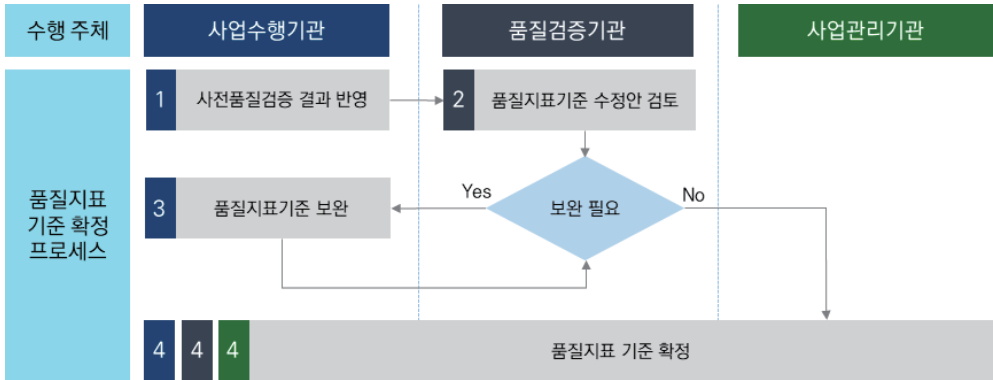


[그림 II-32] 사전품질검증 프로세스

번호	활동	주체	주요 내용	산출물
1	사전품질검증 안내	품질검증기관	<ul style="list-style-type: none"> 품질검증 안내자료 배포 품질검증 교육 제공 	<ul style="list-style-type: none"> 품질검증안내자료 품질검증 양식
		사업관리기관	<ul style="list-style-type: none"> 초기데이터(10% 이상) 제출 안내 	<ul style="list-style-type: none"> 요청 공문
2	사전품질검증 준비	품질검증기관	<ul style="list-style-type: none"> 품질지표 기준서 검토 	<ul style="list-style-type: none"> 초기데이터
		사업수행기관	<ul style="list-style-type: none"> 데이터셋 제출(초기데이터 10% 이상) 품질검증 참고자료 산출물 제출 검증 준비 현황 점검 	
3	사전품질검증 의뢰	사업수행기관	<ul style="list-style-type: none"> 온라인 검증의뢰 	-
4	사전품질검증 수행	품질검증기관	<ul style="list-style-type: none"> 사전품질검증 수행 및 검증 결과 작성 	<ul style="list-style-type: none"> 사전품질검증 결과서
5	사전품질검증 결과 검토	사업수행기관 품질검증기관	<ul style="list-style-type: none"> 검증 결과 상호검토 및 합의 	-
6	사전품질검증 결과 공유	품질검증기관	<ul style="list-style-type: none"> 사전품질검증 결과를 사업관리기관 및 사업수행기관에 공유 	-
7	사전품질검증 결과 확인	사업수행기관 사업관리기관	<ul style="list-style-type: none"> 사전품질검증 결과 확인 ※ 검증 결과 오류는 반드시 수정·반영 	<ul style="list-style-type: none"> 사업관리대장

〈표 II-40〉 사전품질검증 활동 및 산출물

3.2.3 품질지표 기준 확정 프로세스



[그림 Ⅱ-33] 품질지표 기준 확정 프로세스

번호	활동	주체	주요 내용	산출물
1	사전품질검증 결과 반영	사업수행기관	<ul style="list-style-type: none"> 사전품질검증 결과를 품질지표 기준서 v1.0에 반영 항목별 측정조건 변경 반영 구문 정확성 검사규칙 변경 반영 의미 정확성 검사기준서 변경 반영 	<ul style="list-style-type: none"> 품질지표 기준서 v1.0 (수정) 항목별 측정조건 구문 정확성 검사규칙 의미 정확성 검사기준서
2	품질지표 기준 수정안 검토	품질검증기관	<ul style="list-style-type: none"> 사업수행기관이 수정한 품질지표 기준 자료 검토 	-
3	품질지표 기준 보완	사업수행기관	<ul style="list-style-type: none"> 검토내용을 토대로 품질지표 기준 수정 및 보완 	<ul style="list-style-type: none"> 품질지표 기준서 v1.0 (수정) 항목별 측정조건 구문 정확성 검사규칙 의미 정확성 검사기준서
4	품질지표 기준 확정	사업수행기관 품질검증기관 사업관리기관	<ul style="list-style-type: none"> 최종 보완된 내용으로 품질지표 및 목표에 대한 합의 합의된 품질지표기준 확인 	<ul style="list-style-type: none"> 품질지표 기준서 v2.0

〈표 Ⅱ-41〉 품질지표 기준 확정 활동 및 산출물

3.2.4 최종품질검증 프로세스



[그림 II-34] 최종품질검증 프로세스

번호	활동	주체	주요 내용	산출물
1	품질검증안내	품질검증기관 사업관리기관	<ul style="list-style-type: none"> 품질검증 안내자료 배포 최종데이터(100%) 제출 안내 	<ul style="list-style-type: none"> 품질검증안내자료 요청 공문
2	품질검증 준비	품질검증기관 사업수행기관	<ul style="list-style-type: none"> 품질지표 기준서 검토 데이터셋 제출(최종데이터 100%) 품질검증 참고자료 산출물 제출 검증 준비 현황 점검 유효성 증빙자료 제출 	<ul style="list-style-type: none"> 최종데이터 유효성 증빙자료
3	품질검증 의뢰	사업수행기관	<ul style="list-style-type: none"> 온라인 검증의뢰 	-
4	품질검증 수행	품질검증기관	<ul style="list-style-type: none"> 품질검증 수행 및 검증 결과 작성 	<ul style="list-style-type: none"> 품질특성별 검증결과
5	검증결과 검토	품질검증기관 사업수행기관	<ul style="list-style-type: none"> 검증결과 상호검토 및 합의 	-
6	품질검증결과서 발급	품질검증기관	<ul style="list-style-type: none"> 사업관리기관 및 사업수행기관에 품질검증 결과서 발급 	<ul style="list-style-type: none"> 품질검증결과서
7	품질검증결과 확인	사업수행기관 사업관리기관	<ul style="list-style-type: none"> 최종 품질검증결과 확인 	<ul style="list-style-type: none"> 사업관리대장

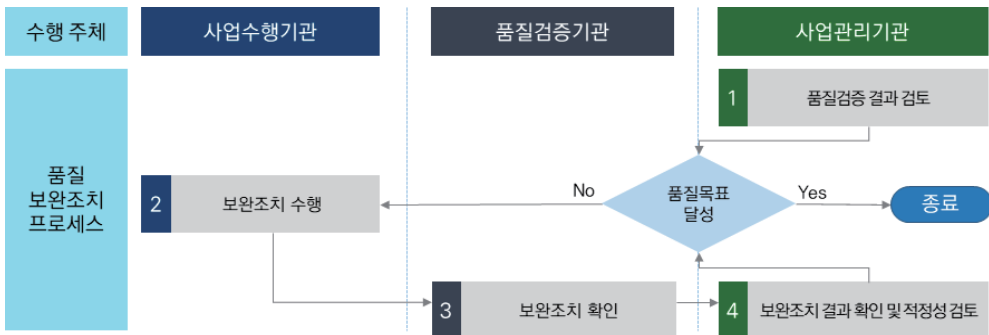
〈표 II-42〉 최종품질검증 활동 및 산출물

● 품질검증항목

번호	품질지표	주요 내용	기준
1	다양성(통계)	• 관심객체, 카테고리, 수집환경 등 인공지능 학습용 데이터의 주요특성을 통계적 방법으로 확인	충분성, 균등성, 편향성 여부 확인
2	다양성(요건)	• 사업수행기관의 구축목표 대비 구축결과물 비율(수량)이 충족하는지 통계적 방법으로 확인	구축목표 대비 구축결과물 비율(수량) 기준 설정
3	구문 정확성	• 라벨 데이터 포맷과 값이 정확하게 입력되어 있고 필수항목 누락 여부 검사	정확도 99.5% 이상 권고
4	의미 정확성	• 어노테이션 값이 의미적으로 정답(GT)인지 확인하는 항목으로 실제적인 정답 비율 확인	정확도 95% 이상 권고
5	유효성	• 구축한 데이터세트를 잘 알려진 인공지능 학습 모델로 훈련시킨 후 목표로 했던 수준의 성능 달성이 가능한지 확인	인공지능 학습 모델 별 적정 기준 설정

〈표 II-43〉 품질검증항목 내용

3.2.5 품질 보완조치 프로세스



〔그림 II-35〕 품질 보완조치 프로세스

번호	활동	주체	주요 내용	산출물
1	품질검증 결과 검토	사업관리기관	• 최종품질검증 결과 달성여부 검토 • 검증 결과에 따라 사업수행기관에 보완조치 시행 요구	• 사업관리대장
2	보완조치 수행	사업수행기관	• 미달성 항목에 대한 보완 수행 (데이터, SI모델 등) • 보완 완료 결과물 업로드	• 보완 완료한 데이터 혹은 SI모델
3	보완조치 확인	품질검증기관	• 사업관리기관에 보완조치 확인 결과 통보	• 보완조치 확인 결과 (공문)
4	보완조치 결과 확인 및 적정성 검토	사업관리기관	• 보완조치 결과 확인 • 필요시 사업수행기관에 재보완 요구	• 사업관리대장

〈표 II-44〉 품질 보완조치 활동 및 산출물

● 품질지표에 따른 보완조치 항목

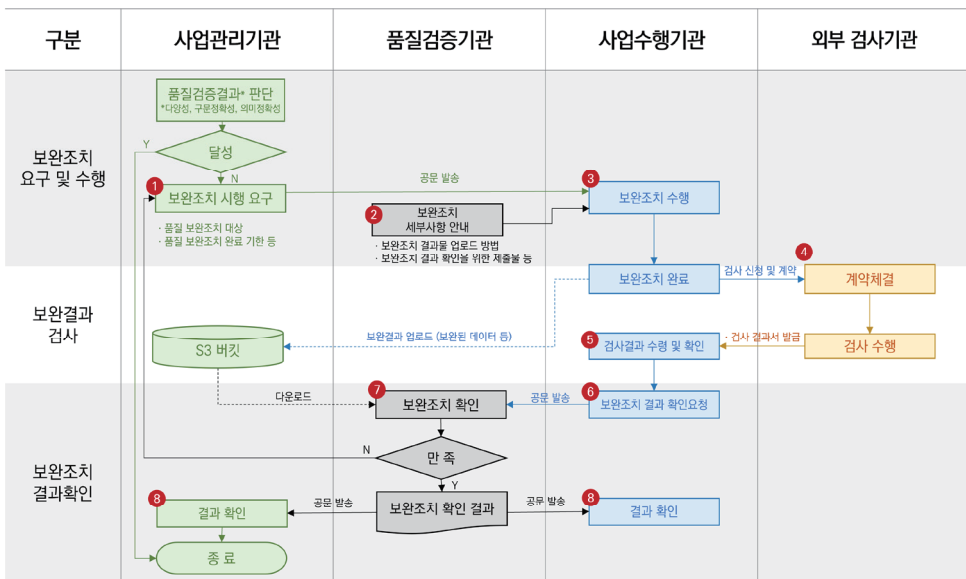
※ 품질지표별 보완 대상, 세부 절차 등 일부 변동 가능

품질지표	보완 대상	검사 주체	세부 절차
다양성	전체 데이터	외부 검사기관	[그림 Ⅱ-36]
구문 정확성			
의미 정확성			
유효성	데이터, SI 모델, 관련 서류	품질검증기관	[그림 Ⅱ-37]

〈표 Ⅱ-45〉 품질 보완조치 대상 구분

● 품질 보완조치 세부 절차

- 다양성, 구문 정확성, 의미 정확성 보완조치 절차

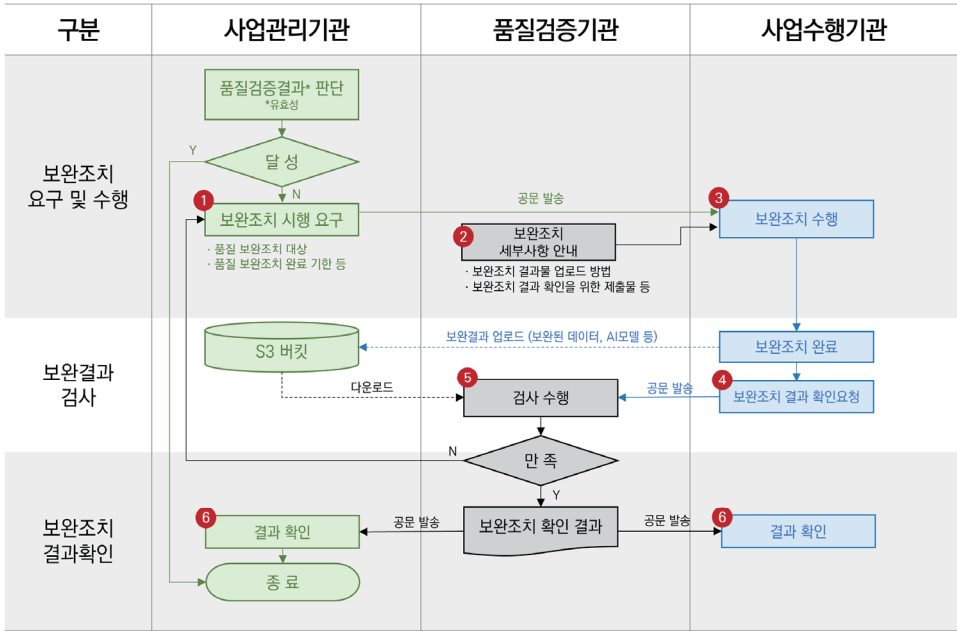


[그림 Ⅱ-36] 품질 보완조치 세부절차(다양성, 구문 정확성, 의미 정확성)

번호	활동	주체	주요 내용
1	보완조치 시행 요구	사업관리기관	<ul style="list-style-type: none"> 사업관리기관은 품질검증 결과에 따라 사업수행기관에 품질 보완조치 시행 요구 공문 발송 품질 보완조치 시행 요구 시, 아래 항목 포함 <ul style="list-style-type: none"> - 품질 보완조치 대상 - 데이터 보완 완료 기한 등
2	보완조치 세부사항 안내	품질검증기관	<ul style="list-style-type: none"> 사업수행기관은 품질검증기관에 연락하여 품질 보완조치에 대한 세부 안내 확인 <ul style="list-style-type: none"> - 품질검증기관은 데이터 업로드 방법, 제출서류, 외부 검사기관 리스트 등에 대해 안내 ※ 계약 추진 및 보완조치 검사 일정 협의 등은 사업수행기관이 직접 외부 검사기관과 진행함
3	보완조치 수행	사업수행기관	<ul style="list-style-type: none"> 사업수행기관은 전체 데이터에 대해 보완 수행 사업수행기관은 보완 완료된 데이터를 S3 버킷에 업로드
4	계약 체결	사업수행기관 외부 검사기관	<ul style="list-style-type: none"> 사업수행기관은 외부 검사기관과 개별적으로 검증 의뢰 사업수행기관과 외부 검사기관은 검사 대상, 검사 기준, 검사 기간 등 협의 후 계약 체결 ※ 사업수행기관은 검증 일정을 사업관리기관, 품질검증기관에 공유하여야 함
5	검사결과 수령 및 확인	사업수행기관	<ul style="list-style-type: none"> 사업수행기관은 외부 검사기관이 발급한 검사결과를 검토 후 이상 없음을 확인
6	보완조치 결과 확인요청	사업수행기관	<ul style="list-style-type: none"> 사업수행기관은 외부 검사기관의 검사결과를 바탕으로 품질검증기관에 보완조치 결과 확인요청 공문 발송 <ul style="list-style-type: none"> - 보완조치 결과 확인요청 공문에 외부 검사기관으로부터 발급받은 검사 결과를 첨부
7	보완조치 확인	품질검증기관	<ul style="list-style-type: none"> 품질검증기관은 사업수행기관의 제출서류(공문 및 외부 검사기관의 검사결과)를 검토 <ul style="list-style-type: none"> - 보완결과 만족 시, 품질검증기관은 보완조치 확인 결과를 사업관리기관 및 사업수행기관에 공문으로 통보 - 보완결과 불만족 시, 품질검증기관은 사업관리기관에 이를 알리고, 사업관리기관은 사업수행기관에 재보완 요구
8	결과 확인	사업관리기관 사업수행기관	<ul style="list-style-type: none"> 사업관리기관과 사업수행기관은 품질검증기관의 보완조치 확인 결과 공문 수령 및 세부내역(엑셀) 최종 확인

〈표 II-46〉 품질 보완조치 세부절차(다양성, 구문 정확성, 의미 정확성)

- 유효성 보완조치 세부절차



[그림 II-37] 품질 보완조치 세부절차(유효성)

번호	활동	주체	주요 내용
1	보완조치 시행 요구	사업관리기관	<ul style="list-style-type: none"> 사업관리기관은 품질검증 결과에 따라 사업수행기관에 품질 보완조치 시행 요구 공문 발송 품질 보완조치 시행 요구 시, 아래 항목 포함 <ul style="list-style-type: none"> 품질 보완조치 대상 데이터/AI모델 보완 완료 기한 등
2	보완조치 세부사항 안내	품질검증기관	<ul style="list-style-type: none"> 사업수행기관은 품질검증기관에 연락하여 품질 보완조치에 대한 세부 안내 확인 <ul style="list-style-type: none"> 품질검증기관은 데이터 업로드 방법, 제출서류 등에 대해 안내
3	보완조치 수행	사업수행기관	<ul style="list-style-type: none"> 사업수행기관은 데이터, AI 모델, 관련 서류 등 보완 수행
4	보완조치 결과 확인요청	사업수행기관	<ul style="list-style-type: none"> 사업수행기관은 보완 완료된 결과물을 S3 버킷에 업로드 후 품질검증기관에 보완조치 결과 확인요청 공문 발송

번호	활동	주체	주요 내용
5	검사 수행 및 보완조치 확인	품질검증기관	<ul style="list-style-type: none"> • 품질검증기관은 사업수행기관이 보완한 결과 및 제출서류를 기반으로 검사 수행 및 보완조치 확인 <ul style="list-style-type: none"> - 보완결과 만족 시, 품질검증기관은 보완조치 확인 결과를 사업관리기관 및 사업수행기관에 공문으로 통보 - 보완결과 불만족 시, 품질검증기관은 사업관리기관에 이를 알리고, 사업관리기관은 사업수행기관에 재보완 요구
6	결과 확인	사업관리기관 사업수행기관	<ul style="list-style-type: none"> • 사업관리기관과 사업수행기관은 품질검증기관의 보완조치 확인 결과 공문 수령 및 세부내역(엑셀) 최종 확인

〈표 II-47〉 품질 보완조치 세부절차(유효성)



부 록

부록 1. 품질관리 기준

부록 2. 품질지표 설정 가이드

부록 3. 개인정보보호 가이드

부록 1 | 품질관리 기준

1 개요

1.1 품질관리 지표 정의

- 한국지능정보사회진흥원이 추진하는 ‘인공지능 학습용 데이터 구축사업’과 인공지능 통합플랫폼 (AI Hub)를 통해 서비스하는 학습데이터 전반의 품질관리를 위한 기준이 필요하며, 다양한 인공지능 학습용 데이터를 수집·가공·활용하는 특성을 반영하여 범용적이고 공통적인 지표로 인공지능 학습용 데이터에 대한 품질관리를 강화하고, 지속적인 품질관리가 가능할 수 있도록 기준 마련하였음
- 목적에 맞게 데이터가 구축되었는지 검사하기 위해, 검사 대상과 방식을 결정하는 것이 중요하며, 구축 목표에 따라 요구사항을 정의하고, 데이터 구축 생애주기에 따라 최적의 품질을 확보할 수 있도록 검사와 개선이 필요하고, 이를 위해 데이터 품질관리 지표를 정의해야 하는데, 데이터 품질관리 지표는 추후 구축된 데이터를 활용할 수요자의 요구사항을 달성하는데 필요한 요소가 잘 갖추어졌는지 확인하기 위해 데이터 품질을 관리하는 기준임
- 따라서, 데이터 품질관리 지표는 “데이터의 품질 수준을 측정하기 위한 관점을 정의한 것으로 무엇을 측정할 것인가에 대한 기준”이라고 정의할 수 있음

1.2 품질관리 지표 구성

- 품질관리 지표는 데이터 생애주기 분석, 인공지능 학습용 데이터 구축 및 품질관점에서 일치성 분석, 데이터 품질관리 기준 분석을 통해 구축 및 활용관점을 반영한 준비성, 완전성, 유용성, 기준 적합성, 기술 적합성, 통계적 다양성, 구문 정확성, 의미 정확성, 알고리즘 적정성, 유효성 등 10가지 지표로 구성
- 수립된 데이터 품질관리의 10개 지표는 데이터 생애주기인 계획, 구축, 운영·활용의 전 단계의 품질관리에 활용하여 데이터 품질 수준을 확보하는 기준이 됨

구분	품질 지표	설명
구축과정	준비성	• 인공지능 학습용 데이터 품질관리를 위해 기본적으로 관리해야 하는 정책, 규정(저작권, 초상권, 개인정보보호 및 정보보호 등에 대한 검토 결과를 포함), 조직, 절차 등을 마련하고, 최신의 내용으로 충실하게 관리되는지를 검사하는 지표
	완전성	• 인공지능 학습용 데이터를 구축함에 있어 물리적인 구조를 갖추고, 정의한 데이터 형식 및 입력값 범위에 맞게 데이터가 저장되도록 설계·구축되었는지를 검사하는 지표
	유용성	• 발주기관(수요자)의 요구사항이 충분히 반영되었는지, 임무정의에 적합한 인공지능 학습용 데이터의 범위와 상세화 정도를 충족시키는지 검사하는 지표
데이터 적합성	기준 적합성	• 구축 데이터가 학습용도로써 적합한지 기준을 선정하기 위해, 다양성, 신뢰성, 충분성, 사실성을 측정하는 지표
	기술 적합성	• 구축 데이터가 학습용도로써 적합한지 기술적으로 판단하기 위해, 파일포맷, 해상도, 선명도, 컬러, 크기, 길이, 음질 등을 측정하는 지표
	통계적 다양성	• 데이터의 편향성을 방지하기 위해, 클래스 분포도, 인스턴스 분포도, 문장길이, 어휘 개수 등을 측정하는 지표
데이터 정확성	구문 정확성	• 어노테이션 데이터를 구성하는 속성 값들과 원래 정의한 데이터 형식 및 입력 값 범위와의 일치성을 측정하는 지표
	의미 정확성	• 데이터의 참값(Ground Truth)을 확인하기 위해, 정확도, 정밀도, 재현율, IoU, ROUGE, BLEU, F β -점수, EM 등을 측정하는 지표
학습모델	알고리즘 적정성	• 알고리즘을 Task 단위로 구분하여, 사업수행기관이 제시하는 학습모델의 Task가 적절한지 판단하는 지표
	유효성	• 학습용 데이터로 훈련시키는데 적합한 인공지능 알고리즘의 유효성을 측정하는 지표

〈표 III-1〉 품질관리 지표

1.3 품질검사 방법

- 데이터 품질검사 방법은 세부지표를 기준으로 검사 대상별로 정량적 검사와 정성적 검사를 수행하며, 세부지표와 검사대상 간 매핑을 통해 해당 지표별 검사 대상과 방법을 사용
- 품질관리 지표는 데이터 생애주기 분석, 인공지능 학습용 데이터 구축 및 품질 관점에서 일치성 분석, 데이터 품질관리 기준 분석을 통해 구축 및 활용관점을 반영한 준비성, 완전성, 유용성, 기준 적합성, 기술 적합성, 통계적 다양성, 구문 정확성, 의미 정확성, 알고리즘 적정성, 유효성 등 10가지 지표로 구성

품질지표			검사대상 데이터종류					검사대상 구축과정					검사 방법			
			동영상	이미지	텍스트	오디오	3D	임무 정의	구축 계획 수립	데이터 획득/ 수집	데이터 정제	데이터 가공	데이터 학습	정량적 검사	정성적 검사	
준비성	계획 수립성	절차준비						●	●	●	●	●			√	
		조직준비						●	●	●	●	●			√	
		도구준비							●	●	●	●			√	
		위험관리						●	●	●	●	●			√	
	체계 준수성	보안준수						●	●	●	●	●			√	
		법·제도 준수						●	●	●	●	●			√	
완전성	수집완전성								●						√	
	정제완전성									●					√	
	가공완전성										●				√	
유용성	사용편의성							●					●		√	
	유연성							●					●		√	
적합성	기준 적합성	다양성	●	●	●	●	●			●	●					√
		신뢰성	●	●	●	●	●			●	●					√
		충분성	●	●	●	●	●			●	●					√
		균일성	●	●	●	●	●			●	●					√
		사실성	●	●	●	●	●			●	●					√
		공평성	●	●	●	●	●			●	●					√
	기술 적합성	파일포맷	●	●	●	●	●								√	
		해상도	●	●											√	
		프레임 레이트	●												√	
		컬러심도	●	●											√	
		어절수			●	●									√	
	통계적 다양성	클래스 분포	●	●	●	●	●								√	
		인스턴스 분포	●	●	●	●	●								√	
		문장길이			●	●									√	
		어휘개수			●	●									√	
정확성	구문 정확성	데이터 구조	●	●	●	●	●								√	
		입력값 범위	●	●	●	●	●								√	
		데이터 형식	●	●	●	●	●								√	
	의미 정확성	정확도	●	●	●	●	●								√	
		정밀도	●	●	●		●								√	
		재현율	●	●	●		●								√	
		※ 정확도, 정밀도, 재현율 지표 이외에 IoU, ROUGE, BLEU, F β -점수, EM 등등의 지표를 과업 목표에 따라 적합하게 선정													√	
유효성	학습 모델 유효성	데이터 종류별 알고리즘 Task 분야	※ 과업 목표에 따라 적합한 품질지표 선정 필요											√		

〈표 III-2〉 품질지표별 검사대상 및 검사 방법

2 구축공정 품질관리 지표

2.1 준비성

- ‘준비성’에 대한 세부지표는 ‘계획수립성’과 ‘체계준수성’으로 구성되며, 검사항목은 계획수립성은 절차준비, 조직준비, 도구준비, 위험관리의 4개 항목으로 구성되며, 체계준수성은 보안준수, 법·제도 준수의 2개 항목으로 구성

품질지표		검사항목
준비성	계획수립성	절차준비
		조직준비
		도구준비
		위험관리
	체계준수성	보안준수
		법·제도 준수

〈표 III-3〉 준비성 검사항목 구성

2.1.1 계획수립성 - 절차준비

2.1.1.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 인공지능 학습용 데이터 구축을 위한 절차가 체계적으로 수립하여 관리 및 수행하고 있는지 확인
검사 대상	산출물	• 사업수행계획서(구축계획서, 품질관리계획서) 등
	공정	• 임무정의, 구축계획수립, 데이터획득/수집, 데이터정제, 데이터가공
검사방법		• 체크리스트

〈표 III-4〉 계획수립성 내 절차준비 검사기준 및 방법

2.1.1.2 검사 내용

분류	체크리스트
임무정의	<ul style="list-style-type: none"> • 발주기관(수요자) 요구사항을 수집하였는가? • 발주기관(수요자) 요구사항을 분석하였는가? • 과제관련 유사 또는 동일한 연구 등에 대한 사전조사를 하였는가? • 과제에 적합한 인공지능 학습모델선정 방안을 마련하였는가? • 과제에 적합한 인공지능 학습모델선정 결과를 제시하였는가? • 과제에 적합한 인공지능 학습모델선정 시 제약사항, 한계점 등에 대한 내용을 제시하였는가? • 인공지능 학습모델 성능지표와 목표를 제시하였는가? • 인공지능 학습모델에 성능을 보장하기 위한 지표와 목표를 제시하였는가?(지표 특징, 선행 연구결과 등에 대한 조사) • 인공지능 학습모델에 성능을 보장하기 위한 지표와 목표에 대한 객관성 확보를 위한 근거를 제시하였는가? (전문가 자문 등)
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습용 데이터셋을 정의하였는가? • 인공지능 학습용 데이터셋(훈련용, 검증용, 시험용) 구성 비율을 제시하였는가? (예, 8:1:1 또는 기타 비율) • 인공지능 학습용 데이터셋 중 시험용 데이터셋 미활용을 위한 관리방안을 제시하였는가? (관리 및 운영 규정, 환경, 접근권한 등) • 원시, 원천, 라벨링, 학습모델 구축을 위한 환경 구성계획을 수립하였는가? • 인공지능 학습용 데이터 분류체계를 정의하였는가? • 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토계획을 수립하였는가? • 정의된 분류체계별로 데이터 획득/수집, 정제, 가공 계획을 수립하였는가? • 인공지능 학습모델 검토를 위한 초기데이터(데이터 규모 10%이내) 선정 방법 및 기준을 마련하였는가? • 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토결과 반영을 위한 절차를 마련하였는가?
데이터 획득/수집	<ul style="list-style-type: none"> • 데이터 선정을 위한 개인정보, 초상권, 저작권, 명예훼손, 보안 등과 관련된 법·제도적인 검토 절차를 마련하였는가? • 데이터 선정 시 수집 기술 및 방법에 대한 검토절차를 마련하였는가? ※ 『제2권. 데이터 구축 안내서 v3.0』의 구축 데이터 정의 참조 • 데이터 획득/수집을 위한 개인정보, 초상권, 저작권, 명예훼손, 보안 등과 관련된 법·제도적인 검토 절차를 마련하였는가? • 데이터 획득/수집 방법 및 기준을 마련하였는가? • 데이터 획득/수집 시 미확보 데이터 획득/수집을 위한 방안을 수립하였는가? • 데이터 획득/수집 기준변경에 대한 절차를 마련하였는가? • 데이터 획득/수집 방법에 대한 교육 및 훈련계획을 수립하였는가? • 데이터 획득/수집 결과에 대한 검수절차를 마련하였는가? • 데이터 획득/수집 결과에 대한 검수 교육 및 훈련계획을 수립하였는가? • 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토결과 반영을 위한 절차를 마련하였는가?

분류	체크리스트
데이터 정제	<ul style="list-style-type: none"> • 데이터 정제 시 개인정보보호 등 비식별화를 위한 기준과 절차를 마련하였는가? • 데이터 정제 방법 및 기준을 마련하였는가? • 데이터 정제 기준변경에 대한 절차를 마련하였는가? • 데이터 정제 방법에 대한 교육 및 훈련계획을 수립하였는가? • 데이터 정제 결과에 대한 검수절차와 기준을 마련하였는가? • 데이터 정제 결과에 대한 검수 기준변경 절차를 마련하였는가? • 데이터 정제 결과에 대한 검수 교육 및 훈련계획을 수립하였는가?
데이터 가공	<ul style="list-style-type: none"> • 데이터 라벨링 방법 및 기준을 마련하였는가? • 데이터 라벨링 기준변경에 대한 절차를 마련하였는가? • 데이터 라벨링 방법에 대한 교육 및 훈련계획을 수립하였는가? • 데이터 라벨링 결과에 대한 검수절차를 마련하였는가? • 데이터 라벨링 결과에 대한 검수 기준변경 절차를 마련하였는가? • 데이터 라벨링 결과에 대한 검수 교육 및 훈련계획을 수립하였는가? • 인공지능 학습모델에 라벨링데이터가 적합한지 검사를 위한 데이터 샘플링 방안을 마련하였는가?

〈표 III-5〉 계획수립성 내 절차준비 검사 내용

2.1.2 계획수립성 - 조직준비

2.1.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 인공지능 학습용 데이터 구축을 위한 조직구성, 역할 및 책임을 체계적으로 수립하여 관리 및 운영하고 있는지 확인
검사 대상	산출물	• 사업수행계획서, 구축계획서, 품질관리계획서 등
	공정	• 구축계획수립, 데이터 획득/수집, 데이터정제, 데이터가공
검사방법		• 체크리스트

〈표 III-6〉 계획수립성 내 조직준비 검사기준 및 방법

2.1.2.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토를 위한 조직의 역할과 책임을 정의하였는가? • 인공지능 학습용 데이터셋 중 시험용 데이터셋 미활용을 위한 관리자를 지정하였는가?

분류	체크리스트
데이터 획득/수집	<ul style="list-style-type: none"> • 데이터 획득/수집을 위한 관련 법·제도적인 검토를 위한 전담조직을 구성하였는가? • 데이터 획득/수집을 위한 조직을 구성하고 담당자를 명시 하였는가? • 데이터 획득/수집을 위한 인력운영 계획을 수립하였는가? • 데이터 획득/수집을 위한 조직의 역할과 책임을 정의하였는가? • 데이터 획득/수집 방법에 대한 교육 및 훈련을 위한 조직을 구성하고 담당자를 명시 하였는가? • 데이터 획득/수집 방법에 대한 교육 및 훈련을 위한 조직 역할과 책임을 정의하였는가? • 데이터 획득/수집 결과에 대한 검수를 위한 조직의 역할과 책임을 정의하였는가?
데이터 정제	<ul style="list-style-type: none"> • 데이터 정제 시 개인정보보호 등 비식별화를 위한 기준 마련을 위한 전담조직을 구성하였는가? • 데이터 정제를 위한 조직을 구성하고 담당자를 명시하였는가? • 데이터 정제를 위한 인력운영 계획을 수립하였는가? • 데이터 정제를 위한 조직의 역할과 책임을 정의하였는가? • 데이터 정제 방법에 대한 교육 및 훈련을 위한 조직을 구성하고 담당자를 명시하였는가? • 데이터 정제 방법에 대한 교육 및 훈련을 위한 조직 역할과 책임을 정의하였는가? • 데이터 정제 결과에 대한 검수를 위한 조직의 역할과 책임을 정의하였는가?
데이터 가공	<ul style="list-style-type: none"> • 데이터 라벨링을 위한 조직을 구성하고 담당자를 명시 하였는가? • 데이터 라벨링을 위한 인력운영 계획을 수립하였는가? • 데이터 라벨링을 위한 조직의 역할과 책임을 정의하였는가? • 데이터 라벨링 방법에 대한 교육 및 훈련을 위한 조직을 구성하고 담당자를 명시 하였는가? • 데이터 라벨링 방법에 대한 교육 및 훈련을 위한 조직의 역할과 책임을 정의하였는가? • 데이터 라벨링 결과에 대한 검수를 위한 조직의 역할과 책임을 정의하였는가?

〈표 III-7〉 계획수립성 내 조직준비 검사 내용

2.1.3 계획수립성 - 도구준비

2.1.3.1 검사기준 및 방법

분류	검사기준 및 방법
검사상세	<ul style="list-style-type: none"> • 인공지능 학습용 데이터 구축을 위한 도구와 환경 구성 계획을 수립하여 관리 및 운영하고 있는지 확인
검사 대상	산출물
	공정
검사방법	<ul style="list-style-type: none"> • 체크리스트

〈표 III-8〉 계획수립성 내 도구준비 검사기준 및 방법

2.1.3.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습모델 선정에 따른 인공지능 학습모델이 정의되어 있는가? • 요구사항 변화에 따른 후보 인공지능 학습모델을 제시하고 있는가?
데이터 획득/수집	<ul style="list-style-type: none"> • 데이터 획득/수집을 위한 작업 도구를 정의하였는가? • 데이터 획득/수집을 위한 작업 도구 활용을 위한 사용자/관리자 매뉴얼이 있는가? • 데이터 획득/수집을 위한 작업 도구에 대한 교육 및 훈련 계획을 수립하였는가? • 데이터 획득/수집을 위한 작업 도구에 대한 교육 및 훈련을 실시하였는가? • 데이터 획득/수집을 위한 작업 도구에 대한 법·제도적(저작권 등) 검토를 하였는가?
데이터 정제	<ul style="list-style-type: none"> • 데이터 정제를 위한 작업 도구를 정의하였는가? • 데이터 정제를 위한 작업 도구 활용을 위한 사용자/관리자 매뉴얼이 있는가? • 데이터 정제를 위한 작업 도구에 대한 교육 및 훈련 계획을 수립하였는가? • 데이터 정제를 위한 작업 도구에 대한 교육 및 훈련을 실시하였는가? • 데이터 정제를 위한 작업 도구에 대한 법·제도적(저작권 등) 검토를 하였는가?
데이터 가공	<ul style="list-style-type: none"> • 데이터 라벨링을 위한 작업 도구를 정의하였는가? • 데이터 라벨링을 위한 작업 도구 활용을 위한 사용자/관리자 매뉴얼이 있는가? • 데이터 라벨링을 위한 작업 도구에 대한 교육 및 훈련 계획을 수립하였는가? • 데이터 라벨링을 위한 작업 도구에 대한 교육 및 훈련을 실시하였는가? • 데이터 라벨링을 위한 작업 도구에 대한 법·제도적(저작권 등) 검토를 하였는가?

〈표 Ⅲ-9〉 계획수립성 내 도구준비 검사 내용

2.1.4 계획수립성 - 위험관리

2.1.4.1 검사기준 및 방법

분류	검사기준 및 방법
검사상세	<ul style="list-style-type: none"> • 인공지능 학습용 데이터 구축을 위한 위험관리체계를 수립하여 관리 및 수행하고 있는지 확인
검사 대상	<div>산출물</div> <ul style="list-style-type: none"> • 사업수행계획서/구축계획서/품질관리계획서 등 <div>공정</div> <ul style="list-style-type: none"> • 구축계획수립/데이터획득/수집/데이터정제/데이터가공
검사방법	<ul style="list-style-type: none"> • 체크리스트

〈표 Ⅲ-10〉 계획수립성 내 위험관리 검사기준 및 방법

2.1.4.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 사업의 위험관리를 위한 계획을 수립하였는가? • 사업의 위험관리를 위한 위험요소를 식별하였는가? • 식별된 위험대응을 위한 활동을 수행하고 있는가?

분류	체크리스트
데이터 획득/수집	<ul style="list-style-type: none"> 데이터 획득/수집 시 위험관리를 위한 계획을 수립하였는가? 데이터 획득/수집 시 위험관리를 위한 위험요소를 식별하였는가? 식별된 위험대응을 위한 활동을 수행하고 있는가?
데이터 정제	<ul style="list-style-type: none"> 데이터 정제 시 위험관리를 위한 계획을 수립하였는가? 데이터 정제 시 위험관리를 위한 위험요소를 식별하였는가? 식별된 위험대응을 위한 활동을 수행하고 있는가?
데이터 가공	<ul style="list-style-type: none"> 데이터 라벨링 시 위험관리를 위한 계획을 수립하였는가? 데이터 라벨링 시 위험관리를 위한 위험요소를 식별하였는가? 식별된 위험대응을 위한 활동을 수행하고 있는가?

〈표 III-11〉 계획수립성 내 위험관리 검사 내용

2.1.5 체계준수성 - 보안준수

2.1.5.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 개인정보보호, 비식별화 조치 등 민감정보 보호와 인공지능 학습용 데이터 보안 관리를 위한 절차와 규정을 수립하여 관리 및 수행하고 있는지 확인
검사 대상	산출물	• 사업수행계획서/구축계획서/품질관리계획서 등
	공정	• 구축계획수립/데이터획득/수집/데이터정제/데이터가공
검사방법		• 체크리스트

〈표 III-12〉 체계준수성 내 보안준수 검사기준 및 방법

2.1.5.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> 인공지능 학습용 데이터셋(훈련용, 검증용, 시험용)에 대한 보안 관리체계를 마련하였는가? (관리 및 운영 규정, 환경, 접근권한 등) 민감정보 보호를 위한 체계를 마련하였는가? (관리 및 운영 규정, 환경, 접근권한 등) ※ 개인정보보호 및 보안을 위한 관계법령, 고시·권고사항 준수
데이터 획득/수집	<ul style="list-style-type: none"> 데이터 획득/수집에 대한 보안관리를 수행하고 있는가?(담당자지정, 운영환경구성, 접근권한관리 등) 민감정보 보호를 위한 활동을 수행하고 있는가? (담당자지정, 로그관리 등) ※ 개인정보보호 및 보안을 위한 관계법령, 고시·권고사항 준수
데이터 정제	<ul style="list-style-type: none"> 데이터 정제 시 개인정보보호 등 비식별화를 수행하고 있는가? 데이터 정제에 대한 보안관리를 수행하고 있는가?(담당자지정, 운영환경구성, 접근권한관리 등) ※ 개인정보보호 및 보안을 위한 관계법령, 고시·권고사항 준수
데이터 가공	<ul style="list-style-type: none"> 데이터 라벨링에 대한 보안관리를 수행하고 있는가? (담당자 지정, 운영환경구성, 접근권한 관리 등) ※ 개인정보보호 및 보안을 위한 관계법령, 고시·권고사항 준수

〈표 III-13〉 체계준수성 내 보안준수 검사 내용

2.1.6 체계준수성 - 법·제도 준수

2.1.6.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 인공지능 학습용 데이터 구축 시 관련 법·제도(개인정보, 초상권, 저작권 등)에 대한 충분한 검토와 해결방안이 제시되었는지 검사
검사 대상	산출물	• 사업수행계획서/구축계획서/품질관리계획서 등
	공정	• 구축계획수립/데이터획득/수집/데이터정제/데이터가공
검사방법		• 체크리스트

〈표 III-14〉 체계준수성 내 법·제도 준수 검사기준 및 방법

2.1.6.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습용 데이터 구축을 위한 관련 법·제도적인 검토를 위한 절차 및 해결방안을 제시하는가? • 인공지능 학습용 데이터 구축을 위한 개인정보활용 동의 절차를 마련하고 수행하고 있는가? • 인공지능 학습용 데이터 구축을 위한 초상권 활용 동의 절차를 마련하고 수행하고 있는가? • 인공지능 학습용 데이터 구축 시 명예훼손 가능성 여부 검토 절차를 마련하고 수행하고 있는가? ※ 개인정보보호 및 보안을 위한 관계법령, 저작권 및 초상권관련 관계법령준수
데이터 획득/수집	<ul style="list-style-type: none"> • 데이터 획득/수집을 위한 관련 법·제도적인 검토를 위한 절차 및 해결방안을 제시하는가? • 데이터 획득/수집 시 저작권 보호 대상인 경우 법에 저촉되지 않는 범위 내에서 수집할 수 있는 방안을 마련하였는가? • 데이터 획득/수집 시 저작권 보호 대상 저작물 활용에 따른 동의절차를 마련하고 수행하는가? • 데이터 획득/수집 시 저작권 보호 대상 저작물 활용에 따른 계약절차를 마련하고 수행하는가? ※ 개인정보보호 및 보안을 위한 관계법령, 저작권 및 초상권관련 관계법령준수
데이터 정제	<ul style="list-style-type: none"> • 데이터 정제를 위한 관련 민감정보 비식별화 조치 등 법·제도적인 검토 절차 및 해결방안을 제시하는가? • 데이터 정제를 위한 비식별화 기법 적용하는 방안을 마련하였는가? ※ 개인정보보호 및 보안을 위한 관계법령, 저작권 및 초상권관련 관계법령준수
데이터 가공	<ul style="list-style-type: none"> • 데이터 라벨링을 위한 법·제도적인 검토 절차 및 해결방안을 제시하는가? ※ 개인정보보호 및 보안을 위한 관계법령, 저작권 및 초상권관련 관계법령준수

〈표 III-15〉 체계준수성 내 법·제도 준수 검사 내용

2.2 완전성

- ‘완전성’에 대한 검사항목은 ‘수집완전성’, ‘정제완전성’, ‘가공완전성’으로 구성

품질지표	검사항목
완전성	수집완전성
	정제완전성
	가공완전성

〈표 III-16〉 완전성 검사항목 구성

2.2.1 수집완전성

2.2.1.1 검사기준 및 방법

분류	검사기준 및 방법	
검사상세	<ul style="list-style-type: none"> 데이터의 편향성 방지방안을 수립하여 관리 및 수행하고 있는지 확인 데이터 획득/수집 방법 및 기준, 교육, 검수에 대한 체계를 수립하여 관리 및 수행하고 있는지 확인 데이터 획득/수집 기준변경을 위한 체계를 수립하여 관리 및 수행하고 있는지 확인 	
검사대상	산출물	구축계획서/품질관리계획서/구축 가이드라인 등
	공정	데이터획득/수집
검사방법	체크리스트	

〈표 III-17〉 수집완전성 검사기준 및 방법

2.2.1.2 검사 내용

분류	체크리스트
데이터 획득/수집	<ul style="list-style-type: none"> 편향성 방지 방안을 마련하였는가? 정의된 데이터 획득/수집 방법 및 기준을 적용하고 있는가? 데이터 획득/수집 기준변경에 대한 절차를 마련하였는가? 데이터 획득/수집 방법에 대한 교육 및 훈련을 실시하였는가? 데이터 획득/수집 결과에 대한 검수절차 및 기준에 따라 수행하고 있는가? 데이터 획득/수집 결과에 대한 검수 기준변경 시 절차에 따라 수행하였는가? 데이터 획득/수집 결과에 대한 검수 교육 및 훈련을 실시하였는가? 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토를 위한 데이터 선별 방안에 따라 선별하였는가?

〈표 III-18〉 수집완전성 검사 내용

2.2.2 정제완전성

2.2.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		<ul style="list-style-type: none"> 개인정보보호를 위한 비식별화 조치 등 민감정보 보호를 위한 절차와 규정을 수립하여 관리 및 수행하고 있는지 확인 데이터 정제 방법 및 기준, 교육, 검수에 대한 체계를 수립하여 관리 및 수행하고 있는지 확인 데이터 정제 기준변경을 위한 체계를 수립하여 관리 및 수행하고 있는지 확인
검사 대상	산출물	• 품질관리계획서/구축 가이드라인 등
	공정	• 데이터정제
검사방법		• 체크리스트

〈표 Ⅲ-19〉 정제완전성 검사기준 및 방법

2.2.2.2 검사 내용

분류	체크리스트
데이터 정제	<ul style="list-style-type: none"> 데이터 정제 시 개인정보보호 등 비식별화를 실시하였는가? 정의된 데이터 정제 방법 및 기준을 적용하고 있는가? 데이터 정제 기준변경에 대한 절차를 마련하였는가? 데이터 정제 방법에 대한 교육 및 훈련을 실시하였는가? 데이터 정제 결과에 대한 검수 절차 및 기준에 따라 수행하고 있는가? 데이터 정제 결과에 대한 검수 기준변경 시 절차에 따라 수행하였는가? 데이터 정제 결과에 대한 검수 교육 및 훈련을 실시하였는가? 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토를 위한 데이터 선별 방안에 따라 선별하였는가?

〈표 Ⅲ-20〉 정제완전성 검사 내용

2.2.3 가공완전성

2.2.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		<ul style="list-style-type: none"> 데이터 라벨링 방법 및 기준, 교육, 검수에 대한 체계를 수립하여 관리 및 수행하고 있는지 확인 데이터 라벨링 기준 변경을 위한 체계를 수립하여 관리 및 수행하고 있는지 확인
검사 대상	산출물	• 구축계획서/품질관리계획서/구축 가이드라인 등
	공정	• 데이터가공
검사방법		• 체크리스트

〈표 Ⅲ-21〉 가공완전성 검사기준 및 방법

2.2.3.2 검사 내용

분류	체크리스트
데이터 가공	<ul style="list-style-type: none"> • 정의된 데이터 라벨링 방법 및 기준을 적용하고 있는가? • 데이터 라벨링 기준변경에 대한 절차를 마련하였는가? • 데이터 라벨링 방법에 대한 교육 및 훈련을 실시하였는가? • 데이터 라벨링 결과에 대한 검수 절차 및 기준에 따라 수행하고 있는가? • 데이터 라벨링 결과에 대한 검수 기준변경 시 절차에 따라 수행하였는가? • 데이터 라벨링 결과에 대한 검수 교육 및 훈련을 실시하였는가? • 인공지능 학습모델에 적합한지 초기데이터(데이터 규모 10%이내)를 활용한 인공지능 학습모델 검토를 위한 데이터 선별 방안에 따라 선별하였는가?

〈표 III-22〉 가공완전성 검사 내용

2.3 유용성

- ‘유용성’에 대한 검사항목은 ‘사용편의성’, ‘유연성’으로 구성

품질지표	검사항목
유용성	사용편의성
	유연성

〈표 III-23〉 유용성 검사항목 구성

2.3.1 사용편의성

2.3.1.1 검사기준 및 방법

분류	검사기준 및 방법
검사상세	<ul style="list-style-type: none"> • 사용자의 요구사항을 만족 시킬 수 있는 학습모델과 학습용 데이터셋이 제공되고 있는지 확인
검사 대상	산출물
	공정
검사방법	<ul style="list-style-type: none"> • 체크리스트

〈표 III-24〉 사용편의성 검사기준 및 방법

2.3.1.2 검사 내용

분류	체크리스트
임무정의	<ul style="list-style-type: none"> 발주자(수요자)와 협의된 결과를 해결하는데 적합한 인공지능 학습모델에 적합한 데이터 셋인가?
데이터학습	<ul style="list-style-type: none"> 인공지능 학습모델이 성능지표를 만족하는가? 인공지능 학습모델이 훈련용 및 검증용 데이터셋을 활용한 결과 성능지표 및 목표를 만족하는가? 인공지능 학습모델이 시험용 데이터셋을 활용한 결과 성능지표 및 목표를 만족하는가? (수행기관 제외) 사용자가 인공지능 학습모델에 사용하기 위한 데이터셋을 검색 및 활용하기 위한 기본정보를 제공하는가? (학습형태 등)

〈표 III-25〉 사용편의성 검사 내용

2.3.2 유연성

2.3.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 사용자 요구사항을 수용할 수 있는 학습모델 선정 및 적용이 유연한지 확인
검사 대상	산출물	• 구축 가이드라인
	공정	• 데이터학습
검사방법		• 체크리스트

〈표 III-26〉 유연성 검사기준 및 방법

2.3.2.2 검사 내용

분류	체크리스트
임무정의	<ul style="list-style-type: none"> 인공지능 학습모델 성능지표를 만족하지 못하는 경우 대안을 제시하고 있는가?
데이터학습	<ul style="list-style-type: none"> 인공지능 학습용 데이터 변경사항 및 이력이 관리되고 있는가?

〈표 III-27〉 유연성 검사 내용

3 데이터 적합성 품질관리 지표

3.1 기준 적합성

- ‘기준 적합성’에 대한 검사항목은 ‘다양성’, ‘신뢰성’, ‘충분성’, ‘균일성’, ‘사실성’, ‘공평성’으로 구성

품질지표	검사항목
기준 적합성	다양성
	신뢰성
	충분성
	균일성
	사실성
	공평성

〈표 III-28〉 기준 적합성 검사항목 구성

3.1.1 다양성

3.1.1.1 검사기준 및 방법

분류	검사기준 및 방법
검사상세	• 공간, 시간 등 환경조건 및 대상 객체 종류와 속성의 변화 정도를 확인
검사 대상	산출물 • 구축 가이드라인/원시데이터
	데이터 종류 • 텍스트/이미지/비디오/오디오/3D
검사방법	• 체크리스트/전수검사

〈표 III-29〉 다양성 검사기준 및 방법

3.1.1.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> 인공지능이 처리해야 하는 실제 세상의 데이터(사물, 사람, 장소, 시간, 환경, 언어 특성 등)와 유사한 특성(예시 참조)이 데이터에 반영되도록 계획을 수립하였는가? 〈예시〉 분야별 데이터 특성(포괄성)

분류	체크리스트												
	<table><tr><th>분야</th><th>설명</th></tr><tr><td>자율주행 데이터</td><td>• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함</td></tr><tr><td>어류행동 데이터</td><td>• 비디오 외에 생육단계, 수질정보 등 포함</td></tr><tr><td>음성 데이터</td><td>• 발화자의 연령, 성별, 지역정보 등 포함</td></tr></table> <p>※ 분야별 특성 정보는 지속적으로 추가 필요</p> <p>• 인공지능이 처리해야 하는 실제 세상의 데이터(사물, 사람, 장소, 시간, 환경, 언어 특성 등)와 유사한 변동성(예시 참조)을 데이터가 갖도록 계획을 수립하였는가? 〈예시〉 분야별 데이터 변동성</p> <table><tr><th>분야</th><th>설명</th></tr><tr><td>자율주행 데이터</td><td>• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영</td></tr></table> <p>※ 분야별 특성 정보는 지속적으로 추가 필요</p>	분야	설명	자율주행 데이터	• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함	어류행동 데이터	• 비디오 외에 생육단계, 수질정보 등 포함	음성 데이터	• 발화자의 연령, 성별, 지역정보 등 포함	분야	설명	자율주행 데이터	• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영
	분야	설명											
	자율주행 데이터	• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함											
	어류행동 데이터	• 비디오 외에 생육단계, 수질정보 등 포함											
	음성 데이터	• 발화자의 연령, 성별, 지역정보 등 포함											
분야	설명												
자율주행 데이터	• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영												
데이터 획득/수집	<p>• 인공지능이 처리해야 하는 실제 세상의 데이터(사물, 사람, 장소, 시간, 환경, 언어 특성 등)와 유사한 특성(예시 참조)이 데이터에 반영되었는가? 〈예시〉 분야별 데이터 특성(포괄성)</p> <table><tr><th>분야</th><th>설명</th></tr><tr><td>자율주행 데이터</td><td>• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함</td></tr><tr><td>어류행동 데이터</td><td>• 비디오 외에 생육단계, 수질정보 등 포함</td></tr><tr><td>음성 데이터</td><td>• 발화자의 연령, 성별, 지역정보 등 포함</td></tr></table> <p>※ 분야별 특성 정보는 지속적으로 추가 필요</p> <p>• 인공지능이 처리해야 하는 실제 세상의 데이터(사물, 사람, 장소, 시간, 환경, 언어 특성 등)와 유사한 변동성(예시 참조)을 데이터가 갖고 있는가? 〈예시〉 분야별 데이터 변동성</p> <table><tr><th>분야</th><th>설명</th></tr><tr><td>자율주행 데이터</td><td>• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영</td></tr></table> <p>※ 분야별 특성 정보는 지속적으로 추가 필요</p>	분야	설명	자율주행 데이터	• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함	어류행동 데이터	• 비디오 외에 생육단계, 수질정보 등 포함	음성 데이터	• 발화자의 연령, 성별, 지역정보 등 포함	분야	설명	자율주행 데이터	• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영
	분야	설명											
	자율주행 데이터	• 위치정보(GPS), 촬영시간, 도로종류, 차량 속도 등 포함											
	어류행동 데이터	• 비디오 외에 생육단계, 수질정보 등 포함											
	음성 데이터	• 발화자의 연령, 성별, 지역정보 등 포함											
분야	설명												
자율주행 데이터	• 넓은 범위의 지역을 대상으로 낮, 밤, 우천, 눈 등 조건하에서 도심도로, 고속도로, 외곽/시골도로를 중심으로 저속, 중속, 고속 주행 환경에서 촬영												

〈표 Ⅲ-30〉 다양성 검사 내용

3.1.2 신뢰성

3.1.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 데이터가 신뢰할 수 있는 사람, 기관, 기업으로부터 수집되었는지를 확인
검사 대상	산출물	• 구축 가이드라인/원시데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
검사방법		• 체크리스트

〈표 Ⅲ-31〉 신뢰성 검사기준 및 방법

3.1.2.2 검사 내용

분류	체크리스트
구축 계획수립	• 데이터 획득/수집 시 획득/수집 출처의 객관성 확보를 위한 계획을 수립하였는가?(출처예시, 사람, 기관, 기업 등)
데이터 획득/수집	• 데이터 획득/수집을 위한 획득/수집 출처에 대한 객관성 확보를 위한 근거를 제시하였는가?

〈표 Ⅲ-32〉 신뢰성 검사 내용

3.1.3 충분성

3.1.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 카테고리(분류체계) 및 인스턴스(분류체계별 특성을 갖고 있는 데이터) 등이 학습에 유용한 수량인지를 확인
검사 대상	산출물	• 구축 가이드라인/원시데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
검사방법		• 체크리스트

〈표 Ⅲ-33〉 충분성 검사기준 및 방법

3.1.3.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습모델에 필요한 분류체계 및 분류체계별 데이터 획득/수집 최소 수량 결정을 위한 절차를 마련하였는가? • 인공지능 학습모델에 필요한 분류체계 및 분류체계별 데이터 획득/수집 최소 수량 결정에 대한 근거를 제시하였는가?
데이터 획득/수집	• 인공지능 학습모델에 필요한 분류체계 및 분류체계별 데이터 획득/수집 최소 수량을 확보하였는가?(예시, MS COCO 데이터셋은 객체 당 5,000개 이상 인스턴스 수집)

〈표 Ⅲ-34〉 충분성 검사 내용

3.1.4 균일성

3.1.4.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 분류/탐지/인식/이해/예측 카테고리(분류체계)별 인스턴스(분류체계별 특성을 갖고 있는 데이터) 수량의 균일성 및 비율이 고려되었는지 확인
검사 대상	산출물	• 구축 가이드라인/원시데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
검사방법		• 체크리스트

〈표 Ⅲ-35〉 균일성 검사기준 및 방법

3.1.4.2 검사 내용

분류	체크리스트
구축 계획수립	<ul style="list-style-type: none"> • 인공지능 학습모델에 필요한 분류체계별 데이터 획득/수집 수량에 대한 적합한 비율 결정을 위한 절차를 마련하였는가? • 인공지능 학습모델에 필요한 분류체계별 데이터 획득/수집 수량에 대한 적합한 비율 결정에 대한 근거를 제시하였는가?
데이터 획득/수집	<ul style="list-style-type: none"> • 인공지능 학습모델에 필요한 분류체계별 데이터 획득/수집 수량에 대한 적합한 비율에 맞게 수량을 확보하였는가?

〈표 Ⅲ-36〉 균일성 검사 내용

3.1.5 사실성

3.1.5.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 분류/탐지/인식/이해/예측 카테고리(분류체계)별 인스턴스(분류체계별 특성을 갖고 있는 데이터) 수량의 균일성 및 비율이 고려되었는지 확인
검사 대상	산출물	• 구축 가이드라인/원시데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
검사방법		• 체크리스트

〈표 Ⅲ-37〉 사실성 검사기준 및 방법

3.1.5.2 검사 내용

분류	체크리스트										
구축 계획수립	<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 데이터 획득/수집이 인위적인 환경인 경우 실제 환경 및 상황의 특성을 반영하기 위한 계획을 수립하였는가? 인공지능 학습용 데이터 구축 시 데이터 획득/수집이 인위적인 환경인 경우 실세계 환경 및 조건이 일관성을 갖도록 계획을 수립하였는가? 데이터 획득/수집 시 다중 데이터 간에 동기화를 하도록 계획을 수립하였는가? 										
데이터 획득/수집	<ul style="list-style-type: none"> 데이터 획득/수집이 인위적인 환경인 경우 실제 환경 및 상황의 특성이 반영된 근거를 제시하였는가? 데이터 획득/수집이 인위적인 환경인 경우 실세계 환경 및 조건이 일관성이 확보된 근거를 제시하였는가? 데이터 획득/수집 시 다중 데이터 간에 동기화된 근거를 제시하였는가? <p>〈예시〉 데이터 획득/수집 시 동기화</p> <table> <tr> <th>분야</th><th>설명</th></tr> <tr> <td>자율주행</td><td>• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함</td></tr> <tr> <td>CCTV</td><td>• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함</td></tr> <tr> <td>영상, 음성, 자막 동시 수집</td><td> <ul style="list-style-type: none"> 영상과 음성을 동기화 해야됨 영상과 자막을 동기화 해야됨 영상, 음성, 자막을 동기화 해야됨 </td></tr> <tr> <td>컴퓨터 비전, 메타버스, 3D 프린팅 등</td><td> <ul style="list-style-type: none"> 이미지와 3D 모델링 데이터의 선형적 특징(직사각형, 원형, 마름모)이 유사해야 함 3D 모델링 데이터에 실제 객체의 부속장치가 모두 표현되어야 함 노멀(Normal) 방향이 정확해야 함 좌표 중심축을 기준으로 3D 모델링 데이터의 3차원 자세가 정확해야 함 메쉬(Mesh)내 홀(Hole)이 존재하지 않아야 함 메쉬(Mesh)와 텍스처(Textures)가 존재해야 함 3D 모델링 특징점간 규격이 실제 객체의 특징점간 규격과 일치해야 함 </td></tr> </table>	분야	설명	자율주행	• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함	CCTV	• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함	영상, 음성, 자막 동시 수집	<ul style="list-style-type: none"> 영상과 음성을 동기화 해야됨 영상과 자막을 동기화 해야됨 영상, 음성, 자막을 동기화 해야됨 	컴퓨터 비전, 메타버스, 3D 프린팅 등	<ul style="list-style-type: none"> 이미지와 3D 모델링 데이터의 선형적 특징(직사각형, 원형, 마름모)이 유사해야 함 3D 모델링 데이터에 실제 객체의 부속장치가 모두 표현되어야 함 노멀(Normal) 방향이 정확해야 함 좌표 중심축을 기준으로 3D 모델링 데이터의 3차원 자세가 정확해야 함 메쉬(Mesh)내 홀(Hole)이 존재하지 않아야 함 메쉬(Mesh)와 텍스처(Textures)가 존재해야 함 3D 모델링 특징점간 규격이 실제 객체의 특징점간 규격과 일치해야 함
분야	설명										
자율주행	• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함										
CCTV	• 비디오의 멀티카메라 촬영 영상은 시간동기화를 맞추어야 함										
영상, 음성, 자막 동시 수집	<ul style="list-style-type: none"> 영상과 음성을 동기화 해야됨 영상과 자막을 동기화 해야됨 영상, 음성, 자막을 동기화 해야됨 										
컴퓨터 비전, 메타버스, 3D 프린팅 등	<ul style="list-style-type: none"> 이미지와 3D 모델링 데이터의 선형적 특징(직사각형, 원형, 마름모)이 유사해야 함 3D 모델링 데이터에 실제 객체의 부속장치가 모두 표현되어야 함 노멀(Normal) 방향이 정확해야 함 좌표 중심축을 기준으로 3D 모델링 데이터의 3차원 자세가 정확해야 함 메쉬(Mesh)내 홀(Hole)이 존재하지 않아야 함 메쉬(Mesh)와 텍스처(Textures)가 존재해야 함 3D 모델링 특징점간 규격이 실제 객체의 특징점간 규격과 일치해야 함 										

〈표 Ⅲ-38〉 사실성 검사 내용

3.1.6 공평성

3.1.6.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 데이터 속에 포함될 수 있는 편향된 데이터를 제거하였는지 확인
검사 대상	산출물	• 구축 가이드라인/원시데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
검사방법		• 체크리스트

〈표 Ⅲ-39〉 공평성 검사기준 및 방법

3.1.6.2 검사 내용

분류	체크리스트
구축 계획수립	• 인공지능 학습용 데이터 구축 시 지역적 편견, 사회적 편견, 인종적 편견 등을 방지하기 위한 계획을 수립하였는가?
데이터 획득/수집	• 인공지능 학습용 데이터 구축 시 지역적 편견, 사회적 편견, 인종적 편견 등의 방지 결과에 대한 근거를 제시하였는가?

〈표 III-40〉 공평성 검사 내용

3.2 기술 적합성

- ‘기술 적합성’에 대한 검사항목은 ‘파일포맷’, ‘해상도’, ‘프레임 레이트’, ‘컬러심도’, ‘어절 수’로 구성

품질지표	검사항목
기술 적합성	파일포맷
	해상도
	프레임 레이트
	컬러심도
	어절 수

〈표 III-41〉 기술 적합성 검사항목 구성

3.2.1 파일포맷

3.2.1.1 검사기준 및 방법

분류		검사기준 및 방법		
검사상세		• 대중적으로 널리 사용되는 대표적인 파일 포맷을 사용하였는지 확인		
검사 대상	산출물	• 원시데이터		
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D		
	내용	• 값		
검사방법		• 인공지능 학습용 데이터 구축 시 정의된 파일 포맷을 원시데이터에 적용하고 있는지 전수조사를 통해 확인		
		〈데이터 유형 별 권장 파일 포맷〉		
		데이터 유형	파일 포맷	비고
		이미지	JPG, PNG, TIFF	• 의료 등 전문분야의 경우 해당분야 표준 준수
		비디오	MP4, AVI	• 무압축 방식으로 프레임 이미지 시퀀스의 묶음형태
		오디오	WAV, PCM	-
		텍스트	-	• UTF-8 인코딩 준수
		정량수치	CSV	• 산업용 센서의 경우 해당분야 표준 준수
		로그	JSON	• 웹표준 준수

〈표 III-42〉 파일포맷 검사기준 및 방법

3.2.1.2 계산식

구분	계산 수식
준수율 계산	$\text{준수율}(\%) = \frac{\text{준수 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-43〉 파일포맷 준수율 계산식

3.2.2 해상도

3.2.2.1 검사기준 및 방법

분류		검사기준 및 방법	
검사상세		• 유형별로 활용 목적에 적합한 해상도를 사용하고 있는지 확인	
검사 대상	산출물	• 원시데이터	
	데이터 종류	• 비디오/이미지	
	내용	• 값	
검사방법		• 인공지능 학습용 데이터 구축 시 정의된 해상도를 원시데이터에 적용하고 있는지 전수조사를 통해 확인	
		〈데이터 유형 별 권장 파일 포맷〉	
		데이터 유형	확인 방법
		비디오	• 인공지능 학습용 데이터 구축 시 정의한 수집 비디오 데이터 해상도(가로, 세로 픽셀수)를 준수하고 있는지 전수조사를 통해 확인 ※ 데이터 확인방식은 전수조사를 기본으로 함
	이미지	• 인공지능 학습용 데이터 구축 시 정의한 수집 이미지 데이터 해상도(가로, 세로 픽셀수 및 인치당 픽셀수 밀도)를 준수하고 있는지 전수조사를 통해 확인 ※ 데이터 확인방식은 전수조사를 기본으로 함	

〈표 III-44〉 해상도 검사기준 및 방법

3.2.2.2 계산식

구분	계산 수식
준수율 계산	$\text{준수율}(\%) = \frac{\text{준수 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-45〉 해상도 준수율 계산식

3.2.3 프레임 레이트

3.2.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 유형별로 활용 목적에 적합한 프레임 레이트를 사용하고 있는지 확인
검사 대상	산출물	• 원시데이터
	데이터 종류	• 비디오
	내용	• 값
검사방법		<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 정의한 수집 비디오 데이터 프레임 레이트(1초 당 사용되는 이미지가 몇 장인지 초당 프레임 수)를 준수하고 있는지 전수조사를 통해 확인 ※ 데이터 확인방식은 전수조사를 기본으로 함

〈표 III-46〉 프레임 레이트 검사기준 및 방법

3.2.3.2 계산식

구분	계산 수식
준수율 계산	$\text{준수율}(\%) = \frac{\text{준수 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-47〉 프레임 레이트 준수율 계산식

3.2.4 컬러심도

3.2.4.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 유형별로 활용 목적에 적합한 컬러심도(픽셀 당 비트수)를 사용하고 있는지 확인
검사 대상	산출물	• 원시데이터
	데이터 종류	• 비디오/이미지
	내용	• 값
검사방법		<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 정의한 수집 비디오 및 이미지 데이터에 대한 컬러심도(픽셀당 비트수)를 준수하고 있는지 전수조사를 통해 확인 ※ 데이터 확인방식은 전수조사를 기본으로 함

〈표 III-48〉 컬러심도 검사기준 및 방법

3.2.4.2 계산식

구분	계산 수식
준수율 계산	$\text{준수율}(\%) = \frac{\text{준수 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-49〉 컬러심도 준수율 계산식

3.2.5 어절 수

3.2.5.1 검사기준 및 방법

분류		검사기준 및 방법	
검사상세		• 유형별로 활용 목적에 적합한 어절 수를 사용하고 있는지 확인	
검사 대상	산출물	• 원시데이터	
	데이터 종류	• 텍스트/오디오	
	내용	• 값	
검사방법		• 인공지능 학습용 데이터 구축 시 정의한 텍스트 및 오디오 데이터 획득/수집 시 최소 어절 수 기준범위를 준수하고 있는지 전수조사를 통해 확인 ※ 데이터 확인방식은 전수조사를 기본으로 함	
		〈데이터 유형 별 권장 최소 기준 범위〉	
		분야	설명
		문어체	• 평균 15어절
		구어체(대화)	• 평균 5어절
※ 단, 권장 최소 기준 범위는 사업특성에 따라 전문가 자문 필요			

〈표 III-50〉 어절 수 검사기준 및 방법

3.2.5.2 계산식

구분	계산 수식
준수율 계산	$\text{준수율}(\%) = \frac{\text{준수 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-51〉 어절 수 준수율 계산식

3.3 통계적 다양성

- ‘통계적 다양성’에 대한 검사항목은 ‘클래스 분포’, ‘인스턴스 분포’, ‘문장길이’, ‘어휘개수’로 구성

품질지표	검사항목
통계적 다양성	클래스 분포
	인스턴스 분포
	문장길이
	어휘개수

〈표 III-52〉 통계적 다양성 검사항목 구성

3.3.1 클래스 분포

3.3.1.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 학습에 유용한 분류체계(카테고리)별 수량의 균일성 및 비율이 고려되었는지 검사
검사 대상	산출물	• 원천데이터/라벨링데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
	내용	• 값
검사방법		<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 정의한 분류체계(카테고리)를 기준으로 어노테이션 데이터에 포함된 클래스 정보를 추출하여 통계값 산출 ※ 클래스 분포 최소 기준은 전문가 자문필요

〈표 III-53〉 클래스 분포 검사기준 및 방법

3.3.1.2 계산식

구분	계산 수식
클래스 분포 계산	※ 계산 방법에 대한 전문가 자문필요

〈표 III-54〉 클래스 분포 계산식

3.3.2 인스턴스 분포

3.3.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 학습에 유용한 분류체계(카테고리)별 인스턴스(분류체계별 특성을 갖고 있는 데이터) 수량의 균일성 및 비율이 고려되었는지 확인
검사 대상	산출물	• 원천데이터/라벨링데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
	내용	• 값
검사방법		<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 정의한 분류체계(카테고리)별 인스턴스 정보를 기준으로 어노테이션 데이터에 포함된 클래스별 인스턴스 정보를 추출하여 통계값 산출 ※ 예시) 총분성 최소 기준 : 클래스 별 최소 1,000개 인스턴스 수집 ※ 인스턴스 분포 최소 기준은 전문가 자문필요

〈표 III-55〉 인스턴스 분포 검사기준 및 방법

3.3.2.2 계산식

구분	계산 수식
인스턴스 분포 계산	※ 계산 방법에 대한 전문가 자문필요

〈표 III-56〉 인스턴스 분포 계산식

3.3.3 문장길이

3.3.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 텍스트 및 오디오 데이터의 문장길이가 학습에 유용한 수량과 수량의 균일성 및 비율이 고려되었는지 확인
검사 대상	산출물	• 원천데이터/라벨링데이터
	데이터 종류	• 텍스트/오디오
	내용	• 값
검사방법		• 인공지능 학습용 데이터 구축 시 정의한 문장길이를 기준으로 어노테이션 데이터에 포함된 문장길이 정보를 추출하여 통계값 산출

〈표 III-57〉 문장길이 검사기준 및 방법

3.3.2.2 계산식

구분	계산 수식
문장길이 분포 계산	※ 계산 방법에 대한 전문가 자문필요

〈표 III-58〉 문장길이 분포 계산식

3.3.4 어휘개수

3.3.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 텍스트 및 오디오 데이터의 어휘개수가 학습에 유용한 수량과 수량의 균일성 및 비율이 고려되었는지 확인
검사 대상	산출물	• 원천데이터/라벨링데이터
	데이터 종류	• 텍스트/오디오
	내용	• 값
검사방법		• 인공지능 학습용 데이터 구축 시 정의한 어휘개수를 기준으로 어노테이션 데이터에 포함된 어휘개수 정보를 추출하여 통계값 산출 ※ 학습에 유용한 수량, 비율 등의 최소 기준에 대한 전문가 자문 필요

〈표 III-59〉 어휘개수 검사기준 및 방법

3.3.2.2 계산식

구분	계산 수식
어휘개수 분포 계산	※ 계산 방법에 대한 전문가 자문필요

〈표 III-60〉 어휘개수 분포 계산식

4 데이터 정확성 품질관리 지표

4.1 구문 정확성

- ‘구문 정확성’에 대한 세부지표는 ‘데이터 구조’, ‘입력값 범위’, ‘데이터 형식’으로 구성

품질지표	검사항목
구문 정확성	데이터 구조
	입력값 범위
	데이터 형식

〈표 III-61〉 구문 정확성 검사항목 구성

4.1.1 데이터 구조

4.1.1.1 검사기준 및 방법

분류	검사기준 및 방법
검사상세	• 어노테이션 데이터 구조를 준수하는지 확인
검사 대상	산출물
	데이터 종류
	내용
검사방법	<ul style="list-style-type: none"> 인공지능 학습용 데이터 구축 시 정의한 데이터 구조를 미준수하는 데이터가 있는지 확인 ※ 데이터 구조에 대한 표준화가 필요

〈표 III-62〉 데이터 구조 검사기준 및 방법

4.1.1.2 계산식

구분	계산 수식
오류율 계산	$\text{오류율}(\%) = \frac{\text{오류 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-63〉 데이터 구조 오류율 계산식

4.1.2 입력값 범위

4.1.2.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 어노테이션 데이터를 구성하는 속성의 값이 입력 유효 값 내에 존재하는지 확인
검사 대상	산출물	• 라벨링데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
	내용	• 값
검사방법		• 인공지능 학습용 데이터 구축 시 정의한 데이터를 구성하는 속성의 값이 입력 유효값을 벗어나는 속성을 포함하고 있는 데이터가 존재하는지 확인 ※ 메타데이터에 대한 표준화가 필요

〈표 III-64〉 입력값 범위 검사기준 및 방법

4.1.2.2 계산식

구분	계산 수식
오류율 계산	$\text{오류율}(\%) = \frac{\text{오류 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-65〉 입력값 범위 오류율 계산식

4.1.3 데이터 형식

4.1.3.1 검사기준 및 방법

분류		검사기준 및 방법
검사상세		• 어노테이션 데이터를 구성하는 데이터가 형식을 준수하는지 확인
검사 대상	산출물	• 라벨링데이터
	데이터 종류	• 텍스트/이미지/비디오/오디오/3D
	내용	• 값
검사방법		• 인공지능 학습용 데이터 구축 시 정의한 데이터 형식을 미준수하는 데이터가 있는지 확인 ※ 메타데이터에 대한 표준화가 필요 ※ 텍스트, 오디오는 ETRI전사규칙을 준수

〈표 III-66〉 데이터 형식 검사기준 및 방법

4.1.3.2 계산식

구분	계산 수식
오류율 계산	$\text{오류율}(\%) = \frac{\text{오류 건수(데이터 수)}}{\text{검사 총 건수(데이터 수)}} \times 100(\%)$

〈표 III-67〉 데이터 형식 오류율 계산식

4.2 의미 정확성

- ‘의미 정확성’에 대한 측정지표는 데이터 종류와 어노테이션 방식에 따라서 ‘정확도’, ‘정밀도’, ‘재현율’, ‘IoU’, ‘ROUGE-L’, ‘BLEU’, ‘Fβ-점수’, ‘EM’, ‘OKS’, ‘MOTA’ 등으로 구성

※ 의미 정확성 설명부분에서는 ‘정확도’, ‘정밀도’, ‘재현율’에 대한 검사기준 및 방법만을 기술하였으며, 그 외의 지표는 『5.2 유효성』에 기술된 지표를 참조

품질지표	측정지표
의미 정확성	정확도
	정밀도
	재현율
	IoU
	ROUGE-L
	BLEU
	Fβ-점수
	EM
	OKS
	MOTA
	...

〈표 III-68〉 의미 정확성 측정지표 구성

4.2.1 정확도

● 검사기준 및 방법

분류		검사기준 및 방법	
검사상세		• 어노테이션과 참값 간의 중첩율을 확인	
검사 대상	산출물	• 라벨링데이터	
	데이터 종류	데이터 종류	어노테이션
		텍스트	내용요약
			번역
			질의응답
		이미지/비디오	라벨링
			바운딩박스
			키포인트
		오디오	세그멘테이션
		3D	전사
	세그멘테이션		
	키포인트		
	큐보이드		
내용	• 값		
검사방법	• 데이터 종류의 분류를 참조하여 데이터 종류별에 해당하는 어노테이션과 참값 간의 중첩율을 계산하여 의미 정확성의 정확도를 확인		

〈표 III-69〉 정확도 검사기준 및 방법

● 계산식은 「5.2 유효성」 지표별 계산식(방식)을 참조

4.2.2 정밀도

● 검사기준 및 방법

분류		검사기준 및 방법		
검사상세		• 생성된 어노테이션 중 참값 어노테이션 비율을 확인		
검사 대상	산출물	• 라벨링데이터		
	데이터 종류	데이터 종류	어노테이션	
		텍스트	말뭉치 태깅	
			라벨링	
		이미지/비디오	바운딩박스	
			키포인트	
			세그멘테이션	
	3D	큐보이드		
내용	• 값			
검사방법		• 데이터 종류의 분류를 참조하여 데이터 종류별로 생성된 어노테이션 중 참값 어노테이션 비율을 계산하여 의미 정확성의 정밀도를 확인		

〈표 III-70〉 정밀도 검사기준 및 방법

- 계산식은 『5.2 유효성』 지표별 계산식(방식)을 참조

4.2.3 재현율

- 검사기준 및 방법

분류		검사기준 및 방법	
검사상세		• 모든 가능한 어노테이션 중 참값 어노테이션 비율을 확인	
검사 대상	산출물	• 라벨링데이터	
	데이터 종류	데이터 종류	어노테이션
		텍스트	말뭉치 태깅
		이미지/비디오	라벨링
			바운딩박스
			키포인트
			세그멘테이션
	3D	큐보이드	
내용	• 값		
검사방법		• 데이터 종류의 분류를 참조하여 데이터 종류별로 모든 가능한 어노테이션 중 참값 어노테이션 비율을 계산하여 의미 정확성의 재현율을 확인	

〈표 Ⅲ-71〉 재현율 검사기준 및 방법

- 계산식은 『5.2 유효성』 지표별 계산식(방식)을 참조

4.2.4 ROUGE-L, BLEU, F β -점수, EM, OKS, MOTA 등

- 검사기준 및 방법, 계산식은 『5.2 유효성』에 기술된 지표를 참조

5 학습모델 품질관리 지표

5.1 알고리즘 적정성

- ‘알고리즘 적정성’에 대한 검사항목은 데이터 유형 별로 구성되며, 각각의 데이터 유형별 관련 알고리즘을 임무(Task)별로 구분하여 ‘알고리즘 적정성’을 평가
- 각각의 Task 별 측정지표는 ‘2. 유효성’에서 설명하며, 본 장에서는 각각의 Task에 대한 개념을 제시하고, 제시되는 Task는 업무의 성격에 따라 ‘언어지능’ 및 ‘시각지능’으로 최상위층이 구분되며, 각각은 [이해/생성], [인지/생성]으로 좀 더 세분화될 수 있음
- 다만, 본 『제1권. 품질관리 가이드라인 v3.0』에서 제시하는 Task는, 한국지능정보사회진흥원에서 공모하는 ‘인공지능 학습용 데이터 구축사업’의 구축 사례들을 기반으로 널리 활용되는 Task 분야를 정리한 것으로 데이터 유형별 Task의 종류는 제시하는 것 외에도 매우 다양하므로, 데이터 구축 목적에 따라 별도의 Task를 정의해야 함

5.1.1 언어 이해

알고리즘 Task 분야		데이터 유형				
한글명	영문명	텍스트	이미지	비디오	오디오	3D
광학 문자 인식	<i>Optical Character Recognition</i>	●	●			
순차적 레이블링	<i>Sequence Labeling</i>	●				
텍스트 분류	<i>Text Classification</i>	●			○	
정보 추출	<i>Information Extraction</i>	●			○	
질의 응답	<i>Question Answering</i>	●			○	
음성 인식	<i>Speech Recognition</i>	●			●	
오디오 분류	<i>Audio Classification</i>	●	●		●	

※ 검은색 표기(●)는 학습모델에 직접적으로 적용되는 데이터유형을 표시함

※ 하얀색 표기(○)는 원시데이터에서 정제 작업을 거친 후 학습모델에 적용되는 데이터유형을 표시함

〈표 III-72〉 언어 이해 Task별 데이터 유형 구성

5.1.1.1 광학 문자 인식 (OCR, Optical Character Recognition)

구분	내용
개념	• 사람이 쓰거나 기계로 인쇄한 문자의 영상을 이미지 스캐너로 획득하여 기계가 읽을 수 있는 문자로 변환하는 작업
세부 내용	• 기존의 정형화된 숫자나 텍스트 인식에서 벗어나, 인공지능 기술이 접목되어 필기체, 이미지, 비정형 데이터까지 인식 범위를 넓히고 있음

〈표 III-73〉 광학 문자 인식의 개념 및 세부 내용

5.1.1.2 순차적 레이블링 (Sequence Labeling)

구분	내용
개념	• 입력 시퀀스에 대하여 레이블 시퀀스를 각각 부여하는 작업으로, 태깅 작업이 대표적인 순차적 레이블링의 예시임
세부 내용	• 태깅을 해야하는 단어 데이터를 X, 레이블에 해당되는 태깅 정보 데이터는 y라고 할 때, X와 y데이터의 쌍(pair)은 병렬 구조를 가지며, 각 데이터는 정수 인코딩 과정을 거친 후, 모든 데이터의 길이를 동일하게 맞춰주기 위한 패딩(Padding) 작업을 거쳐 학습용 데이터로 구축됨

〈표 III-74〉 순차적 레이블링의 개념 및 세부 내용

5.1.1.3 텍스트 분류 (Text Classification)

구분	내용
개념	• 문장이나 문서를 적절한 범주로 지정하는 작업으로, 범주는 선택한 데이터 세트에 따라 다르며 주제 범위가 다양할 수 있음
세부 내용	<ul style="list-style-type: none"> • 텍스트 분류 문제에는 감정 분류, 뉴스 분류, 인용 의도 분류 등이 있으며, 텍스트 분류 기능을 평가하기 위한 벤치마크 데이터 세트에는 GLUE, AGNews 등이 있음 • 최근 몇 년 동안 XLNet 및 RoBERTa와 같은 딥러닝 기술은 텍스트 분류 문제에서 가장 큰 성능 향상을 달성함

〈표 III-75〉 텍스트 분류의 개념 및 세부 내용

5.1.1.4 정보 추출 (Information Extraction)

구분	내용
개념	• 자연어에서 유용한 정보를 자동으로 수집하기 위해 비정형화된 텍스트 데이터 내에서 정형화된 데이터를 추출하는 작업
세부 내용	<ul style="list-style-type: none"> • 고유 명사 등을 판별하는 개체명 인식(Named Entity Recognition), 개체명 간의 관계를 판별하는 관계 추출(Relation Extraction), 중의성 해소(word sense disambiguation), 대표 단어/문장 추출 등이 가장 대표적인 정보 추출 문제이며, MRC(Machine Reading & Comprehension)의 기반이 되는 작업이라 할 수 있음

〈표 III-76〉 정보 추출의 개념 및 세부 내용

5.1.1.5 질의 응답 (Question Answering)

구분	내용
개념	<ul style="list-style-type: none"> 질문(일반적으로 독해 질문)에 답변하는 작업으로서, 제공된 컨텍스트를 기반으로 답변할 수 없는 질문이 제시되면 기권함
세부 내용	<ul style="list-style-type: none"> 질의 응답은 커뮤니티 질문 답변 및 지식 기반 질문 답변과 같은 도메인별 작업으로 나눌 수 있음 평가 질문 응답 시스템을 위한 인기 있는 벤치마크 데이터 세트에는 SQuAD, HotPotQA, bAbI, TriviaQA, WikiQA 등이 있으며, 질문 답변 모델은 일반적으로 EM 및 F1과 같은 성능지표로 평가함 최근 질의 응답 분야 최고 성능 모델은 T5 및 XLNet 등이 있음

〈표 III-77〉 질의 응답의 개념 및 세부 내용

5.1.1.6 음성 인식 (Speech Recognition)

구분	내용
개념	<ul style="list-style-type: none"> 사람이 말하는 음성 언어를 인공지능이 해석해 그 내용을 문자 데이터로 전환하는 작업
세부 내용	<ul style="list-style-type: none"> 일반적으로 STT(Speech to Text)와 동일한 Task로 인식되며, Speaker Verification이나 Speaker Recognition과 같이 개별 사용자의 음성을 식별하는 데에 중점을 둔 Task도 존재함

〈표 III-78〉 음성 인식의 개념 및 세부 내용

5.1.1.7 오디오 분류 (Audio Classification)

구분	내용
개념	<ul style="list-style-type: none"> 오디오 정보를 특정 시간 단위로 분할(오디오 클립)하여, 신호의 특징점을 분석하고 특정 클래스로 분류하는 작업
세부 내용	<ul style="list-style-type: none"> 소리는 음성과 음향으로 구분이 되며, 일반적으로 오디오 분류(Audio Classification)이라고 정의할 때는 음향에 대한 분류를 의미함 오디오 분류 작업은 주석이 달린 오디오 데이터로 시작하며, 기계는 어떻게 듣고 무엇을 들어야 하는지 배우기 위해 이 데이터가 필요함 구글 오디오셋이 오디오 분류의 대표적인 데이터셋임

〈표 III-79〉 오디오 분류의 개념 및 세부 내용

5.1.2 언어 생성

알고리즘 Task 분야		데이터 유형				
한글명	영문명	텍스트	이미지	비디오	오디오	3D
텍스트 요약	Text Summary	●			○	
텍스트 생성	Text Generation	●				
기계 번역	Machine Translation	●			○	
음성 합성	Speech Synthesis	●			●	

※ 검은색 표기(●)는 학습모델에 직접적으로 적용되는 데이터유형을 표시함

※ 하얀색 표기(○)는 원시데이터에서 정제 작업을 거친 후 학습모델에 적용되는 데이터유형을 표시함

〈표 III-80〉 언어 생성 Task별 데이터 유형 구성

5.1.2.1 텍스트 요약 (Text Summary)

구분	내용
개념	• 상대적으로 큰 원문을 핵심 내용만 간추려서 상대적으로 작은 요약문으로 변환하는 작업
세부 내용	• 원문에서 중요한 핵심 문장 또는 단어구를 뽑아 이들로 구성하는 '추출적 요약(Extractive Summarization)'과 원문에 없던 문장이라도 핵심 문맥을 반영하여 새로운 문장을 생성해서 요약하는 '추상적 요약(Abstractive Summarization)'으로 구분됨

〈표 III-81〉 텍스트 요약의 개념 및 세부 내용

5.1.2.2 텍스트 생성 (Text Generation)

구분	내용
개념	• 기계학습을 이용하여 문장을 생성하는 작업
세부 내용	• LSTM(Long Short Term Memory)/RNN(Recurrent Neural Network) 나 Softmax 알고리즘을 적용하여, 인간이 문장을 만드는 과정을 최대한 구현하는 방향으로 Language Modeling이 발전하고 있음

〈표 III-82〉 텍스트 생성의 개념 및 세부 내용

5.1.2.3 기계 번역 (Machine Translation)

구분	내용
개념	• 인공지능을 이용해 인간이 사용하는 자연어를 다른 자연어로 번역하는 작업
세부 내용	• 자동 번역으로 부르기도 하며, 구글 번역, 네이버 파파고, SYSTRAN, 바벨피시, 빙 번역기 등 다양한 기계 번역기가 존재함 • 최근에는 기계 번역이 스스로 학습해 자신의 정확도를 향상하는 역 번역(Back Translation)과 디노이징(Denoising) 같은 기술을 적용하여 발전하고 있음

〈표 III-83〉 기계 번역의 개념 및 세부 내용

5.1.2.4 음성 합성 (Speech Synthesis)

구분	내용
개념	• 텍스트나 입술 움직임과 같이 다른 매체정보로부터 음성을 생성하는 작업
세부 내용	<ul style="list-style-type: none"> • 음성 인식에 대응하는 음성 합성은 주로 텍스트 정보를 오디오 정보로 변환하는 데 사용되며 음성 지원 서비스 및 모바일 응용 프로그램과 같은 응용 프로그램에서 사용됨 • 이 외에도 시각 장애인이 텍스트 콘텐츠를 읽을 수 있도록 돕는 보조 기술에도 사용될 수 있음

〈표 III-84〉 음성 합성의 개념 및 세부 내용

5.1.3 시각 인지

알고리즘 Task 분야		데이터 유형				
한글명	영문명	텍스트	이미지	비디오	오디오	3D
이미지 분류	Image Classification		●	○		
3D 포인트 클라우드 분류	3D Point Cloud Detection					●
객체 인식	Object Detection		●	○		
3D 객체 인식	3D Object Detection		○	○		●
얼굴 인식	Face Recognition		●			
키폰트 검출	Keypoint Detection		●	○		
3D 키폰트 검출	3D Keypoint Detection					○
행동 인식	Action Recognition		●	○		
자세 추정	Pose Estimation		●	○		
3D 자세 추정	3D Pose Estimation					○
비디오 분류	Video Classification			●		
3D 객체 분류	3D Object Classification					○
비디오 인식	Video Recognition			●		

※ 검은색 표기(●)는 학습모델에 직접적으로 적용되는 데이터유형을 표시함

※ 하얀색 표기(○)는 원시데이터에서 정제 작업을 거친 후 학습모델에 적용되는 데이터유형을 표시함

〈표 III-85〉 시각 인지 Task별 데이터 유형 구성

5.1.3.1 이미지 분류 (Image Classification)

구분	내용
개념	• 특정 이미지를 알고리즘에 입력했을 때, 그 이미지가 어떤 클래스에 속하는지 알려주는 작업
세부 내용	• 단순히 객체의 큰 범위만 분류하는 것이 아닌, 해당 속성이나 상세한 클래스까지 분류해내는 경우 이미지 인식 또는 객체 인식이라고 할 수 있음

〈표 III-86〉 이미지 분류의 개념 및 세부 내용

5.1.3.2 3D 포인트 클라우드 분류 (3D Point Cloud Classification)

구분	내용
개념	• 특정 포인트 클라우드를 알고리즘에 입력했을 때, 그 포인트 클라우드가 어떤 클래스에 속하는지 알려주는 작업
세부 내용	• 단순히 객체의 큰 범위만 분류하는 것이 아닌, 해당 속성이나 상세한 클래스까지 분류해내는 경우 3D 객체 인식이라고 할 수 있음

〈표 III-87〉 3D 포인트 클라우드 분류의 개념 및 세부 내용

5.1.3.3 객체 인식 (Object Detection)

구분	내용
개념	• 이미지 내에서 특정 클래스의 객체 인스턴스를 인식하는 작업
세부 내용	<ul style="list-style-type: none"> • 객체 인식을 위한 방법은 1단계 방법과 2단계 방법의 두 가지 주요 유형으로 분류할 수 있음 • 1단계 방법은 추론 속도를 우선시하며, 대표적인 모델로는 YOLO, SSD, RetinaNet이 있음 • 2단계 방법은 탐지 정확도를 우선시하며, 대표적인 모델로는 Faster R-CNN, Mask R-CNN, Cascade R-CNN이 있음 • 객체 인식 작업과 관련된 가장 인기 있는 데이터셋은 MSCOCO이며, 모델은 일반적으로 평균 정밀도 평가모델에 따라 평가됨

〈표 III-88〉 객체 인식의 개념 및 세부 내용

5.1.3.4 3D 객체 인식 (3D Object Detection)

구분	내용
개념	• 3D 영상/이미지 내에서 특정 클래스의 객체 인스턴스를 감지하는 작업
세부 내용	<ul style="list-style-type: none"> • 3D 객체 인식은 일반적인 객체인식과 유사하지만, 입력값으로 RGB 이미지와 함께 depth 값을 받아서 학습하고, 추론시에는 RGB 이미지만을 입력으로 받아 depth 값을 추정하는 것이 중요한 차이라고 할 수 있음 • 객체 인식 분야를 3차원 객체 인식으로 범위를 확장하면, 물체의 크기, 위치, 방향 등을 알 수 있으므로 자율 주행 및 이미지 검색, 증강 현실에서 다양한 분야에 응용 가능함

〈표 III-89〉 3D 객체 인식의 개념 및 세부 내용

5.1.3.5 얼굴 인식 (Face Recognition)

구분	내용
개념	• 얼굴을 사용하여 개인의 신원을 식별하거나 확인하는 작업
세부 내용	<ul style="list-style-type: none"> • 얼굴 인식 시스템은 사진, 비디오 또는 실시간으로 사람을 식별하는 데 사용할 수 있음 • 안면 인식은 생체 인식 보안의 범주로서, 다른 형태의 생체 인식 소프트웨어에는 음성 인식, 지문 인식, 눈 망막 또는 홍채 인식 등이 있음 • 얼굴 인식 Task는 Face Verification과 Face Identification과 같이 세부적인 Task로 구체화 됨

〈표 III-90〉 얼굴 인식의 개념 및 세부 내용

5.1.3.6 키포인트 검출 (Keypoint Detection)

구분	내용
개념	• 동작 인식을 위해 필요한 주요 키포인트를 검출하는 작업
세부 내용	<ul style="list-style-type: none"> • 키포인트는 관심 포인트라고도 하며, 이미지에서 흥미롭거나 눈에 띄는 것을 정의하는 공간적 위치 또는 이미지의 포인트를 지칭함 • 특징으로는, 이미지 회전, 축소, 변환, 왜곡 등에 대해 변하지 않음 • 얼굴을 대상으로 할 때는 Facial Landmark Detection, 신체를 대상으로 할 때는 Pose Estimation 이라는 Task로 좀 더 구체화 됨

〈표 III-91〉 키포인트 인식의 개념 및 세부 내용

5.1.3.7 3D 키포인트 검출 (3D Keypoint Detection)

구분	내용
개념	• 3D 데이터에서 동작 인식을 위해 필요한 주요 키포인트를 검출하는 작업
세부 내용	<ul style="list-style-type: none"> • 키포인트는 관심 포인트라고도 하며, 3D 데이터에서 흥미롭거나 눈에 띄는 것을 정의하는 공간적 위치 포인트를 지칭함 • 특징으로는, 3D 데이터 회전, 축소, 변환, 왜곡 등에 대해 변하지 않음

〈표 III-92〉 3D 키포인트 인식의 개념 및 세부 내용

5.1.3.8 행동 인식 (Action Recognition)

구분	내용
개념	• 비디오 데이터 스트림에서 객체 행동의 감지 및 인식을 하는 작업
세부 내용	<ul style="list-style-type: none"> • 실시간 CCTV 비디오 데이터 스트림에서 객체 행동의 감지 및 인식은, 잠재적으로 위험한 시나리오의 이상 감지를 위해 활용될 수 있음 • 최첨단 딥러닝 기술과 '경로 서명'을 결합하여 객체 행동을 해석하기 위한 일반화된 프레임워크를 개발하며, 문제를 효율적이고 효과적으로 해결하는 것을 목표로 함

〈표 III-93〉 행동 인식의 개념 및 세부 내용

5.1.3.9 자세 추정 (Pose Estimation)

구분	내용
개념	• 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업
세부 내용	<ul style="list-style-type: none"> • 기술적으로 키포인트 검출(Keypoint Detection)이나 행동 인식(Action Recognition) Task와 유사한 문제로 정의됨 • 게임, 의료, AR 및 스포츠를 포함한 광범위한 분야에서 응용 프로그램을 찾는 확장성으로 인해 많은 연구가 진행되고 있음

〈표 III-94〉 자세 추정의 개념 및 세부 내용

5.1.3.10 3D 자세 추정 (3D Pose Estimation)

구분	내용
개념	• 3D 데이터에서 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업
세부 내용	<ul style="list-style-type: none"> • 기술적으로 3D 키포인트 검출 (3D Keypoint Detection) Task와 유사한 문제로 정의됨 • 게임, 의료, AR 및 스포츠를 포함한 광범위한 분야에서 응용 프로그램을 찾는 확장성으로 인해 많은 연구가 진행되고 있음

〈표 III-95〉 3D 자세 추정의 개념 및 세부 내용

5.1.3.11 비디오 분류 (Video Classification)

구분	내용
개념	• 프레임이 주어진 비디오에서 특정 프레임 구간을 선택하고 클래스 레이블을 부여하는 작업
세부 내용	<ul style="list-style-type: none"> • 좋은 비디오 레벨 분류기는 정확한 프레임 레이블을 제공할 뿐만 아니라 비디오에 있는 다양한 프레임의 기능과 주석을 고려하여 전체 비디오를 설명할 수 있음 • 프레임과 비디오를 설명하는 데 필요한 레이블의 세분성은 작업에 따라 다르며, 일반적인 작업에는 비디오에 하나 이상의 전역 레이블을 할당하고 비디오 내부의 각 프레임에 하나 이상의 레이블을 할당하는 것이 포함됨

〈표 III-96〉 비디오 분류의 개념 및 세부 내용

5.1.3.12 3D 객체 분류 (3D Object Classification)

구분	내용
개념	• 3D 이미지/영상 데이터를 알고리즘에 입력했을 때, 그 3D 이미지/영상이 어떤 클래스에 속하는지 알려주는 작업
세부 내용	• 단순히 객체의 큰 범위만 분류하는 것이 아닌, 해당 속성이나 상세한 클래스까지 분류해내는 경우 3D 객체 인식이라고 할 수 있음

〈표 III-97〉 3D 객체 분류의 개념 및 세부 내용

5.1.3.13 비디오 인식 (Video Recognition)

구분	내용
개념	• 비디오를 활용하여 실시간 객체 식별 및 비디오 프레임을 분류하는 작업
세부 내용	<ul style="list-style-type: none"> • 컴퓨터는 수천 개의 비디오 스트림을 보고 프레임 단위로 수신하는 정보를 이해하는 과정을 거치며, Tracking 하는 업무로 정의될 수 있음 • 카메라를 사용하여 시간이 지남에 따라 움직이는 물체를 찾아 연속적인 비디오 프레임에서 대상 물체를 연결하는 것으로 구성됨

〈표 III-98〉 비디오 인식의 개념 및 세부 내용

5.1.4 시각 생성

알고리즘 Task 분야		데이터 유형				
한글명	영문명	텍스트	이미지	비디오	오디오	3D
이미지 생성	Image Generation		●			
이미지 합성	Image Synthesis		●			
3D 객체 복원	3D Object Reconstruction		○	○		○
2D 이미지를 통한 사점 합성	2D View Synthetic		○			○

※ 검은색 표기(●)는 학습모델에 직접적으로 적용되는 데이터유형을 표시함

※ 하얀색 표기(○)는 원시데이터에서 정제 작업을 거친 후 학습모델에 적용되는 데이터유형을 표시함

〈표 III-99〉 시각 생성 Task별 데이터 유형 구성

5.1.4.1 이미지 생성 (Image Generation)

구분	내용
개념	• 인공지능이 다양한 노이즈(Noise) 입력을 받아, 원하는 카테고리 내 새로운 이미지를 생성하는 작업
세부 내용	• GAN 기술을 이용하여 이미지 정보들의 변환과 재구축으로 진짜처럼 보이는 가짜 이미지 형성이 가능함

〈표 III-100〉 이미지 생성의 개념 및 세부 내용

5.1.4.2 이미지 합성 (Image Synthesis)

구분	내용
개념	• 인공지능이 이미지나 비디오를 입력 받아, 다른 형태나 정보를 지닌 이미지 또는 비디오로 변환하는 작업
세부 내용	• GAN 기술을 이용하여 적은 양의 정보로 원본 이미지를 예측하거나, 사진을 특정한 방식의 이미지로 전환하는 것이 가능함

〈표 III-101〉 이미지 합성의 개념 및 세부 내용

5.1.4.3 3D 객체 복원(3D Object Reconstruction)

구분	내용
개념	• 인공지능이 영상/이미지/포인트 클라우드로부터 객체(물체, 사람 등)의 3차원 구조를 복원(reconstruction)하는 작업
세부 내용	• 3D GAN 기술을 이용하여 영상/이미지/포인트 클라우드 정보들의 변환과 재구축으로, 실제 객체와 기하학적인 정확도뿐만 아니라 현실감이나 사실성 측면에서 매우 유사한 객체 형성이 가능함

〈표 III-102〉 3D 객체 복원 생성의 개념 및 세부 내용

5.1.4.4 2D 이미지를 통한 시점 합성(View Synthetic)

구분	내용
개념	• 인공지능을 활용하여 이미지를 인공적으로 3D로 전환하는 작업
세부 내용	• 서로 다른 시점을 가진 여러 개의 카메라들을 이용하여 객체를 촬영한 뒤, 촬영된 영상들과 각 카메라들의 촬영 정보들을 이용하여 해당 객체에 대한 3차원 정보를 추정

〈표 III-103〉 2D 이미지 합성의 개념 및 세부 내용

5.2 유효성

- ‘유효성’에 대한 검사항목은 알고리즘 Task 분야 별로 구성되며, 각각의 Task 별로 유효성 측정지표를 사용할 수 있음
- 본 장에서는 각각의 Task에 대한 상세 설명을 데이터 유형별로 제시하였고, 제시되는 유효성 측정지표는 데이터의 유형에 따라 ‘이미지 데이터’, ‘비디오 데이터’, ‘오디오 데이터’, ‘텍스트 데이터’, ‘3D 데이터’로 구분
- 다만, 본 『제1권 품질관리 가이드라인 v3.0』에서 제시하는 유효성 측정지표는 한국지능정보사회진흥원에서 공모하는 ‘인공지능 학습용 데이터 구축사업’의 구축 사례들을 기반으로 자주 활용되는 측정지표를 정리한 것으로, 제시하는 것 이외에도 매우 다양하므로 데이터 구축 목적에 따라 적절한 측정지표 및 목표를 정의해야 함
- 또한, 유효성 목표 설정 시 다음과 같은 원칙을 준수해야 함

〈유효성 목표 설정 원칙〉

- 최신의 논문, 보고서, 공식 블로그 등에서 공식 보고된 값을 참고하여 유효성 목표치 기준 수립
 - 해당 구축사업과 가장 유사한 주제의 SoTA(State of The Art) 자료(논문, 보고서 등)에서 공식 보고되는 값을 기준으로 유효성 목표치를 설정하는 것이 원칙
 - 해당 구축사업의 이해관계자 간에 목표치에 대해 논의할 때, 반드시 상기 자료를 근거로 하고 출처를 명확히 제시
 - 다만, 선행 연구 분야에 해당하는 임무이기 때문에 SoTA에 해당 자료를 참고하기 힘든 경우 이해관계자 및 해당 임무·도메인 관련 전문가 의견을 반영하여 설정

〈표 III-104〉 유효성 목표 설정 원칙

- 〈표 III-102〉, 〈표 III-109〉, 〈표 III-125〉, 〈표 III-129〉에 기재된 ‘알고리즘 Task 분야’ 별 ‘라벨 유형’이나 ‘유효성 측정지표’의 경우, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음
- 이것은 한국지능정보사회진흥원에서 공모하는 ‘인공지능 학습용 데이터 구축사업’ 내 모든 과제에 대해서 동일한 기준으로 유효성 측정지표 및 목표의 타당성을 검증하지 못하는 것이 첫 번째 이유이며, 각각의 데이터 분야별 요구하는 유효성 기준이 다를 수 있다는 것이 두 번째 이유로 본 『제1권 품질관리 가이드라인 v3.0』에 기술된 내용을 참고로 하되, 과업의 목표에 따라 적합한 측정지표를 선정하는 것이 무엇보다 중요함

5.2.1 텍스트 데이터

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
텍스트 분류 (Text Classification)	Machine Translation	BLEU	73.25
		F β -점수	91.3
	Class Labeling	정확도	98.68
		AUROC	99
		F β -점수	90.8
		MAE	59
		MSE	71
		PCC	82.28
		정밀도	86.1
		QWK	81.16

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
	Text Summary	$F\beta$ -점수	89.85
		TOP-K 정확도	95
		ROUGE-L	37.55
언어 이해 (Language Understanding)	Machine Translation	BLEU	77.22
	Sequence Labeling	$F\beta$ -점수	88.53
		정밀도	86.1
질의 응답 (Question Answering)	Class Labeling	정확도	77.78
		$F\beta$ -점수	84.64
음성 인식 (Speech Recognition)	Class Labeling	정확도	76.05

※ 해당 내용은 한국지능정보사회진흥원에서 공모하는 ‘인공지능 학습용 데이터 구축사업’에 대한 품질검증결과를 기반으로 작성되었으며, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음

〈표 III-105〉 텍스트 데이터 유효성 품질지표 구성

5.2.1.1 $F\beta$ -점수

구분	내용															
개념	<ul style="list-style-type: none"> 정밀도와 재현율의 가중치 조화 평균으로 정밀도 또는 재현율에 가중치를 주어 계산되는 평가지표 															
세부 내용	<ul style="list-style-type: none"> 데이터 Label이 불균형 구조일 때, 모델의 성능을 정확하게 평가할 수 있으며, 성능을 하나의 숫자로 표현할 수 있음 매개변수인 베타(beta) 점수로 재현율의 가중치를 결정하며, 정밀도에 가중치를 부여하는 경우 $\beta < 1$, 재현율에 가중치를 부여하는 경우 $\beta > 1$를 적용 															
계산식(방식)	<table border="1"> <thead> <tr> <th colspan="2" rowspan="2"></th><th colspan="2">Predicted Condition</th></tr> <tr> <th>Positive(PP)</th><th>Negative (PN)</th></tr> </thead> <tbody> <tr> <td rowspan="2">Actual Condition</td><td>Positive (P)</td><td>True Positive (TP)</td><td>False Negative (FN)</td></tr> <tr> <td>Negative (N)</td><td>False Positive (FP)</td><td>True Negative (TN)</td></tr> </tbody> </table> <p> ※ TP: 실제 True인 정답을 True라고 예측 ※ FP: 실제 False인 정답을 True라고 예측 ※ FN: 실제 True인 정답을 False라고 예측 ※ TN: 실제 False인 정답을 False라고 예측 </p>					Predicted Condition		Positive(PP)	Negative (PN)	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)	Negative (N)	False Positive (FP)	True Negative (TN)
		Predicted Condition														
		Positive(PP)	Negative (PN)													
Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)													
	Negative (N)	False Positive (FP)	True Negative (TN)													

구분	내용
	$F_{\beta}\text{-점수} = (1 + \beta^2) \times \frac{\text{precision} \times \text{recall}}{(\beta^2 \times \text{precision}) + \text{recall}}$ $\text{Precision} = \frac{TP}{TP + FP}$ $\text{Recall} = \frac{TP}{TP + FN}$
관련 출처	<ul style="list-style-type: none"> “Metric: f1”. Hugging Face. 2021년 12월 23일 접속. https://huggingface.co/metrics/f1 “fbeta_score”. 2023년 1월 3일 접속. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.fbeta_score.html F-beta score. 2023년 1월 3일 접속. https://hasty.ai/docs/mp-wiki/metrics/f-beta-score

〈표 III-106〉 F 베타-점수의 개념 및 세부 내용

5.2.1.2 Accuracy (정확도)

구분	내용													
개념	<ul style="list-style-type: none">알고리즘이 데이터 포인트를 올바르게 분류하는 빈도를 측정하는 평가지표													
세부 내용	<ul style="list-style-type: none">기계 학습 분류 알고리즘의 정확도는 알고리즘이 데이터 포인트를 올바르게 분류하는 빈도를 측정하는 한 가지 방법으로, 정확도는 모든 데이터 포인트 중에서 올바르게 예측된 데이터 포인트의 수라고 할 수 있음정확도는 다양한 비율의 참/거짓 양성/음성을 사용하는 다른 품질지표인 정밀도 및 재현율과 함께 사용되기도 함 <div></div>													
계산식(방식)	<table><tr><th colspan="2" rowspan="2"></th><th colspan="2">Predicted Condition</th></tr><tr><th>Positive(PP)</th><th>Negative (PN)</th></tr><tr><th rowspan="2">Actual Condition</th><th>Positive (P)</th><td>True Positive (TP)</td><td>False Negative (FN)</td></tr><tr><th>Negative (N)</th><td>False Positive (FP)</td><td>True Negative (TN)</td></tr></table>			Predicted Condition		Positive(PP)	Negative (PN)	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)	Negative (N)	False Positive (FP)	True Negative (TN)
				Predicted Condition										
		Positive(PP)	Negative (PN)											
Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)											
	Negative (N)	False Positive (FP)	True Negative (TN)											

구분	내용
	$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$
관련 출처	<ul style="list-style-type: none"> “Accuracy(error rate)”. DeepAI, 2021년 12월 23일 접속, https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate

〈표 III-107〉 Accuracy의 개념 및 세부 내용

5.2.1.3 BLEU (Bilingual Evaluation Understudy Score)

구분	내용
개념	<ul style="list-style-type: none"> 자연어를 기계 번역하여 다른 종류의 자연어로 생성된 문장을 평가하기 위한 평가지표
세부 내용	<ul style="list-style-type: none"> BLEU는 생성된 문장을 참조 문장으로 평가하기 위한 평가지표로서, 완벽한 일치는 1.0의 점수가 나오는 반면 완벽한 불일치는 0.0의 점수가 나옴 측정 기준은 n-gram(n개의 연속적인 단어 나열)에 기반하며, n은 1~4 사이의 범위를 가진다. 여기서 n-gram precision(정밀도)란, n-gram 중 일치하는 개수를 전체 n-gram의 수로 나눈 값을 의미한다.
계산식(방식)	$\text{BLEU} = \min \left(1, \frac{\text{output-length}}{\text{reference-length}} \right) \left(\prod_{i=1}^4 \text{precision}_i \right)^{1/4}$
관련 출처	<ul style="list-style-type: none"> “Metric: BLEU”. Hugging Face, 2021년 12월 23일 접속, https://huggingface.co/metrics/bleu

〈표 III-108〉 BLEU의 개념 및 세부 내용

5.2.1.4 EM (Exact Matching)

구분	내용
개념	<ul style="list-style-type: none"> 실제 참값과 정확하게 일치하는 비율을 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> 음성인식 분야 중, 특히 Machine Translation에서 정답과 예측값의 정확도를 측정하기 위한 지표
관련 출처	<ul style="list-style-type: none"> “Exact Matching”. MatchIt, 2021년 12월 23일 접속, https://kosukeimai.github.io/MatchIt/reference/method_exact.html

〈표 III-109〉 EM의 개념 및 세부 내용

5.2.1.5 ROUGE-L (Recall-Oriented Understudy for Gisting Evaluation-Longest Common Subsequence)

구분	내용
개념	<ul style="list-style-type: none"> 가장 긴 일치하는 단어 시퀀스를 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> ROUGE-L (Recall-Oriented Understudy for Gisting Evaluation-LCS)를 사용하여 가장 긴 일치하는 단어 시퀀스를 측정함 LCS 개념을 적용한 평가지표 사용 시, 단어의 연속적 일치가 필요 없고 문장 수준의 어순을 반영한 순차적 일치를 측정하므로 유연하다는 장점이 있음
계산식(방식)	$\text{ROUGE-L}(\text{candidate}, \text{references}) = \max_k \{ \text{ROUGE-L}_{\text{single}}(\text{candidate}, \text{references}_k) \}$ $\text{ROUGE-L}_{\text{single}}(\text{candidate}, \text{reference}) = \frac{(1 + \beta^2) R_{\text{lcs}}(\text{candidate}, \text{reference}) P_{\text{lcs}}(\text{candidate}, \text{reference})}{R_{\text{lcs}}(\text{candidate}, \text{reference}) + \beta^2 P_{\text{lcs}}(\text{candidate}, \text{reference})}$ $R_{\text{lcs}}(\text{candidate}, \text{reference}) = \frac{\sum_{r_j \in \text{reference}} \text{LCS}_j(\text{candidate}, r_j) }{\text{numWords}(\text{reference})}$ $P_{\text{lcs}}(\text{candidate}, \text{reference}) = \frac{\sum_{r_j \in \text{reference}} \text{LCS}_j(\text{candidate}, r_j) }{\text{numWords}(\text{candidate})}$
관련 출처	<ul style="list-style-type: none"> “ROUGE”. ACL Anthology. 2021년 12월 23일 접속,. https://aclanthology.org/W04-1013/

〈표 III-110〉 ROUGE-L의 개념 및 세부 내용

5.2.1.6 TQA (Translation Quality Assessment)

구분	내용						
개념	• 원문의 의미를 포착하고, 적절한 단어가 있으며 오류가 없고 일관성 있게 번역되어 있는지 측정하는 평가지표						
세부 내용	• TQA(Translation Quality Assessment) 는 4가지 파라미터로 정의되며, 각각 Meaning(의미), Wording / Expression(단어 및 표현), Errors(오류), Consistency(일관성) 임						
계산식(방식) (예시)	Words Checked:		265				
	QA Result:		Borderline Pass		Quality Indicator: 92,5		
	Error points allowed:		1,325		Error points calculated: 1		
	Error category		Critical	Medium	Minor	Repeat	Instruction
	Accuracy		0	0	1	0	1
	Terminology - Adherence		0	0	0	0	0
	Terminology - Consistency		0	0	0	0	0
	Appeal		0	0	0	5	2
	Language - Tone & Voice		0	0	0	0	0
	Language - Grammar & Syntax		0	0	0	0	0
	Language - Punctuation & Spelling		0	0	0	0	0
	Compliance - Style Guide & Country Standards		0	0	0	0	0
	Compliance - Project Instructions&Formal		0	0	0	0	0
Total		0	0	1	5	3	

구분	내용
관련 출처	<ul style="list-style-type: none"> • “The limitations of a quality assessment”. vertaalt, 2021년 12월 23일 접속. https://www.vertaalt.nu/blog/the-limitations-of-a-quality-assessment/ • 인공지능 데이터 구축·활용 가이드라인-전문분야 학습용 한영 번역 말뭉치. AI-Hub. 2021년 12월 23일 접속. https://aihub.or.kr/sites/default/files/Sample_data/구축활용가이드북_2020-01/006.전문분야_한영번역말뭉치_데이터_구축_가이드라인.pdf

〈표 III-111〉 TQA의 개념 및 세부 내용

5.2.2 이미지 데이터

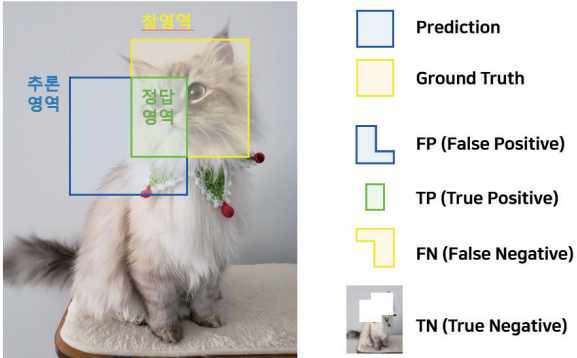
알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
객체 인식 (Object Detection)	Bounding Box	EAO	73.24
		FPS	20.93
		IoU	79.98
		mAP	96.87
		mAP@IoU	93.6
		mIoU	74.74
		Top-K 정확도	100
		재현율	80.48
		정확도	90.81
		Fβ-점수	96
	Keypoint	mAP	87.26
	Polygon	AP	38.78
		AR	72.24
		BLEU	81.92
		FID	10.44
		mAP	93.4
		mAP@IoU	89.6
		mIoU	74.74
		Top-K 정확도	99.74
		정확도	81.24
		Fβ-점수	90.77

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
	Polyline	Correlation	71
		mAP@IoU	71.2
키포인트 검출 (Keypoint Detection)	Keypoint	mAP	99.29
		정확도	94.48
얼굴 인식 (Face Recognition)	Bounding Box	Kinship Verification 정확도	70.66
		MAE	3.11
		정확도	67.2
	Keypoint	AP	98.76
		FID	7.94
		Rank-K 정확도	81.9
		정확도	99.31
	Polyline	Rank-K 정확도	81.9
	Tagging	Rank-K 정확도	81.9
이미지 분류 (Image Classification)	Bounding Box	AUROC	98.24
		Grid Cell 정확도	97.47
		mAP	99.1
		Top-K 정확도	95.3
		정확도	91.73
		F β -점수	86.04
	Polygon	mAP	97.4
		mIoU	85.93
		Overall Pixel 정확도	95.95
		F β -점수	99.6

※ 해당 내용은 한국지능정보사회진흥원에서 공모하는 '인공지능 학습용 데이터 구축사업'에 대한 품질검증결과를 기반으로 작성되었으며, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음

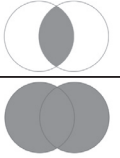
〈표 Ⅲ-112〉 이미지 데이터 유효성 품질지표 구성

5.2.2.1 DICE-Coefficient

구분	내용
개념	<ul style="list-style-type: none"> 두 개의 데이터 세트 간의 유사성을 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> DICE-Coefficient는 두 데이터 세트 간의 유사성을 측정하는 통계 도구이며, 2개의 중첩 영역을 두 이미지의 총 픽셀(pixel) 수로 나누어 계산함 인공지능으로 생성된 이미지 분할 알고리즘 검증에 특화되어있으며, 이미지 분할에서 가장 일반적으로 사용되는 품질지표임 참고로 DICE-coefficient는 F1-score의 값과 같다. 따라서 F1-score를 이미지에 적용한 개념이라고 볼 수 있다. $* \text{DICE-coefficient} = \text{F1 score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}$
계산식(방식)	 <p> $\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{정답영역}}{\text{정답영역} + \text{추론영역}}$ $\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{정답영역}}{\text{정답영역} + \text{참영역}}$ $\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} = \frac{\text{정답영역}}{\text{정답영역} + \text{추론영역} + \text{참영역}}$ $\text{DICE-coefficient} = \frac{2 * \text{Area of Intersection}}{\text{Area of Union} + \text{Area of Intersection}} = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}} = \frac{\text{정답영역} + \text{정답영역}}{\text{추론영역} + \text{참영역}}$ </p>
관련 출처	<ul style="list-style-type: none"> Daniel J Bell, "Dice similarity coefficient", Radiopaedia, 2021년 12월 23일 접속, https://radiopaedia.org/articles/dice-similarity-coefficient

〈표 III-113〉 DICE-Coefficient의 개념 및 세부 내용

5.2.2.2 mIoU (mean Intersection Over Union)

구분	내용
개념	<ul style="list-style-type: none"> 개체 범주 세분화 방법의 성능을 측정하는 데 사용하는 평가지표로서, IoU 값들에 대한 평균값을 의미함
세부 내용	<ul style="list-style-type: none"> IoU(intersection-over-union)는 Jaccard Index라고도 알려져 있으며, Segmentation과 Object Detection에서 많이 활용되는 품질지표임
계산식(방식)	$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$ 
관련 출처	<ul style="list-style-type: none"> “Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation”. ResearchGate. 2021년 12월 23일 접속. https://www.researchgate.net/publication/311531910_Optimizing_Intersection-Over-Union_in_Deep_Neural_Networks_for_Image_Segmentation

〈표 III-114〉 mIoU의 개념 및 세부 내용

5.2.2.3 Recall (재현율)

구분	내용			
개념	• 분류 모델에 일반적으로 사용하며, 올바르게 분류했는지 측정하는 평가지표			
세부 내용	• 분류 모델에 일반적으로 사용되는 성능 메트릭인 재현율(Recall)은 올바르게 분류된 양성 비율로서, 재현율은 일반적으로 “진정한 긍정률”, “민감도” 및 “적중률”이라고도 함			
계산식(방식)			Predicted Condition	
			Positive(PP)	Negative (PN)
	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)
		Negative (N)	False Positive (FP)	True Negative (TN)
$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$				
관련 출처	• “Recall”. C3.ai. 2021년 12월 23일 접속. https://c3.ai/glossary/data-science/recall/			

〈표 III-115〉 Recall의 개념 및 세부 내용

5.2.2.4 Precision (정밀도)

구분	내용			
개념	• 분류 모델에 일반적으로 사용되며, 정확하게 예측했는지 측정하는 평가지표			
세부 내용	• 분류 모델에 일반적으로 사용되는 성능지표인 정밀도는 정확하게 예측한 양성 비율로서, 정밀도는 일반적으로 재현율과 반비례 관계임			
계산식(방식)			Predicted Condition	
			Positive(PP)	Negative (PN)
	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)
		Negative (N)	False Positive (FP)	True Negative (TN)
		$\text{Precision} = \frac{TP}{TP + FP} = \frac{\text{TP}}{\text{TP} + \text{FP}}$		
관련 출처	• “Precision”. C3.ai. 2021년 12월 23일 접속. https://c3.ai/glossary/machine-learning/precision/			

〈표 III-116〉 Precision의 개념 및 세부 내용

5.2.2.5 OA (Overall Accuracy)

구분	내용			
개념	• 각각의 객체에 대한 검출 정확도 총합으로서 표현되는 품질지표			
세부 내용	• Overall Accuracy(OA)는 기본적으로 모든 객체를 대상으로, 각각의 객체 검출 이 올바르게 매핑 되었는지 총합으로 계산함 • 해당 품질지표는 계산 및 이해가 쉽고 전체적인 정확도를 표현하는데 유용하지만, 편향된 정확도에 대해서 편중(bias)이 발생할 수 있음			
계산식(방식)			Predicted Condition	
			Positive(PP)	Negative (PN)
	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)
		Negative (N)	False Positive (FP)	True Negative (TN)
		$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$		
관련 출처	• “Accuracy Metrics”. Humboldt State University – GSP216 Introduction to remote sensing, 2021년 12월 23일 접속, http://gsp.humboldt.edu/olm_2019/courses/GSP_216_Online/lesson6-2/metrics.html			

〈표 III-117〉 OA의 개념 및 세부 내용

5.2.2.6 Kinship Verification 정확도

구분	내용													
개념	<ul style="list-style-type: none">주어진 얼굴 이미지 쌍 간의 친족 관계를 확인하는 평가지표													
세부 내용	<ul style="list-style-type: none">Kinship Verification 정확도(자동 친족 확인 시스템)는 주어진 얼굴 이미지 쌍 간의 친족 관계를 확인하도록 설계된 평가지표로서, 부모와 유전자 중복으로 자식에게서 비슷하게 나타나는 ‘외모’, ‘좋아하는 것과 싫어하는 것’, ‘행동’ 및 ‘목소리’ 등 많은 유사 특성을 검출하여 정확도를 측정함													
계산식(방식)	<table><tr><th colspan="2" rowspan="2"></th><th colspan="2">Predicted Condition</th></tr><tr><th>Positive(PP)</th><th>Negative (PN)</th></tr><tr><th rowspan="2">Actual Condition</th><th>Positive (P)</th><td>True Positive (TP)</td><td>False Negative (FN)</td></tr><tr><th>Negative (N)</th><td>False Positive (FP)</td><td>True Negative (TN)</td></tr></table> $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$			Predicted Condition		Positive(PP)	Negative (PN)	Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)	Negative (N)	False Positive (FP)	True Negative (TN)
				Predicted Condition										
		Positive(PP)	Negative (PN)											
Actual Condition	Positive (P)	True Positive (TP)	False Negative (FN)											
	Negative (N)	False Positive (FP)	True Negative (TN)											
관련 출처	<ul style="list-style-type: none">Moumita Mukherjee et al. Electronics and Telecommunication Engineering, National Institute of Technology, “Kinship verification using Compound Local Binary Pattern and Local Feature Discriminant Analysis”, 2019													

※ Kinship Verification Accuracy란, Accuracy 품질지표를 유전성 분석 분야에 적용한 것으로 별도의 품질지표로 보기 어려울 수도 있으나, '인공지능 학습용 데이터 구축사업' 내 다양한 품질지표를 소개하는 측면에서 해당 내용을 작성함

〈표 III-118〉 Kinship Verification 정확도의 개념 및 세부 내용

5.2.2.7 Overall Pixel 정확도

구분	내용
개념	<ul style="list-style-type: none"> 올바르게 어노테이션이 수행되었는지 지정된 픽셀의 비율을 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> 전체 픽셀 정확도(Overall Pixel Accuracy)는 올바르게 레이블이 지정된 픽셀의 비율을 측정하여, 검출에 대한 정확도를 판단함 다만, 불균형한 가공 산출물로 인한 결과의 편중(bias)이 발생할 수 있음
계산식(방식)	$OP = \frac{\sum_{i=1}^L C_{ii}}{\sum_{i=1}^L G_i}$ <p>※ 계산식 관련 세부 내용은 아래 출처 참조</p>
관련 출처	<ul style="list-style-type: none"> Gabriela Csurka et al. Xerox Research Centre Europe, "What is a good evaluation measure for semantic segmentation?", 2013

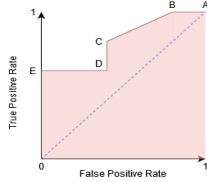
〈표 III-119〉 Overall Pixel 정확도의 개념 및 세부 내용

5.2.2.8 mAP (mean Average Precision)

구분	내용
개념	<ul style="list-style-type: none"> 각 쿼리에 대한 평균 정밀도 점수의 평균을 의미하는 평가지표로서, 클래스가 여러 개인 경우 각 클래스당 AP를 구한 뒤 그 합을 클래스의 개수로 나누어 계산함
세부 내용	<ul style="list-style-type: none"> 쿼리 집합의 평균 정밀도(mAP, mean Average Precision)는 각 쿼리에 대한 평균 정밀도 점수의 평균으로서, 아래 수식에서 Q는 쿼리의 수를 의미함 수식이 의미하는 바는, 주어진 쿼리 q에 대해 해당 평균 정밀도(AP)를 계산한 다음, 이러한 모든 AP 점수의 평균을 mAP라는 단일 숫자로 제공하여 모델이 얼마나 좋은지를 수치화함
계산식(방식)	$mAP = \frac{\sum_{q=1}^Q AveP(q)}{Q}$
관련 출처	<ul style="list-style-type: none"> “Evaluation measures”,Wikipedia. 2021년 12월 23일 접속. https://en.wikipedia.org/wiki/Evaluation_measures_(information_retrieval)#Average_precision

〈표 III-120〉 mAP의 개념 및 세부 내용

5.2.2.9 AUROC (Area Under Receiver Operating Characteristic)

구분	내용
개념	<ul style="list-style-type: none"> 이진 분류 문제에 대한 평가지표
세부 내용	<ul style="list-style-type: none"> ROC(Receiver Operator Characteristic) 곡선은 이진 분류 문제에 대한 평가 지표로서, 다양한 임계값에서 FPR에 대한 TPR을 플롯하고 본질적으로 ‘신호’를 ‘잡음’에서 분리하는 확률 곡선을 의미함 AUROC는 AUC(Area Under the Curve)와 같으며, 분류기가 클래스를 구별하는 능력의 척도이며 ROC 곡선의 요약으로 사용되는데, AUROC의 값이 최대 값인 1에 가까울수록 Positive 클래스와 Negative 클래스를 구별하는 모델의 성능이 향상됨
계산식(방식)	
관련 출처	<ul style="list-style-type: none"> “AUC-ROC Curve – The Star Performer!”, Analytics Vidhya. 2021년 12월 23일 접속. https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/

〈표 III-121〉 AUROC의 개념 및 세부 내용

5.2.2.10 RMSE (Root Mean Squared Error)

구분	내용
개념	<ul style="list-style-type: none"> 예측 품질을 평가하기 위해, 실제 값에서 예측이 얼마나 떨어져 있는지 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> 제곱 평균 제곱근 오차 또는 제곱 평균 제곱근 편차는 예측 품질을 평가하는데 가장 일반적으로 사용되는 측정 방법 중 하나로서, 유클리드 거리를 사용하여 측정된 실제 값에서 예측이 얼마나 떨어져 있는지 측정함 RMSE를 계산하려면 각 데이터 포인트에 대한 잔차(예측과 진실의 차이)를 계산하고, 각 데이터 포인트에 대한 잔차의 노름(norm)을 계산한 뒤, 잔차의 평균을 계산하고 그 평균의 제곱근을 취함 RMSE는 각 예측 데이터 포인트에서 실제 측정값을 사용하고 필요로 하기 때문에 지도 학습 애플리케이션에서 일반적으로 사용될 수 있음
계산식	$RMSE = \sqrt{\frac{\sum_{i=1}^N \ y(i) - \hat{y}(i)\ ^2}{N}}$
관련 출처	<ul style="list-style-type: none"> “Root Mean Square Error”. C3.ai. 2021년 12월 23일 접속. https://c3.ai/glossary/data-science/root-mean-square-error-rmse/


〈표 III-122〉 RMSE의 개념 및 세부 내용

5.2.2.11 OKS (Object Keypoint Similarity)

구분	내용
개념	<ul style="list-style-type: none"> 관절의 추정 좌표와 참값 좌표의 유사성을 나타내는 평가지표
세부 내용	<ul style="list-style-type: none"> OKS (Object Keypoint Similarity)는 관절의 추정 좌표와 참값 좌표의 유사성의 평균을 나타내는 평가지표이며, 인물의 추정 자세와 정답 자세가 완전히 일치하면 1이 됨 Object Detection의 IoU(Intersection over Union)와 유사하며 Keypoint Detection에서 유사성 측정 방법으로 사용됨
계산식(방식)	$OKS = \sum_i [\exp(\frac{-d_i^2}{2s^2k_i^2})\delta(v_i > 0)] / \sum_i \delta(v_i > 0)$
관련 출처	<ul style="list-style-type: none"> “Evaluate: Keypoints”, COCO Common Objects in Context, 2021년 12월 23일 접속, https://cocodataset.org/#keypoints-eval

〈표 III-123〉 OKS의 개념 및 세부 내용

5.2.2.12 PCK (Percentage of Correct Keypoints)

구분	내용
개념	<ul style="list-style-type: none"> 키포인트의 추정 좌표와 참값 좌표의 거리가 임계값 기준을 만족하는지 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> PCK(Percentage of Correct Keypoints)는 키포인트 관절점의 추정 좌표와 참값 좌표의 거리가 임계값 기준을 만족하는지 측정하며, 임계값은 바운딩박스의 사이즈를 통해 결정된다. <div style="display: flex; align-items: center;">  <div style="margin-left: 20px;"> <p>●: 키포인트 정답 좌표 ●: 키포인트 추정 좌표 □: 임계값 바운딩 박스</p> <p>추정좌표 \hat{x}가 정답 좌표 x 중심의 임계값 바운딩박스 안에 들어오는 검사 →PCK는 바운딩박스 안에 들어오는 정답 확률</p> </div> </div>
계산식(방식)	<ul style="list-style-type: none"> PCKh@0.5 is when the threshold = 50% of the head bone link PCK@0.2 = Distance between predicted and true joint < 0.2 * torso diameter
관련 출처	<ul style="list-style-type: none"> Sergio Escalera Guerrero, ChaLearn, Computer Vision Center (UAB) and University of Barcelona, "Tutorial on Datasets, Annotations and Metrics in Computer Vision". 2015

〈표 III-124〉 PCK의 개념 및 세부 내용

5.2.3 비디오 데이터

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
객체 인식 (Object Detection)	Bounding Box	AUROC	94.55
		mAP	98.2
		mAP@IoU	99.835
		mIoU	80.43
		Sensitivity	92
		Specificity	99
		정확도	91.53
		Fβ-점수	99.29
	Polygon	AUROC	95.38
		DSC	94.14
		MAE	99.6

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
		mAP	98.8
		mAP@IoU	99.45
		mIoU	98.9
		NRMSE	92.5
		RE	95.7
		Sensitivity	86.16
		정확도	91.53
		F β -점수	78.04
	Polyline	mAP	79.17
		mAP@IoU	83.1
		RMSE	0.14
이미지 분류 (Image Classification)	Bounding Box	AUROC	100
		IoU	89
		mAP	90.1
		RMSE	99.944
		Sensitivity	90.32
		Top-K 정확도	100
		정확도	97.5
		F β -점수	99.89
	Keypoint	MRE	1.737
		SMAPE	85.91
	Polygon	AUROC	95.33
		DSC	99.47
		IoU	85
		Top-K 정확도	80.12
		정확도	100
		F β -점수	98.45
	Polyline	Correlation	83
		DSC	95
		Sensitivity	81
		F β -점수	82.61
행동 인식 (Action Recognition)	Bounding Box	1DF1	30.5
		AP	77

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
		mAP	82.2
		MOTA	60.69
		Top-K 정확도	96.99
		$F\beta$ -점수	98.92
	Keypoint	BLEU	16.33
		mAP	82.2
		Mean Gaze Tracking Error	1.64
		Mean Position Error(pixel)	3.26
		PCK	97.6
		Top-K 정확도	96.99
		정확도	97.21
		$F\beta$ -점수	80.87
	Polygon	mIoU	89.2
		$F\beta$ -점수	96.33
	Polyline	mIoU	89.2
		MOTA	60.69
		$F\beta$ -점수	80.87
비디오 분류 (Video Classification)	Bounding Box	AUROC	100
		mA	84.59
		Sensitivity	100
		Specificity	100
		정확도	91.53
		$F\beta$ -점수	77.41
	Polyline	정밀도	88
	Tagging	micro F1	51.38
		$F\beta$ -점수	65.51
자세 추정 (Pose Estimation)	Bounding Box	mASS (meanAverageShapeSimilarity)	68.82
		정확도	88.51
		$F\beta$ -점수	99.32
	Keypoint	mASS	68.82

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
		(meanAverage ShapeSimilarity)	
		정확도	88.51
		Fβ-점수	99.32
비디오 인식 (Video Recognition)	Polygon	PSNR	33.2
	Segmentation	PSNR	41.3
키포인트 검출 (Keypoint Detection)	Keypoint	mAP	92.98
	Segmentation	mAP	87.71

※ 해당 내용은 한국지능정보사회진흥원에서 공모하는 ‘인공지능 학습용 데이터 구축사업’에 대한 품질검증결과를 기반으로 작성되었으며, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음

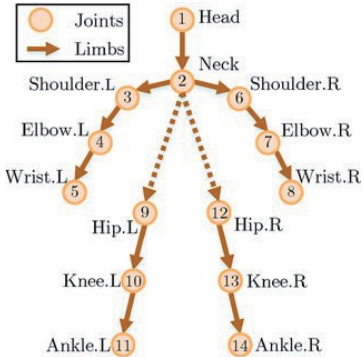
〈표 Ⅲ-125〉 비디오 데이터 유효성 품질지표 구성

5.2.3.1 IDF1 (Identification F1-Score)

구분	내용
개념	<ul style="list-style-type: none"> 참값 정확도의 평균 개수를 계산하여 식별 정밀도를 표현하는 평가지표
세부 내용	<ul style="list-style-type: none"> ID 정밀도(Identification Precision)와 ID 재현율(Identification Recall)은 trade-off 추적에 초점을 맞춘 평가지표인 반면, IDF1은 ID 정밀도와 ID 재현율 간의 조화 평균(harmonic mean)을 계산함
계산식(방식)	$IDF_1 = \frac{2 IDTP}{2 IDTP + IDFP + IDFN}$ $IDFN = \sum_{\tau \in AT} \sum_{t \in T_\tau} m(\tau, \gamma_m(\tau), t, \Delta)$ $IDFP = \sum_{\gamma \in AC} \sum_{t \in T_\gamma} m(\tau_m(\gamma), \gamma, t, \Delta)$ $IDTP = \sum_{\tau \in AT} \text{len}(\tau) - IDFN = \sum_{\gamma \in AC} \text{len}(\gamma) - IDFP$
관련 출처	<ul style="list-style-type: none"> Ergys Ristani et al. “Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking”, 2016

〈표 Ⅲ-126〉 IDF1의 개념 및 세부 내용

5.2.3.2 MPJPE (Mean Per Joint Position Error)

구분	내용
개념	<ul style="list-style-type: none"> 관절의 추정 좌표와 참값 좌표의 거리의 평균값을 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> MPJPE (Mean Per Joint Position Error)는 모든 관절의 추정 좌표와 참값 좌표의 거리(단위: mm)를 평균하여 측정하는 평가지표이며, 값이 작을수록 정확도가 높다고 판단함 Estimated and Ground-truth 3D Pose의 root 관절을 정렬한 후 계산함 
계산식(방식)	$MPJPE = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \ (J_i^{(t)} - J_{root}^{(t)}) - (\hat{J}_i^{(t)} - \hat{J}_{root}^{(t)}) \ _2.$
관련 출처	<ul style="list-style-type: none"> Jinbao Wang et al. "Deep 3D human pose estimation: A review", 2020

〈표 III-127〉 MPJPE의 개념 및 세부 내용

5.2.3.3 NME (Normalized Mean Errors)

구분	내용
개념	<ul style="list-style-type: none"> 작업자에 따라서 라벨링 결과가 일정하지 않아, 해당 데이터의 불확실성 특성을 판단하기 위한 평가지표
세부 내용	<ul style="list-style-type: none"> NME(Normalized Mean Errors), 즉 정규화 오차는 Facial Landmark Detection 분야에서 많이 사용됨 NME를 사용하는 이유는 Facial Landmark의 특성상 데이터 구축 과정에서 작업자에 따라 라벨링 결과의 위치가 일정하지 않으므로, 해당 데이터의 특성을 반영하기 위해서 사용함
계산식(방식)	$NME = \frac{1}{N} \sum_{i=1}^N \frac{\sqrt{\Delta x_i^2 + \Delta y_i^2}}{d}$ <p>where d denotes the normalized term and N is the number of facial landmarks.</p>
관련 출처	<ul style="list-style-type: none"> Shenqi Lai et al, "Enhanced Normalized Mean Error loss for Robust Facial Landmark detection", 2019

〈표 III-128〉 NME의 개념 및 세부 내용

5.2.3.4 MOTA (Multiple Object Tracking Accuracy)

구분	내용
개념	<ul style="list-style-type: none"> 다수 객체를 추적하는데 있어 정확도를 직관적으로 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> MOTA(Multiple Object Tracking Accuracy)는 다양한 목적 함수 평가 범위에서 각각 여러 성능 측정에 따라 확률 최적화 알고리즘을 조정하는 데 특화되어 있음
계산식	$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t}$ <ul style="list-style-type: none"> 여기서 m_t는 t시간에 발생한 추적 실책 수, fp_t는 t시간에 발생한 거짓 양성 수, mme_t는 t시간에 발생한 부정합 추적결과 수를 의미함
관련 출처	<ul style="list-style-type: none"> Keni Bernardin et al, "Multiple Object Tracking Performance Metrics and Evaluation in a Smart Room Environment", 2006.

〈표 III-129〉 MOTA의 개념 및 세부 내용

5.2.4 오디오 데이터

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
오디오 분류 (Audio Classification)	Class Labeling	정확도	94.445
		Fβ-점수	97
	Tagging	정확도	94
	Text Summary	정확도	98.99
		Fβ-점수	99.35
		정밀도	99.56
		재현율	99.15
언어 이해 (Language Understanding)	Class Labeling	CER(문자)	91.18
		EER	97.12
		WER(단어)	88.99
	Sequence Labeling	CER(문자)	90.24
		EER	97.12
		Fβ-점수	34
		WER(단어)	79.79
	Speech to Text	CER(문자)	90.24

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
		EER	97.12
		ESTOI	82
		WER(단어)	79.01
	Tagging	WER(단어)	88.99
음성 인식 (Speech Recognition)	Class Labeling	BLEU	23.7
		CER(문자)	96.6
	Machine Translation	BLEU	23.7
	Sequence Labeling	BLEU	23.7
		WER(단어)	85.85
	Speech to Text	BLEU	23.7
		CER(문자)	96.6
		WER(단어)	84.83
텍스트 분류 (Text Classification)	Class Labeling	CER(문자)	99.22
	Speech to Text	CER(문자)	84.59
음성 합성 (Speech Synthesis)	Class Labeling	MOS	95.9
	Sequence Labeling	MOS	95.91
	Speech to Text	MOS	95.91
	Tagging	MOS	95.91

※ 해당 내용은 한국지능정보사회진흥원에서 공모하는 '인공지능 학습용 데이터 구축사업'에 대한 품질검증결과를 기반으로 작성되었으며, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음

〈표 III-130〉 오디오 데이터 유효성 품질지표 구성

5.2.4.1 CER(Character Error Rate, 문자)

구분	내용
개념	• 주어진 문장에 대해 공백을 포함한 총 문자 수(n)를 필요한 문자의 최소 삽입(i), 대체(s) 및 삭제(d) 수와 비교하여 글자 오류율을 측정하는 평가지표
세부 내용	• CER(Character Error Rate)는 글자 오류율을 측정하는 지표로 주로 한글과 같이 조사를 사용하는 언어에 사용
계산식(방식)	$CER = \frac{S + D + I}{N}$
관련 출처	• “Metric: cer”, Hugging Face, 2021년 12월 23일 접속, https://huggingface.co/metrics/cer

〈표 III-131〉 CER의 개념 및 세부 내용

5.2.4.2 WER (Word Error Rate, 글자)

구분	내용
개념	<ul style="list-style-type: none"> 주어진 문장에 대해 공백을 포함한 총 단어 수(n)를 필요한 단어의 최소 삽입(i), 대체(s) 및 삭제(d) 수와 비교하여 단어 오류율을 측정하는 평가지표
세부 내용	<ul style="list-style-type: none"> WER(Word Error Rate)는 단어 오류율을 측정하는 지표로 영어에 많이 사용
계산식(방식)	$WER = \frac{S_w + D_w + I_w}{N_w}$
관련 출처	<ul style="list-style-type: none"> “Metric: wer”, Hugging Face, 2021년 12월 23일 접속, https://huggingface.co/metrics/wer

〈표 III-132〉 WER의 개념 및 세부 내용

5.2.4.3 Perplexity (복잡도)

구분	내용
개념	<ul style="list-style-type: none"> 확률 분포 또는 확률 모형이 표본을 얼마나 잘 예측하는지 나타내는 평가지표로서, Perplexity(복잡도)가 낮으면 확률 분포가 표본을 잘 예측한다는 것을 나타냄
세부 내용	<ul style="list-style-type: none"> 기본적으로 외부 데이터를 이용하지 않고 언어모델 자체의 성능을 평가하겠다는 아이디어이기 때문에, 언어모델을 기반으로 한 다양한 Task의 간접적인 성능 지표로 사용할 수는 있지만 특정 Task가 정해진 경우에는 그 Task에 보다 적합한 품질지표를 사용하는 것이 일반적인 추세임 즉, 전통적인 n-gram 언어모델에서부터 사용하던 metric이지만 최근에는 그다지 많이 사용되지 않음
계산식(방식)	$PP(W) = \sqrt[N]{\frac{1}{P(w_1, w_2, \dots, w_N)}}$
관련 출처	<ul style="list-style-type: none"> “Perplexity in Language Models”, towards data science, 2021년 12월 23일 접속, https://towardsdatascience.com/perplexity-in-language-models-87a196019a94

〈표 III-133〉 Perplexity의 개념 및 세부 내용

5.2.4.4 MOS (Mean Opinion Score)

구분	내용
개념	<ul style="list-style-type: none"> 음성 통화 성능평가에 주로 사용되는 통계에 의한 주관적 평가 지표
세부 내용	<ul style="list-style-type: none"> MOS(Mean Opinion Score)는 지난 수십 년 동안 전반적인 음성 통화 품질을 측정하기 위해 가장 일반적으로 사용되었음

구분	내용
	<ul style="list-style-type: none"> 해당 품질지표는 ITU-T(International Telecommunications Union)에서 표준화되어 공개된 지표임
계산식(방식)	$MOS = \frac{\sum_{n=1}^N R_n}{N}$
관련 출처	<ul style="list-style-type: none"> "M.O.S. MEAN OPINION SCORE", Arimas, 2021년 12월 23일 접속, https://arimas.com/2017/08/01/mos-mean-opinion-score/

〈표 III-134〉 MOS의 개념 및 세부 내용

5.2.5 3D 데이터

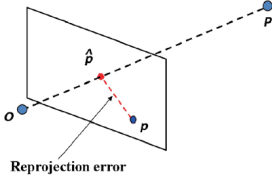
알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
3D 객체 인식 (3D Object Detection)	Bounding Box	mAP	76.6
		mAP@IoU	95.72
		mIoU	85.1
	Cuboid	정확도	99.25
		mAP	99.99
		mAP@IoU	95.72
		mIoU	97.01
	Polygon	정확도	99.25
		mAP	81.2
		mAP@IoU	87.38
		mIoU	68.39
	Polyline	mAP@IoU	79
		mIoU	85.1
3D 포인트 클라우드 (3D Point Cloud Classification)	Cuboid	mAcc	84.2
		mIoU	76.45
		OA	98.7
3D 자세 추정 (3D Pose Estimation)	Bounding Box	2D reprojection 정확도	90.23
		TOP-K 정확도	82.31
	Polygon	2D reprojection 정확도	90.23
3D 객체 복원 (3D Object Reconstruction)	Cuboid	mAP	50

알고리즘 Task 분야	라벨 유형	유효성 측정지표	정량목표(예시)
2D 이미지를 통한 시점 합성 (2D View Synthetic)	Bounding Box	$F\beta$ -점수	91.3
	Cuboid	mAP	60
	Polygon	$F\beta$ -점수	91.3
3D 키포인트 검출 (3D Keypoint Detection)	Keypoint	AP@oks	99
		mAP	97.55

※ 해당 내용은 한국지능정보사회진흥원에서 공모하는 '인공지능 학습용 데이터 구축사업'에 대한 품질검증결과를 기반으로 작성되었으며, 일반적으로 인공지능 분야에서 통용되는 품질 기준과 상이할 수 있음

〈표 III-135〉 3D 데이터 유효성 품질지표 구성

5.2.5.1 2D Reprojection Error

구분	내용
개념	<ul style="list-style-type: none"> 투영된 점과 측정된 점 사이의 거리에 해당하는 기하학적 오차를 측정하는 평가 지표
세부 내용	<ul style="list-style-type: none"> 3D 점의 추정치가 얼마나 근접하게 점의 실제 투영을 재생하는지를 정량화하기 위해 사용됨 재투영 오류(Reprojection Error)를 최소화하며, 두 이미지 사이의 점 대응에서 오류를 추정하는 데 사용됨
계산식(방식)	 <p style="text-align: center;">Reprojection error</p>
관련 출처	<ul style="list-style-type: none"> "Augmented Lagrangian-based Algorithm for Projective Reconstruction from Multiple Views with Minimization of 2D Reprojection Error". SEMANTIC SCHOLAR. 2021년 12월 23일 접속. https://www.semanticscholar.org/paper/Augmented-Lagrangian-based-Algorithm-for-Projective-Mai-Hung/fb3b33345a25d36cdcece5966abf8086d2d0f36

〈표 III-136〉 2D Reprojection Error의 개념 및 세부 내용

6 기타

6.1 합성 데이터 개념

- 합성 데이터(Synthetic data)란 실제 데이터(Real data)의 대체재로서 컴퓨터 시뮬레이션이나 알고리즘을 통해 생성하는 임의의 데이터를 의미

* MIT Technology Review는 2022년 10대 혁신 기술 중 하나로 '인공지능을 위한 합성 데이터(Synthetic Data for AI)'를 선정



[그림 III-1] 실제 데이터와 합성 데이터 비교

- 합성 데이터는 실제 데이터와 같은 분포를 가지고 있기 때문에 실제 자료와 비슷한 결과를 얻을 수 있어 인공지능 모델 학습 시 사용
 - 특히 개인정보가 포함된 데이터는 획득하거나 사용하는데 제약이 있어 개인정보보호 법령으로부터 자유로운 합성 데이터를 활용
- 합성데이터는 완전 합성 데이터(Fully synthetic data), 부분 합성 데이터(Partially synthetic data), 복합 합성 데이터(Hybrid synthetic data) 등으로 분류됨
 - 완전 합성 데이터(Fully synthetic data) : 공개하려고 하는 데이터에 측정된 실제 데이터가 하나도 없이 모두 가상으로 생성된 데이터
 - 부분 합성 데이터(Partially synthetic data) : 공개하려는 변수 중 일부만을 선택하여 합성 데이터로 대체한 데이터
 - 복합 합성 데이터(Hybrid synthetic data) : 일부 변수들의 값을 합성 데이터로 생성하고 생성된 합성 데이터와 실제 데이터를 모두 이용 하여 또 다른 일부 변수들의 값을 다시 도출하는 방법으로 생성

6.2 '22년 인공지능 학습용 데이터 중 합성 데이터 품질지표(의미 정확성, 유효성)

과제명	품질지표		데이터 유형				
	의미 정확성	유효성	비디오	이미지	오디오	텍스트	3D
주요질환 이미지 합성데이터(X-ray)	정확도(%)	FID VTT-정확도(%) 분류모델의유사도-정확도(%)	○	○			
주요질환 이미지 합성데이터(CT)	정확도(%)	FID VTT-정확도(%) 분류모델의유사도-정확도(%)	○	○			
주요질환 이미지 합성데이터(MRI)	정확도(%)	FID VTT-정확도(%) 분류모델의유사도-정확도(%)	○	○			
감성 및 발화스타일 동시 고려 음성합성 데이터	정확도(%)	MOS(likert 5점)			○	○	
한국인 얼굴 합성을 위한 발화 모습 이미지	정확도(%) IoU(%)	FID	○				

〈표 III-137〉 '22년 인공지능 학습용 데이터 합성 데이터 품질지표

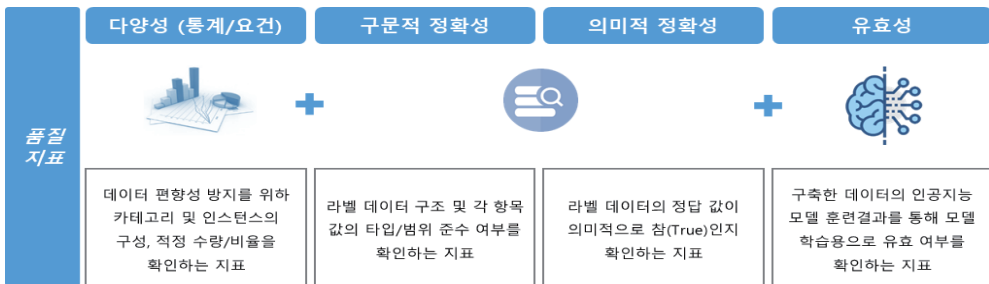
※ 합성데이터는 데이터 증강(Data Augmentation)에 활용되는 데이터로 실제 데이터(Real Data)의 지표 준용

부록 2 | 품질지표 설정 가이드

1 개요

1.1 품질지표

- 인공지능 모델(알고리즘)이 활용 가능한 형태의 데이터를 구축하기 위한 다양성(통계/요건), 구문 정확성, 의미 정확성, 유효성 확보 필요



[그림 III-2] 인공지능 학습용 데이터 품질지표

- 본 가이드는 인공지능 학습데이터 구축사업 추진 시 구축 유형에 따른 품질지표 설정을 돕기 위해 다양성(통계/요건), 구문 정확성, 의미 정확성, 유효성별 지표 가이드를 제공



[그림 III-3] 인공지능 학습용 데이터 품질지표 설정 가이드 구성

1.2 학습데이터 구축 유형

1.2.1 데이터 유형별 학습 Task

- 인공지능 학습데이터 구축 시 ‘텍스트’, ‘이미지’, ‘비디오’, ‘오디오’, ‘3D’ 데이터 종류별로 적합한 학습 Task를 선택함
- 이때, 학습 Task는 인공지능 학습데이터 구축 목표 또는 임무에 적합하도록 선택함

학습 Task	개념	데이터 유형				
		텍스트	이미지	비디오	오디오	3D
텍스트 분류 (Text Classification)	<ul style="list-style-type: none"> 문장이나 문서를 적절한 범주로 지정하는 작업으로, 범주는 선택한 데이터 세트에 따라 다르며 주제 범위가 다양할 수 있음 	●			○	
이미지 분류 (Image Classification)	<ul style="list-style-type: none"> 특정 이미지를 알고리즘에 입력했을 때, 그 이미지가 어떤 클래스에 속하는지 알려주는 작업 		●	○		
비디오 분류 (Video Classification)	<ul style="list-style-type: none"> 프레임이 주어진 비디오에서 특정 프레임 구간을 선택하고 레이블을 부여하는 작업 		●	○		
오디오 분류 (Audio Classification)	<ul style="list-style-type: none"> 오디오 정보를 특정 시간 단위로 분할(오디오 클립)하여, 신호의 특이점을 분석하고 특정 클래스로 분류하는 작업 	●	●		○	
3D 포인트클라우드 분류 (3D Point Cloud Classification)	<ul style="list-style-type: none"> 포인트 클라우드를 알고리즘에 입력했을 때, 어떤 클래스에 속하는지 알려주는 작업 					●
객체 인식 (Object Detection)	<ul style="list-style-type: none"> 이미지 내에서 특정 클래스의 객체 인스턴스를 인식하는 작업 		●	○		
키포인트 검출 (Keypoint Detection)	<ul style="list-style-type: none"> 동작 인식을 위해 필요한 주요 키포인트를 검출하는 작업 		●	○		
3D 객체 인식 (3D Object Detection)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 특정 클래스의 객체 인스턴스를 감지하는 작업 		○	○		●
3D 키포인트 검출 (3D Keypoint Detection)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 동작 인식을 위해 필요한 주요 키포인트를 검출하는 작업 					○
자세 추정 (Pose Estimation)	<ul style="list-style-type: none"> 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업 		●	○		
3D 자세 추정 (3D Pose Estimation)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업 					○

학습 Task	개념	데이터 유형				
		텍스트	이미지	비디오	오디오	3D
순차적 레이블링 (Sequence Labeling)	• 입력 시퀀스에 대하여 레이블 시퀀스를 각각 부여하는 작업으로, 태깅 작업이 대표적인 순차적 레이블링의 예시임	●				
질의 응답 (Question Answering)	• 질문(일반적으로 독해 질문)에 답변하는 작업으로서, 제공된 컨텍스트를 기반으로 답변할 수 없는 질문이 제시되면 기권함	●			○	
얼굴 인식 (Face Recognition)	• 얼굴을 사용하여 개인의 신원을 식별하거나 확인하는 작업		●			
행동 인식 (Action Recognition)	• 비디오 데이터 스트림에서 객체 행동의 감지 및 인식을 하는 작업		●	○		
비디오 인식 (Video Recognition)	• 비디오를 활용하여 실시간 객체 식별 및 비디오 프레임을 분류하는 작업			●		
음성 인식 (Speech Recognition)	• (전사)사람이 말하는 음성 언어를 인공지능이 해석해 그 내용을 문자 데이터로 전환하는 작업	●			○	
	• (자연어처리)사람이 말하는 음성 언어를 기계가 읽을 수 있는 형식으로 변환하는 작업	●			●	
3D 객체 복원 (3D Object Reconstruction)	• 영상/이미지/포인트 클라우드로부터 객체(물체, 사람 등)의 3차원 구조를 복원(reconstruction)하는 작업		○	○		○
텍스트 요약 (Text Summary)	• 상대적으로 큰 원문을 핵심 내용만 간추려서 상대적으로 작은 요약문으로 변환하는 작업	●			○	
음성 합성 (Speech Synthesis)	• 컴퓨터 또는 기타 장치를 사용하여 사람의 음성을 인공적으로 시뮬레이션하는 작업	●			●	
2D 이미지를 통한 시점 합성 (2D View Synthetic)	• 컴퓨터 또는 기타 장치를 사용하여 이미지를 인공적으로 3D로 전환하는 작업		○			○
기계 번역 (Machine Translation)	• 인공지능을 이용해 인간이 사용하는 자연어를 다른 자연어로 번역하는 작업	●			○	

〈범례〉 ○ : 원시데이터(획득/수집한 데이터 포맷), ● : 원천데이터(학습 시 활용하는 데이터 포맷)

〈표 III-138〉 데이터 유형별 학습 Task 종류

1.2.2 라벨링 방식

- 라벨링은 구축사업 목적 및 특성에 따라 다양한 라벨링 방식을 사용할 수 있으며, 대표적인 라벨링 방식은 다음과 같음

라벨링 유형	설명
세그멘테이션 (Segmentation)	• 이미지, 영상 데이터의 모든 객체를 구분하고 클래스 부여
바운딩 박스 (Bounding Box)	• 객체를 직사각형 박스안에 포함되도록 작업하여 클래스 부여
폴리라인 (Polyline)	• 여러 개의 점을 가진 선을 활용하여 인도, 차선, 경로 등 선형 객체를 구분하고 인식하기 위한 작업
폴리곤 (Polygon)	• 다각형 모양으로 객체의 테두리를 따라 점을 찍어 여백 없이 정확히 객체만을 구분하여 클래스 부여
키포인트 (Keypoint)	• 안면, 신체 등 객체의 주요 특징점을 지정하는 정밀하고 섬세한 작업
큐보이드 (Cuboid)	• 2D로 작업할 수 없는 3D 객체들을 정육면체로 생성
전사 (Transcription)	• 음성을 텍스트로 변환(전사)
태깅 (Tagging)	• 원본 지문(텍스트)에서 질문에 대한 답 생성 • 데이터(객체)의 주제나 카테고리 분류

〈표 III-139〉 대표적인 라벨링 유형

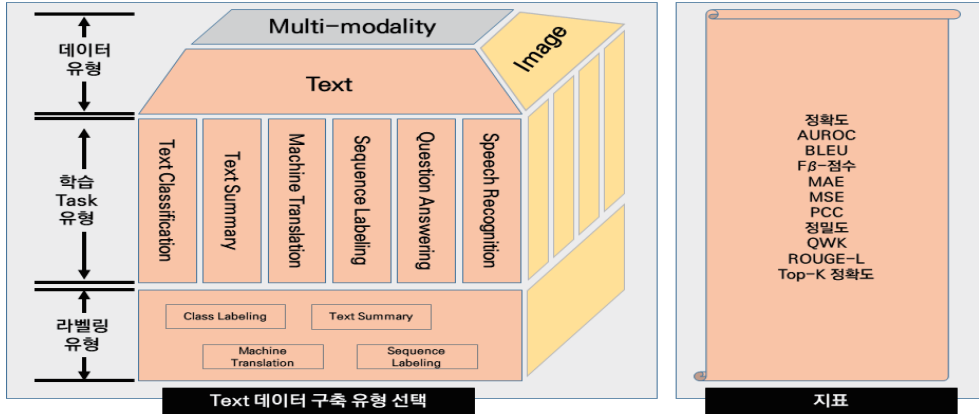
1.2.3 데이터, 학습 Task, 라벨링 유형 및 지표

1.2.3.1 텍스트 데이터

- 텍스트 데이터는 정보검색, 대화분석, 질의응답, 명령어 이해, 언어모델 학습 등의 자연어 처리 AI 기술개발을 위해 구축되며, 주요 학습 Task 유형으로는 ‘텍스트 분류(Text Classification)’, ‘텍스트 요약(Text Summary)’, ‘기계 번역(Machine Translation)’, ‘순차적 레이블링(Sequence Labeling)’, ‘질의 응답(Question Answering)’, ‘음성 인식(Speech Recognition)’ 등이 있음

※ 기타 학습 Task 유형으로 광학문자인식(Optical Character Recognition), 정보추출(Information Extraction), 텍스트 생성(Text Generation) 등이 있음

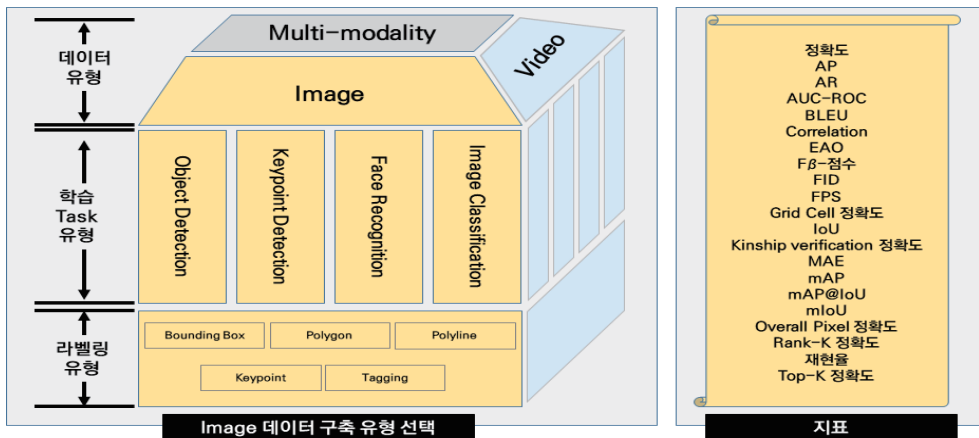
- 텍스트 데이터는 학습 Task 유형별로 적합한 라벨링을 선택하는 것이 필요하며, 선택된 라벨링 유형에 따라 적합한 지표를 적용



[그림 III-4] 텍스트 데이터

1.2.3.2 이미지 데이터

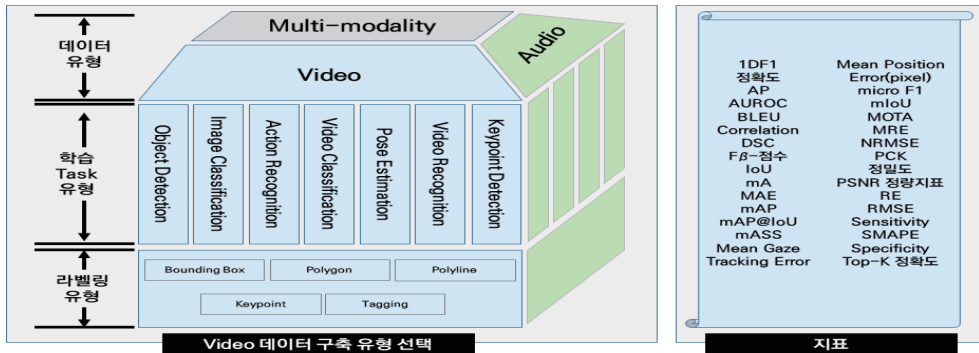
- 이미지 데이터는 시간의 흐름에 따라 지속적인 움직임을 담지 않고 피사체를 고정하여 촬영한 데이터를 의미하며, 주요 학습 Task 유형으로는 ‘객체 인식(Object Detection)’, ‘키포인트 검출(Keypoint Detection)’, ‘얼굴 인식(Face Recognition)’, ‘이미지 분류(Image Classification)’ 등이 있음
- 이미지 데이터는 학습 Task 유형별로 적합한 라벨링을 선택하는 것이 필요하며, 선택된 라벨링 유형에 따라 적합한 지표를 적용



[그림 III-5] 이미지 데이터

1.2.3.3 비디오 데이터

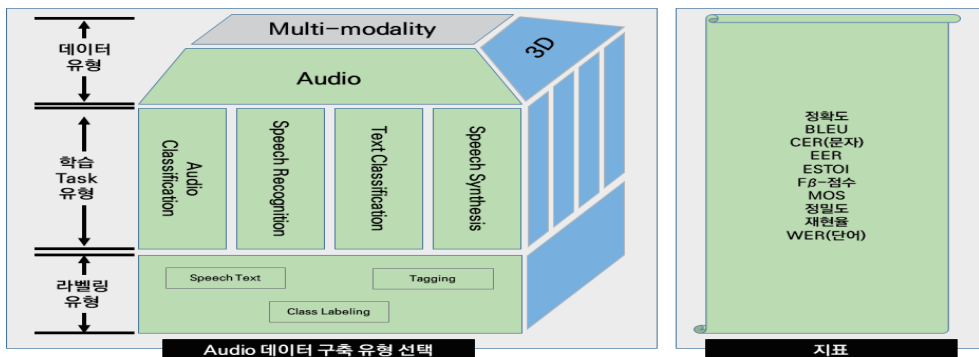
- 비디오 데이터는 연속적인 프레임을 연결하여 만든 스트리밍 이미지 데이터를 의미하며, 일정 시간동안 연결을 가진 동영상상을 뜻하기도 함. 주요 학습 Task 유형으로는 ‘객체 인식(Object Detection)’, ‘이미지 분류(Image Classification)’, ‘행동 인식(Action Recognition)’, ‘비디오 분류(Video Classification)’, ‘자세 추정(Pose Estimation)’, ‘비디오 인식(Video Recognition)’, ‘키포인트 검출(Keypoint Detection)’ 등이 있음
- 비디오 데이터는 학습 Task 유형별로 적합한 라벨링을 선택하는 것이 필요하며, 선택된 라벨링 유형에 따라 적합한 지표를 적용



[그림 III-6] 비디오 데이터

1.2.3.4 오디오 데이터

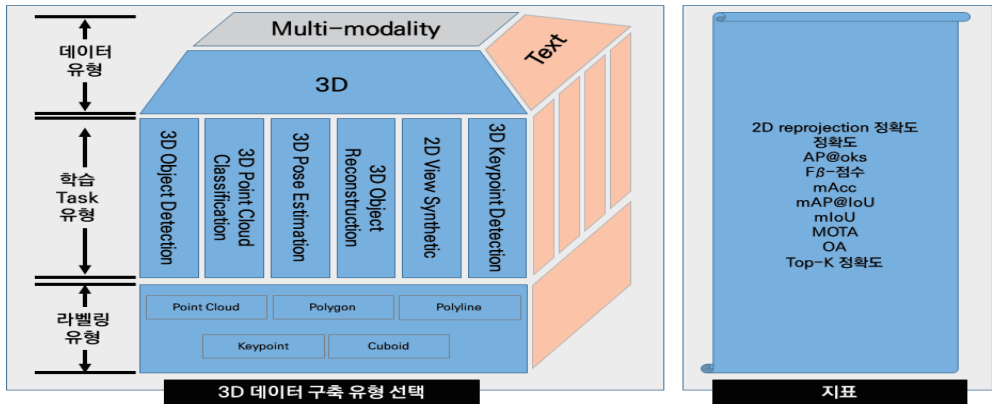
- 오디오 데이터는 녹음이나 녹취 등을 통해 획득된 자연어/비자연어 소리 데이터를 의미하며, 주요 학습 Task 유형으로는 ‘오디오 분류(Audio Classification)’, ‘음성 인식(Speech Recognition)’, ‘텍스트 분류(Text Classification)’, ‘음성 합성(Speech Synthesis)’ 등이 있음
- 오디오 데이터는 학습 Task 유형별로 적합한 라벨링을 선택하는 것이 필요하며, 선택된 라벨링 유형에 따라 적합한 지표를 적용



[그림 III-7] 오디오 데이터

1.2.3.5 3D 데이터

- 3D 데이터는 이미지, 비디오, 센서(LiDAR 등) 등을 통해 획득된 3D 이미지, 포인트 클라우드 데이터를 의미하며, 주요 학습 Task 유형으로는 ‘3D 객체 인식(3D Object Detection)’, ‘3D 포인트클라우드 분류(3D Point Cloud Classification)’, ‘3D 자세 추정(3D Pose Estimation)’, ‘3D 객체 복원(3D Object Reconstruction)’, ‘2D 이미지를 통한 시점 합성(2D View Synthetic)’, ‘3D 키포인트 검출(3D Keypoint Detection)’ 등이 있음
- 3D 데이터는 학습 Task 유형별로 적합한 라벨링을 선택하는 것이 필요하며, 선택된 라벨링 유형에 따라 적합한 품질지표를 적용



[그림 III-8] 3D 데이터

2 품질지표 선택가이드

2.1 다양성

2.1.1 개요

- 데이터는 활용 목적을 달성할 수 있도록 지도학습 계열 인공지능이 처리해야 하는 실제 세상의 데이터와 유사한 특성을 가진 데이터로 구성
- 이를 위해 학습에 유용한 모든 특성 정보를 포함하는 포괄성이 있어야 하며, 특성 정보가 학습에 유용한 범위에서 다양하게 변화하는 변동성이 있어야 하고, 학습에 지역적 편견, 사회적 편견, 인종적 편견 등의 의도가 없도록 편향성을 방지해야 함

〈참고: 특성정보〉

- 특성 정보란 데이터에 포함된 내용 중 학습에 활용되는 속성으로 다음과 같은 종류가 있음
 - 사물
 - 지역
 - 사람(성별, 연령대 등)
 - 장소
 - 시간(오전, 오후, 야간 등)
 - 환경(실외, 실내, 날씨 등)
 - 언어 특성(방언, 외국어 등)

〈표 Ⅲ-140〉 다양성 항목 특성정보

- 인공지능 학습용 데이터 구축에서 다양성 항목 정의 시 사업 목표(임무), 데이터 유형, 어노테이션 대상 등을 고려하여 다양성을 정의해야 함. 이때 메타데이터와 관련된 항목의 경우 사전 계획을 수립하기 비교적 용이하기 때문에 달성 가능한 명확한 목표 수치를 설정
- 단, 어노테이션과 관련된 항목의 경우 사업 목표(임무)를 정의 시 사전 정의가 어렵기 때문에 구축 결과에 대한 주요 통계 정보를 제공하는 목적을 가지고 정의

2.1.2 다양성 항목 정의 절차

- 다양성 항목을 정의하는 절차는 ① 학습에 유용한 모든 특성 정보를 포함하는 포괄성 항목을 정의하고, ② 포괄성 항목에 대한 범위 및 범주를 고려한 변동성을 검토한 후 ③ 편향성을 방지하기 위한 데이터 비율 구성 결정의 절차로 수행

〈다양성 항목 도출 절차〉



〔그림 Ⅲ-9〕 다양성 항목 도출 절차

- ① 포괄성 항목 정의 시 학습에 유용한 모든 특성 정보를 고려하여 항목을 정의
 - ☞ 〈예〉 사물, 지역, 시간, 날씨, 화자, 도메인(임무에 따라 지역별 날씨 등과 같이 조합 항목도 가능)
- ② 변동성을 고려하기 위해 포괄성 항목을 범위 구분 후 범주화
 - ☞ 〈예〉 사물 세부 범주, 데이터 수집 지역(〈예〉 도심, 산간, 해변), 시간대 범위(〈예〉 오전, 오후, 야간, 새벽), 날씨 종류(〈예〉 맑음, 흐림, 눈, 비), 화자 연령대(〈예〉 유아(~13세),

청소년(14~19세), 청년(20~39세), 장년(40~59세), 노년(60세~)), 도메인(〈예〉 쇼핑, 날씨, 관광) 정의

- ③ 편향성 방지를 위해 단계 ②에서 도출한 범주별 적정 데이터 비율(%)을 고려하여 다양성 항목에 대한 구성비를 결정

2.1.3 다양성 항목 도출

- 다양성 항목(속성) 도출 시 ① 사업 목표(임무)별, ②어노테이션 대상별, ③데이터 유형별 항목을 참고해서 항목을 도출
- ① 사업 목표(임무)를 정의하고 사업 목표(임무)에 적합한 데이터 유형과 다양성 지표를 선택

임무	텍스트	이미지	비디오	오디오	3D	예시 항목	대상
객체/행위에 대한 분류/탐지	●	●	○	○	○	클래스별 인스턴스(객체/행위) 분포 또는 최소/평균 수 이미지당 최소/평균 클래스/인스턴스 수 ※ 이미지 내 여러 클래스/인스턴스를 분류/탐지하는 경우 인스턴스 크기(해상도)범위별 분포 또는 최소/평균 수 ※ 인스턴스 크기가 중요한 임무이거나 크기 구간별로 분류/탐지 성능 측정하는 경우	어노테이션
질한/위험의 수준 예측		●	○			수준별 인스턴스(객체/행위) 최소/평균 수	어노테이션
이상(질한/위험) 유무 예측		●	○			정상/이상 분포	어노테이션
자율주행	●	●	○		○	위치 정보(GPS 좌표 등), 시간(주/야간), 도로 종류(도심/고속 도로, 국도 등), 차량 속도(저/중/고속) 등	메타데이터
생물 성장(생육) 예측	●	●	○		○	환경 정보(공기 질/수질, 기온/수온 등)	메타데이터
비디오 요약			●			비디오 당 범주 수(평균)	어노테이션
번역/기계 독해	●			○		주제별, 언어별, 질의유형별 분포	어노테이션
음성인식/음성합성	●			●		화자 속성 분포, 도메인, 환경 관련 정보 분포	메타데이터

〈범례〉 ○ : 원시데이터(획득/수집한 데이터 포맷), ● : 원천데이터(학습 시 활용하는 데이터 포맷)

〈표 III-141〉 사업 목표(임무)별 다양성 항목 도출

- ② 데이터 유형별로 다양성 지표 결정 시 이미지의 경우는 데이터 수집 도구, 환경, 시점을 고려하고, 음성/텍스트 경우는 화자, 문장, 특성, 매체 등을 고려하여 적합한 다양성 지표를 선택하여 적용함

데이터 유형	예시 항목	대상
이미지	촬영 기기 종류(위성, 항공, 드론, CCTV, 블랙박스 등)	메타데이터
	촬영 환경(시간대, 날씨, 계절 등), 촬영 장소(행정/지리적 지역), (기타) 촬영 높이, 방향, 각도 등	
	촬영 시점(일인칭, 삼인칭)	
음성/텍스트	화자/저자 특성(연령대, 성별, 출신/거주지역 등) 분포	메타데이터
	문장길이, 핵심어 분포	
	주제, 도메인, NER(Named Entity Recognition), 질의 유형, 감성 및 감정 종류	
	Device, 매체 종류	

〈표 III-142〉 데이터 유형별 다양성 항목 도출

- ③ 어노테이션 대상별로 다양성 지표 결정 시에는 데이터 유형을 고려하여 적합한 다양성 지표를 선택

대상	텍스트	이미지	비디오	오디오	3D	예시 항목	대상
사람(이미지, 음성)	●	●	○	●	○	연령대, 성별, 출신 지역(인종) 분포 등	메타데이터
지형(항공/드론 촬영)		●	○		○	촬영 고도(저고도(fine), 고고도(Coarse)), 촬영기기 종류(위성, 항공, 드론) 분포 등	메타데이터

〈범례〉 ○ : 원시데이터(획득/수집한 데이터 포맷), ● : 원천데이터(학습 시 활용하는 데이터 포맷)

〈표 III-143〉 어노테이션 대상별 다양성 항목 도출

- 단, 다양성 항목 선정 시 다양성에 영향을 주지 않은 항목은 제외함

〈참고 : 다양성 항목 도출 시 제외 대상〉

- 값이 달라져도 데이터 다양성에 영향을 주지 않는 속성은 다양성 항목으로 도출하지 않는다.
〈예〉 촬영 각도(각도 값을 바꾸어도 객체의 특성에 별다른 차이가 없는 경우)
- 데이터 증식(augmentation) 기능의 파라미터에 해당하는 속성(속성의 변화에 따른 데이터를 자동으로 생성할 수 있으므로 데이터 수집 시 별도로 수집할 필요가 없음)
〈예〉 크기(해상도) 변경, 이동, 회전, 상하 반전, 명암 조절, 회색 화(gray scale) 등

〈표 III-144〉 다양성 항목 도출 제외 대상

2.1.4 다양성 항목에 대한 통계/요건 항목 분류

- ‘다양성(통계)’ 항목은 도출된 다양성 항목 중 인공지능 학습모델에 사용하기에 유의미한 수량인 통계적 충분성을 만족하는 항목으로 선택
- ‘다양성(요건)’ 항목은 도출된 다양성 항목 중 분류/탐지/인식 등의 카테고리별 인스턴스 수량의 적정 비율을 확보하는 통계적 균일성을 만족하는 항목으로 선택

※ 항목 선별 시 정보통신단체표준(국문표준) TTA-KO-10.1339 제정일: 2021년 12월 08일,
“지도학습을 위한 데이터 품질 관리 요구사항” 참조

〈항목 예시〉

- ’22년도 2~80번 “필라테스 동작 데이터” 과제임무는 필라테스 동작분류로 다양성 지표 항목은 아래와 같이 4개 항목이 도출되었으며,

구분	항목
다양성 항목	① 동작시연자 성별분포
	② 동작시연자 연령분포
	③ 동작 클래스
	④ 동작시연자 숙련도분포

- 도출된 4개 항목 중 통계적 충분성을 확보하는 ①, ② 항목이 ‘다양성(통계)’ 항목으로 선택함
- 도출된 4개 항목 중 통계적 균일성을 확보하는 ③, ④ 항목이 ‘다양성(요건)’ 항목으로 선택함

2.1.5 다양성 항목 지표 기준 및 검사 방법

- 다양성 항목에 대한 지표는 단일 값 또는 분포도 중첩률을 기준으로 목표를 설정하고 구축 시 자체적으로 지표 달성에 대한 검사를 통해 다양성을 만족하는지 확인
- 다양성 항목이 단일값인 경우 지표 기준 및 검사 방법

〈지표가 단일 값인 경우 지표 설정 및 검사 방법〉

- 예시 1) 이미지 당 평균 바운딩박스 수
 - 지표 목표치 설정 기준 : MS COCO, Open Images 등 잘 알려진 데이터세트의 통계를 참고하여 임무에 적합하게 설정
 - 검사 방법 : 전체 바운딩박스 수 / 전체 이미지 수
 - 목표 달성 기준 : 목표치 대비 결과치가 높을 경우
- 예시 2) XX 범주 최소 ~이상
 - 지표 목표치 설정 기준 : MS COCO, Open Images 등 잘 알려진 데이터세트의 통계를 참고하여 임무에 적합하게 설정
 - 검사 방법 : 경상도 출신 화자음성 최소 1만 건 이상
 - 목표 달성 기준 : 목표치 대비 결과치가 높을 경우

〈표 III-145〉 단일값 지표 설정 및 검사 방법

- 다양성 항목이 분포도 중첩률인 경우 지표 기준 및 검사방법

〈지표가 분포도 중첩률인 경우 지표 설정 및 검사 방법〉

- 지표 목표치 설정 기준
 - 지표 목표치 설정 기준 : 분포가 매우 균일할 필요는 없으며, 현실적인 수집 상황 등을 고려해서 설정
 〈예〉 '21년도 반려동물 질병 진단을 위한 영상 데이터(근골격계): 질환 및 정상 이미지의 비율을 크게 차이 나게 설정해도 타당한 근거가 있다면 가능함

〈타당한 근거사유 예시〉

해당 질환에 대해 당 사업 데이터 출처인 중대형 동물병원내원 비율이 낮아지고 유전적 소인 차단 치료 증가 및 실내 생활 비율 증가에 따라 해당 질병 발생을 저하되고 있음

- 검사 방법 : 도출한 속성의 목표 구성비 막대그래프와 결과 구성비 막대그래프 간의 길이 중첩 비율을 계산

$$\text{구성비 중첩률(\%)} = \frac{\sum_{k=1}^n \text{중첩막대길이}_k}{\sum_{k=1}^n \text{최대막대길이}_k} \times 100$$

- 목표 달성 기준 : 목표치 대비 결과치가 높을 경우

〈검사 예시〉

- 다양성 항목 : IoU 50% 이상
- 목표대비 결과 비율

구분	목표	결과	중첩률(IoU)
질환	40%	50%	81.8%
정상	60%	50%	

- 검사 결과 구성비 중첩률(%) = (40 + 50) / (50 + 60) * 100 = 81.8

〈표 III-146〉 중첩률 지표 설정 및 검사 방법

2.2 구문 정확성

2.2.1 개요

- 인공지능 학습용 데이터 구축사업에서 구문 정확성의 대상은 라벨링 데이터로, 라벨링 데이터의 구조, 입력값 범위 및 형식이 초기 설계 과정에서 정의한 데이터 속성에 따라 정확하게 입력되어 있어야 하며, 누락된 정보가 없도록 작성

〈예시 : 데이터 속성에 대한 라벨링 데이터(JSON 형식)〉

〈데이터 속성〉						〈라벨링 데이터〉	
구분	항목명	타입	필수 여부	설명	범위	비고	
Raw_Data_Info.	Raw_Data_ID	string	Y	원시 파일명	D1_220401_IS01_E0001_IM		<pre> { "Raw_Data_Info": { "Raw_Data_ID": "D1_220401_IS01_E0001_IM", "Copyrightier": "(주)○○○○", "Location": "IS01", "Date": "2022-04-01", "Length": 10, "Resolution": "1920, 1080", "FPS": 60, "F-stop": "F/1.8", "Exposure_Time": "1/1043", "Target_type": "Static", "Angle": "Eye", "File_Extension": "mp4" }, "Source_Data_Info": { "Source_Data_ID": "D1_220401_IS01_E0001_IM_L03_0001", "Time_start": "00:00:13:05", "Time_end": "00:00:15:02", "Occlusion_level": "L01", "File_Extension": "jpg" }, "Learning_Data_Info": { "Path": "//I-210501/", "File_Extension": "JSON", "Json_Data_ID": "D1_220401_IS01_E0001_IM_L03_0001", "Annotations": [{ "Class_ID": "Human", "Object_type": "Target_object", "Object_num": "1", "Type": "Bounding_Box", "Type_value": [1.0, 0.903, 574.634, 576.0], "Class_size": 2345, "Occlusion_size": 1232, "Occlusion_rate": "52%", }] } }</pre>
	Copyrightier	string	Y	소유자	(주)○○○○		
	Location	string	Y	촬영장소 ID	"IS01" ~ "OD01"		
	Date	string	Y	촬영일자(yyyy-mm-dd)	2022-04-01		
	Length	integer	Y	영상길이 초(sec)	10		
	Resolution	string	Y	해상도	1920, 1080		
	FPS	integer	Y	FPS	60		
	F-Stop	string	Y	조리개 수치	F1.8~"F11"		
	Exposure_Time	string	Y	노출 시간	1/1043~"1/40"		
	Target_type	string	Y	폐색유형	"Static"(일정형), "Dynamic"(유동형)		
	Angle	string	Y	촬영 각도	Low(로우앵글), "Eye"(아이앵글), "High"(하이앵글)		
	File_Extension	string	Y	원시 파일 확장자	mp4		
Source_Data_Info.	Source_Data_ID	string	Y	이미지 파일명	D1_220401_IS01_E0001_IM_L03_0001		
	Time_start	string	N	시작시간	00:00:13:05		
	Time_end	string	N	종료시간	00:00:15:02		
	Occlusion_level	string	N	폐색등급	"L01" ~ "L05"		
Learning_Data_Info.	File_Extension	string	Y	이미지 확장자	jpg		
	Path	string	Y	Json 폴더 경로	//I-210501/		
	File_Extension	string	Y	Json	JSON		
	Json_Data_ID	string	Y	Json 파일명	D1_220401_IS01_E0001_IM_L03_0001		
	Annotations	array	Y		[]		
	Class_ID	string	Y	객체 ID	Human		
	Object_Type	string	Y	객체유형	"Target_object", "Occlusion_object"		
	Object_num	string	Y	Object 번호	"1"~"10"		
	Type	string	Y	어노테이션 종류	"Segmentation", "Bounding_box"		
	Type_value	array	Y	[x.y, x.y, ...]	[1.0, 0.903, 574.634, 576.0, ...]		
	Class_size	integer	Y	객체크기(pixel)	2345		
	Occlusion_size	integer	Y	폐색크기(pixel)	1232		
	Occlusion_rate	string	Y	폐색비율(%)	"0%"~"95%"		

〈표 III-147〉 데이터 속성에 대한 라벨링 데이터 예시

2.2.2 지표 설정 기준

- 구문 정확성은 라벨링 데이터에 대한 무결성, 구조 정확성, 형식 정확성으로 구성되는데, 무결성은 라벨링 데이터가 읽기 실행에 이상이 없어야 되는 것은 필수적인 사항이므로 기준치 설정에서는 제외되며, 구조 정확성과 형식 정확성에 대한 지표 설정기준은 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

〈구문 정확성 지표 설정기준〉

- 구문 정확성의 지표 설정 대상은 구조 정확성과 형식 정확성임

구분	기준치(예시)	연관 품질지표
구조 정확성	99.5%	구문 정확성 - 데이터 구조
형식 정확성	99.5%	구문 정확성 - 입력값 범위, 데이터 형식

〈표 Ⅲ-148〉 구문 정확성 지표 설정기준

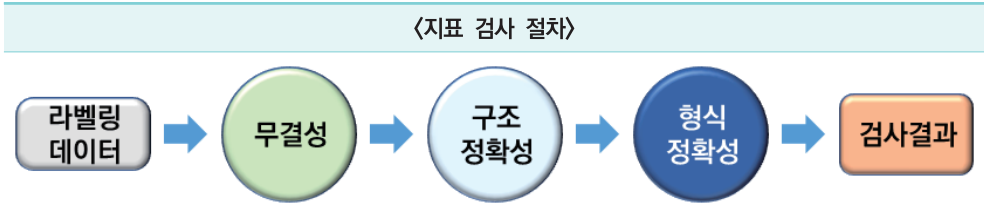
〈구문 정확성 오류 예시〉

○ JSON 스키마 기준 구문 정확성 오류 항목 예시

타입	설명
array	• array를 구성하는 아이템의 최소 개수에 미달
	• array를 구성하는 아이템의 최대 개수를 넘음
string	• 문자열 최소 길이보다 짧음
	• 문자열 최대 길이보다 길음
	• 문자열 허용 포맷외 포맷이 사용
	• format 항목 : date(날짜), regExp(정규표현식), email(이메일), ipv4(ip주소), ipv6(ip주소), url(URL)
	• date(날짜), regExp(정규표현식)에 한하여 expression 미작성
object	• 허용 가능 문자열 외 문자가 존재
	• 허용 가능하지 않은 객체 내 속성
number	• 숫자 최소 허용 값에 미달
	• 숫자 최대 허용 값을 초과
	• 숫자 허용 배수를 위배
	• 허용 가능 숫자를 위배
기타	• 항목의 값 null 허용 가능하지 않은 값인데 null 표기

2.2.3 지표 검사 방법

- 구문 정확성에 대한 검사 절차는 라벨링 데이터에 대한 ① 무결성 확인, ② 구조 정확성 검사, ③ 형식 정확성 검사의 단계로 수행



[그림 III-10] 지표 검사 절차

- ① 무결성 확인은 라벨링 데이터가 읽기 실행 시 이상이 없는 경우만 구조/형식에 대한 정확성 검사를 수행하고, 이상이 있는 경우는 검사를 중지하고 문제점을 확인 및 조치
- ② 구조 정확성은 라벨링 파일 내 항목에 대한 구조적 문제를 점검하는 것으로 필수항목 누락 여부(구문 규칙(스키마)에 정의된 필수항목(객체, 속성)의 누락 여부)와 유효 항목 여부(구문 규칙(스키마)에 정의되지 않은 항목(객체, 속성)이 라벨링 파일에 있는지 확인)를 검사
- ③ 형식 정확성은 라벨링 데이터의 항목별 형식을 점검하는 것으로 라벨링 데이터 내 규칙위반 여부(Value 없음, Value Type 에러, 특정 항목의 상세 검증 조건위반 등)를 검사
- 구조 정확성과 형식 정확성은 “정확도 = 100 - [(오류건수 / 전체 항목 수) × 100]”로 계산하고, 그 결과값을 목표치와 비교하여 달성 여부 확인

2.3 의미 정확성

2.3.1 개요

- 의미 정확성이란 데이터 라벨링(정답 어노테이션)의 품질을 측정하기 위한 지표로 예를 들어, 임무가 이미지 탐지인 경우 의미 정확성은 인스턴스를 둘러싸는 바운딩박스의 위치 정보와 분류 태그 작업이 정답 수준에 맞게 잘 되어 있는지를 판별하여 통계를 내는 것임
- 이러한 의미 정확성 지표를 적용 시 사업 목표(임무), 데이터 종류와 데이터 라벨링(또는 데이터 어노테이션) 방법에 따라 적합한 지표와 목표를 설정
- 또한, 정답 어노테이션의 품질을 확인하는 의미 정확성 지표는 정답 어노테이션이 기재된 학습용 데이터로 임무에 적합한 인공지능 모델을 훈련하고 평가해서 드러나는 성능을 확인하는 유효성 지표의 부분 집합이라고 볼 수 있음

- 단, 의미 정확성 지표 측정 시 기재한 분류 클래스 명(코드) 하나만 확인하면 되므로 임계치 변화에 기반하는 지표(AUC, AP 등)는 의미 정확성 지표로 선택하는 것은 지양
- 이러한 의미 정확성 지표 선택 시 『1.2.1 데이터 유형별 학습 Task』의 데이터 유형별 학습 Task를 참조하여 선택함

2.3.2 텍스트

- 인공지능 학습데이터 구축 목표 또는 임무 정의 시 적합한 데이터 종류가 텍스트인 경우 의미 정확성 지표를 선정하기 위해 구축 목표 또는 임무에 적합한 학습 Task를 결정

학습 Task	개념
텍스트 분류 (Text Classification)	<ul style="list-style-type: none"> 문장이나 문서를 적절한 범주로 지정하는 작업으로, 범주는 선택한 데이터 세트에 따라 다르며 주제 범위가 다양할 수 있음
텍스트 요약 (Text Summary)	<ul style="list-style-type: none"> 상대적으로 큰 원문을 핵심 내용만 간추려서 상대적으로 작은 요약문으로 변환하는 작업
기계 번역 (Machine Translation)	<ul style="list-style-type: none"> 인공지능을 이용해 인간이 사용하는 자연어를 다른 자연어로 번역하는 작업
순차적 레이블링 (Sequence Labeling)	<ul style="list-style-type: none"> 입력 시퀀스에 대하여 레이블 시퀀스를 각각 부여하는 작업으로, 태깅 작업이 대표적인 순차적 레이블링의 예시임
질의 응답 (Question Answering)	<ul style="list-style-type: none"> 질문(일반적으로 독해 질문)에 답변하는 작업으로서, 제공된 컨텍스트를 기반으로 답변할 수 없는 질문이 제시되면 기권함
음성 인식 (Speech Recognition)	<ul style="list-style-type: none"> 사람이 말하는 음성 언어를 인공지능이 해석해 그 내용을 문자 데이터로 전환하는 작업

〈표 III-149〉 텍스트 학습 Task(의미 정확성)

- 학습 Task가 결정되면 적합한 라벨링 유형과 품질측정항목(예시)을 참고하여 품질과 관련된 품질지표 및 정량목표 범위(예시)를 선택하여 활용

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
텍스트 분류 (Text Classification)	Class Labeling	정확도	100	<ul style="list-style-type: none"> KSIC 분류 관찰기록에 대한 역량평가 기술용어 문법적 정확성 분류 라벨 및 영역 정확도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
				<ul style="list-style-type: none"> • 분류 체계 정확도 • 비윤리 어휘 적합성 • 비윤리문장의 문형 매핑 • 비윤리성 유형 분류 정확도 • 에세이 난이도 • 에세이 타입 • 우울한 기분 진단 분류 • 유·불리 판단 • 의도분류(민원분야/술어) • 주제 분류 • 학생 학년 • 화행 분류
	Speech to Text	정확도	100	<ul style="list-style-type: none"> • 우울한 기분 진단 분류 정확도
	Tagging	정확도	100	<ul style="list-style-type: none"> • 개인정보 비식별화 및 민감 이슈 처리 • 개체명 • 신조어 라벨링
텍스트 요약 (Text Summary)	Class Labeling	정확도	100	<ul style="list-style-type: none"> • Scoring • 에세이 주제 • 요약문 분류 정확성 • 키워드 정확도
	Sequence Labeling	정확도	99.53	<ul style="list-style-type: none"> • 원문-요약문 주제(의미) 일치도
	Text Summary	정확도	99.4	<ul style="list-style-type: none"> • 영상에 대한 요약문 연관성 • 요약문 품질 정확도
기계 번역 (Machine Translation)	Machine Translation	정확도	100	<ul style="list-style-type: none"> • 번역 정확도 • 영-한 번역 정확도
		PTQA	88	<ul style="list-style-type: none"> • 영-한 번역 점수
순차적 레이블링 (Sequence Labeling)	Class Labeling	정확도	86.46	<ul style="list-style-type: none"> • “비윤리 문장”, “비도덕(T/F)분류”, “비도덕 강도” 간의 관계적합성
	Sequence Labeling	정확도	99.99	<ul style="list-style-type: none"> • 음성-스크립트(대본) 일치도
	Tagging	정확도	100	<ul style="list-style-type: none"> • 단락-문장 구조 정확도 • 판결문 주장/사실/판단 태깅

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
질의 응답 (Question Answering)	Class Labeling	리커트 척도	63.4	• 유사질의 다양성
	Sequence Labeling	정확도	99.67	• 답변 문장 품질 • 영상에 대한 질의응답 • 지문-질의-응답 간 내용 일치성
	Speech to Text	정확도	99.16	• 질의 문장 품질
음성 인식 (Speech Recognition)	Tagging	정확도	94.66	• 화자 라벨링 정확도

〈표 III-150〉 텍스트 학습 Task 분야별 품질측정항목(의미 정확성)

2.3.3 이미지

- 인공지능 학습데이터 구축 목표 또는 임무 정의 시 적합한 데이터 종류가 이미지인 경우 의미 정확성 지표를 선정하기 위해 구축 목표 또는 임무에 적합한 학습 Task를 결정

학습 Task	개념
객체 인식 (Object Detection)	• 이미지 내에서 특정 클래스의 객체 인스턴스를 인식하는 작업
키폰트 검출 (Keypoint Detection)	• 동작 인식을 위해 필요한 주요 키폰트를 검출하는 작업
얼굴 인식 (Face Recognition)	• 얼굴을 사용하여 개인의 신원을 식별하거나 확인하는 작업
이미지 분류 (Image Classification)	• 특정 이미지를 알고리즘에 입력했을 때, 그 이미지가 어떤 클래스에 속하는지 알려주는 작업

〈표 III-151〉 이미지 학습 Task(의미 정확성)

- 학습 Task가 결정되면 적합한 라벨링 유형과 품질측정항목(예시)을 참고하여 품질과 관련된 측정지표 및 목표치를 선택하여 활용

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
객체 인식 (Object Detection)	Bounding Box	$F\beta$ -점수	100	<ul style="list-style-type: none"> 바운딩박스 정확도 분류 라벨 정확도
	Keypoint	$F\beta$ -점수	99.56	<ul style="list-style-type: none"> 2D 키포인트 정확도
	Polygon	$F\beta$ -점수	99.95	<ul style="list-style-type: none"> 폴리곤 라벨 정확도 폴리곤 정확도
키포인트 검출 (Keypoint Detection)	Keypoint	$F\beta$ -점수	99.92	<ul style="list-style-type: none"> 분류 라벨 정확도(키포인트) 키포인트 정확도
얼굴 인식 (Face Recognition)	Bounding Box	정밀도	98.78	<ul style="list-style-type: none"> 바운딩박스 정확도
	Keypoint	정확도	100	<ul style="list-style-type: none"> 가상인물 기반 인물 스케치 설명문 기반 몽타주 스케치
	Polygon	정밀도	97.99	<ul style="list-style-type: none"> 폴리곤 정확도
이미지 분류 (Image Classification)	Bounding Box	정확도	99.99	<ul style="list-style-type: none"> 바운딩박스 정확도
		$F\beta$ -점수	99.66	<ul style="list-style-type: none"> 객체 라벨링 기상특보별 어노테이션 바운딩박스 정확도(객체, 행동)
	Keypoint	정확도	98.48	<ul style="list-style-type: none"> 키포인트 정확도
		$F\beta$ -점수	100	<ul style="list-style-type: none"> 분류 라벨 정확도(동작)
	Polygon	정확도	83.01	<ul style="list-style-type: none"> 폴리곤 정확도
		$F\beta$ -점수	99.5	<ul style="list-style-type: none"> 폴리곤 정확도
		정밀도	94.69	<ul style="list-style-type: none"> 폴리곤 라벨 정확도
	Polyline	$F\beta$ -점수	97.77	<ul style="list-style-type: none"> 폴리라인 정확도
	Tagging	정확도	99.78	<ul style="list-style-type: none"> 가상인물 기반 설명문 작성

〈표 III-152〉 이미지 학습 Task 분야별 품질측정항목(의미 정확성)

2.3.4 비디오

- 인공지능 학습데이터 구축 목표 또는 임무 정의 시 적합한 데이터 종류가 비디오인 경우 의미 정확성 지표를 선정하기 위해 구축 목표 또는 임무에 적합한 학습 Task를 결정

학습 Task	개념
객체 인식 (Object Detection)	<ul style="list-style-type: none"> 이미지 내에서 특정 클래스의 객체 인스턴스를 인식하는 작업
이미지 분류 (Image Classification)	<ul style="list-style-type: none"> 특정 이미지를 알고리즘에 입력했을 때, 그 이미지가 어떤 클래스에 속하는지 알려주는 작업
행동 인식 (Action Recognition)	<ul style="list-style-type: none"> 비디오 데이터 스트림에서 객체 행동의 감지 및 인식을 하는 작업
비디오 분류 (Video Classification)	<ul style="list-style-type: none"> 프레임이 주어진 비디오에서 특정 프레임 구간을 선택하고 레이블을 부여하는 작업
자세 추정 (Pose Estimation)	<ul style="list-style-type: none"> 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업
비디오 인식 (Video Recognition)	<ul style="list-style-type: none"> 비디오를 활용하여 실시간 객체 식별 및 비디오 프레임을 분류하는 작업
키폰트 검출 (Keypoint Detection)	<ul style="list-style-type: none"> 동작 인식을 위해 필요한 주요 키폰트를 검출하는 작업

〈표 III-153〉 비디오 학습 Task(의미 정확성)

- 학습 Task가 결정되면 적합한 라벨링 유형과 품질측정항목(예시)을 참고하여 품질과 관련된 측정지표 및 목표치를 선택하여 활용

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
객체 인식 (Object Detection)	Bounding Box	정확도	100	<ul style="list-style-type: none"> 바운딩박스 정확도 분류 라벨 및 영역 정확도 얼굴 영역 바운딩 박스
		F β -점수	100	<ul style="list-style-type: none"> 라벨링 정확도 바운딩박스 라벨 정확도 바운딩박스 정확도 분류 라벨 정확도
		정밀도	100	<ul style="list-style-type: none"> 2D 바운딩박스 라벨 정확도 LiDAR

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
				<ul style="list-style-type: none"> • (얼굴, 객체) 바운딩박스 라벨 정확도 • 바운딩박스 정확도 • 분류 라벨 정확도 <기타> <ul style="list-style-type: none"> • 열화상, 엽록소형광, 다각도, 탐부 RGB, 통과차량 라벨 정확도
		재현율	100	<ul style="list-style-type: none"> • 2D 바운딩박스 라벨 정확도 • LiDAR • 다각도 RGB • (얼굴, 객체 등) 바운딩박스 라벨 정확도 • 바운딩박스 정확도 • 분류 라벨 정확도 <기타> <ul style="list-style-type: none"> • 열화상, 엽록소형광, 탐부 RGB, 통과차량 라벨 정확도
	Cuboid	F β -점수	94.86	<ul style="list-style-type: none"> • 포인트 클라우드 정확도
	Keypoint	정확도	96.85	<ul style="list-style-type: none"> • 얼굴마커 프로젝션 정확도
		정밀도	99.74	<ul style="list-style-type: none"> • 도트(dot) 라벨 정확도
	Polygon	정확도	100	<ul style="list-style-type: none"> • 분류 영역 정확도 • 분할 라벨 정확도 <기타> <ul style="list-style-type: none"> • 경동맥 세그멘테이션 정확도 • 뼈 세그멘테이션 정확도 • 척추체 세그멘테이션 정확도 • 퇴행성 경추 협착증/퇴행성 요추 디스크 라벨링 정확도 • 흉추 척수증 라벨링 정확도
		F β -점수	99.85	<ul style="list-style-type: none"> • 폴리곤 라벨 정확도 • 폴리곤 정확도
		MOTA	98.75	<ul style="list-style-type: none"> • 폴리라인 라벨 정확도
		정밀도	100	<ul style="list-style-type: none"> • 바운딩박스 라벨 정확도 • 폴리곤 라벨 정확도 • 폴리곤 정확도 • 세그멘테이션 정확도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
	Polyline	재현율	100	<ul style="list-style-type: none"> 바운딩박스 라벨 정확도 폴리곤 라벨 정확도 폴리곤 정확도 세그멘테이션 정확도
		정확도	100	<ul style="list-style-type: none"> 분류 라벨 정확도 〈기타〉 경동맥 세그멘테이션 정확도 골다공증성 골절 통증 라벨링 정확도 만성신경통증 라벨링 정확도 뼈 라인 정확도 퇴행성 경추 협착증/퇴행성 슬관절 라벨링/퇴행성 요추 디스크/퇴행성 요추 협착증 라벨링 정확도 회전근개 파열 라벨링 정확도
		정밀도	97.21	<ul style="list-style-type: none"> 폴리라인 라벨 정확도
		재현율	95.89	<ul style="list-style-type: none"> 폴리라인 라벨 정확도
	Bounding Box	정확도	100	<ul style="list-style-type: none"> 분류 라벨 정확도 세그멘테이션 정확도 이미지 바운딩박스 정확도
이미지 분류 (Image Classification)	Class Labeling	$F\beta$ -점수	99.05	<ul style="list-style-type: none"> 분류 라벨 정확도
		정밀도	100	<ul style="list-style-type: none"> 분류 라벨 정확도
		정확도	99.6	<ul style="list-style-type: none"> 한국어 문장의 재난 정보 분류 정확도
	Keypoint	$F\beta$ -점수	96.1	<ul style="list-style-type: none"> 키포인트 정확도
	Polygon	$F\beta$ -점수	99.2	<ul style="list-style-type: none"> 폴리곤 정확도
		정확도	99.91	<ul style="list-style-type: none"> OCR 라벨 정확도
		정밀도	100	<ul style="list-style-type: none"> 폴리곤 라벨 정확도
		재현율	100	<ul style="list-style-type: none"> OCR 라벨 정확도 폴리곤 라벨 정확도
행동 인식 (Action Recognition)	Bounding Box	정확도	92	<ul style="list-style-type: none"> 구간 라벨 정확도
		정밀도	99.24	<ul style="list-style-type: none"> 보행자 라벨 정확도
		재현율	97.99	<ul style="list-style-type: none"> 보행자 라벨 정확도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
	Keypoint	정확도	95.47	<ul style="list-style-type: none"> 수어 영상 전사 정확도 (형태소/비수지 가공)
	Sequence Labeling	정확도	97.21	<ul style="list-style-type: none"> 주제 제스처 간의 의미 및 행동 일치성
	Tagging	F β -점수	98.59	<ul style="list-style-type: none"> 태깅 정확도
비디오 분류 (Video Classification)	Bounding Box	정확도	100	<ul style="list-style-type: none"> 부위별 구조 어노테이션 정확도 분류 영역 정확도 분류 라벨 정확도 <기타> BMD class 분류 라벨링 정확도 소장 병변 분류 라벨 정확도 수면 영상과 라벨 데이터의 시간 동기화 수술단계 분류 라벨 정확도 폐/병변 어노테이션 정확도 해부학적 구조물 및 병변 세그멘테이션 정확도
		정밀도	97.49	<ul style="list-style-type: none"> 바운딩박스 라벨 정확도
		재현율	96.85	<ul style="list-style-type: none"> 바운딩박스 라벨 정확도
	Class Labeling	정확도	99.69	<ul style="list-style-type: none"> 분류 라벨 정확도
	Keypoint	정확도	99.13	<ul style="list-style-type: none"> 어노테이션 정확도
	Polygon	정확도	99.6	<ul style="list-style-type: none"> 분류 라벨 및 분할 정확도 분류 라벨 및 영역 정확도 부위별 구조 분류 라벨 정확도 질환 분류 라벨 정확도 뇌전이암 세그멘테이션 정확도
		정밀도	98.55	<ul style="list-style-type: none"> 폴리곤 라벨 정확도
		재현율	98.24	<ul style="list-style-type: none"> 폴리곤 라벨 정확도
	Polyline	정밀도	94.98	<ul style="list-style-type: none"> 폴리라인 라벨 정확도
		재현율	97.36	<ul style="list-style-type: none"> 폴리라인 라벨 정확도
	Tagging	동일응답 비율	79.67	<ul style="list-style-type: none"> 하이라이트 구간 라벨링 결과 간 유사도
		평균 코사인 유사도	82.28	<ul style="list-style-type: none"> 하이라이트 구간 태깅 결과 간 코사인 유사도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
자세 추정 (Pose Estimation)	Keypoint	정확도	96.88	• 분류 라벨 정확도
비디오 인식 (Video Recognition)	Bounding Box	정합도	95.13	• 실화상-열화상 동기화 정합도
	Sequence Labeling	정확도	100	• 멀티모달 데이터 시점 일치
키폰트 검출 (Keypoint Detection)	Keypoint	정확도	100	• 분류 라벨 정확도 • 스켈레톤 투사 정확도 • 키폰트 정확도
		AP@oks	98.68	• 키폰트 정확도
		F β -점수	99.3	• 키폰트 정확도
		OKS	99.96	• 키폰트 정확도
		라벨 분류 정확도	98.89	• 키폰트 속성 분류 정확도
		정밀도	97.6	• 키폰트 정확도 • 키폰트 라벨 정확도
		재현율	100	• 키폰트 정확도 • 키폰트 라벨 정확도

〈표 III-154〉 비디오 학습 Task 분야별 품질측정항목(의미 정확성)

2.3.4 오디오

- 인공지능 학습데이터 구축 목표 또는 임무 정의 시 적합한 데이터 종류가 비디오인 경우 의미 정확성 지표를 선정하기 위해 구축 목표 또는 임무에 적합한 학습 Task를 결정

학습 Task	개념
오디오 분류 (Audio Classification)	• 오디오 정보를 특정 시간 단위로 분할(오디오 클립)하여, 신호의 특이점을 분석하고 특정 클래스로 분류하는 작업
음성 인식 (Speech Recognition)	• 사람이 말하는 음성 언어를 인공지능이 해석해 그 내용을 문자 데이터로 전환하는 작업
텍스트 분류 (Text Classification)	• 문장이나 문서를 적절한 범주로 지정하는 작업으로, 범주는 선택한 데이터 세트에 따라 다르며 주제 범위가 다양할 수 있음
음성 합성 (Speech Synthesis)	• 컴퓨터 또는 기타 장치를 사용하여 사람의 음성을 인공적으로 시뮬레이션 하는 작업

〈표 III-155〉 오디오 학습 Task(의미 정확성)

- 학습 Task가 결정되면 적합한 라벨링 유형과 품질측정항목(예시)을 참고하여 품질과 관련된 측정지표 및 목표치를 선택하여 활용

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
오디오 분류 (Audio Classification)	Class Labeling	정확도	99.99	<ul style="list-style-type: none"> • 구문 정확도 • 분류 라벨 정확도 • 소리 분류 정확도 • 소음클래스 일치도 • 전사 정확도 • 주제 분류 정확도 • 주제별 분포 정확도 • 카테고리 분류 정확도 • 화자 규모 정확도
	Sequence Labeling	정확도	99.97	<ul style="list-style-type: none"> • 노트 가사 / 밴딩 / 바이브레이션 / 호흡 일치성 • 노트 시작 시간과 지속 시간 일치성 • 발화 시작 / 끝 위치 정확도
	Tagging	정확도	100	<ul style="list-style-type: none"> • 발화 시작 / 끝 시간 라벨링 • 발화 시작 시간과 끝 시간 정확도 • 발화시간/종료시간 일치성 • 태깅 정보 정확도
음성 인식 (Speech Recognition)	Class Labeling	정확도	99.31	<ul style="list-style-type: none"> • 감정표현 정확도 • 음성품질 정확도
	Machine Translation	정확도	71.29	<ul style="list-style-type: none"> • 감정 태깅 정보 정확도
	Sequence Labeling	정확도	99.99	<ul style="list-style-type: none"> • 노트 가사/밴딩/바이브레이션/호흡 일치성 • 노트 시작 시간과 지속 시간 일치성 • 노트 피치 일치성
	Speech to Text	정확도	99.99	<ul style="list-style-type: none"> • 전사 정확도 • 전사규칙 부합도 • 전사품질 정확도 • 발화자 의도 정확성
		리커트 척도	82.2	<ul style="list-style-type: none"> • 번역 정확도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
	Tagging	정확도	100	<ul style="list-style-type: none"> 발화시작 시간, 발화 종료 시간 라벨 정확도 음성구간 어노테이션 정확도 전사 정확도
텍스트 분류 (Text Classification)	Class Labeling	정확도	99.23	<ul style="list-style-type: none"> 문장 품질 상담유형/주제분류 정확도 카테고리 분류 정확도
	Sequence Labeling	정확도	100	<ul style="list-style-type: none"> 발화스크립트 중복 확인
음성 합성 (Speech Synthesis)	Class Labeling	정확도	96.98	<ul style="list-style-type: none"> 감성 상태 분류 정확도 감정 상태 분류 정확도 발화스타일 분류 정확도 캐릭터 감정 분류 정확도
	Sequence Labeling	정확도	99.2	<ul style="list-style-type: none"> 발화 시작 시간, 발화 종료 시간 라벨 정확도
	Speech to Text	정확도	99.99	<ul style="list-style-type: none"> 전사 정확도
	Tagging	정확도	94.53	<ul style="list-style-type: none"> 발화 주제와 발화 음성 간의 적합성

〈표 III-156〉 오디오 학습 Task 분야별 품질측정항목(의미 정확성)

2.3.5 3D

- 인공지능 학습데이터 구축 목표 또는 임무 정의 시 적합한 데이터 종류가 비디오인 경우 의미 정확성 지표를 선정하기 위해 구축 목표 또는 임무에 적합한 학습 Task를 결정

학습 Task	개념
3D 객체 인식 (3D Object Detection)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 특정 클래스의 객체 인스턴스를 감지하는 작업
3D 포인트클라우드 분류 (3D Point Cloud Classification)	<ul style="list-style-type: none"> 포인트 클라우드를 알고리즘에 입력했을 때, 어떤 클래스에 속하는지 알려 주는 작업
3D 자세 추정 (3D Pose Estimation)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 객체와 객체의 주요 부분에 대한 위치, 방향을 식별하여 클래스를 분류하는 작업

학습 Task	개념
3D 객체 복원 (3D Object Reconstruction)	<ul style="list-style-type: none"> 영상/이미지/포인트 클라우드로 부터 객체(물체, 사람 등)의 3차원 구조를 복원(reconstruction)하는 작업
2D 이미지를 통한 시점 합성(View Synthetic) 모델	<ul style="list-style-type: none"> 컴퓨터 또는 기타 장치를 사용하여 이미지를 인공적으로 3D로 전환하는 작업으로 서로 다른 시점을 가진 여러개의 카메라들을 이용하여 객체를 촬영한 뒤, 촬영된 영상들과 각 카메라들의 촬영 정보들을 이용하여 해당 객체에 대한 3차원 정보를 추정
3D 키포인트 검출 (3D Keypoint Detection)	<ul style="list-style-type: none"> 3D 영상/이미지 내에서 동작 인식을 위해 필요한 주요 키포인트를 검출하는 작업

〈표 III-157〉 3D 학습 Task(의미 정확성)

- 학습 Task가 결정되면 적합한 라벨링 유형과 품질측정항목(예시)을 참고하여 품질과 관련된 측정지표 및 목표치를 선택하여 활용

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
3D 객체 인식 (3D Object Detection)	Bounding Box	F β -점수	99.4	<ul style="list-style-type: none"> 바운딩박스 정확도 소방시설 객체 바운딩박스 정확도
		정밀도	94.74	<ul style="list-style-type: none"> 바운딩박스 라벨 정확도
		재현율	98.08	<ul style="list-style-type: none"> 바운딩박스 라벨 정확도
	Cuboid	F β -점수	99.01	<ul style="list-style-type: none"> 큐보이드 정확도
		MOTA	96.13	<ul style="list-style-type: none"> 큐보이드 라벨 정확도 (추적 대상)
		정밀도	98.79	<ul style="list-style-type: none"> 2.5D 바운딩박스 라벨 정확도 큐보이드 라벨 정확도 (LiDAR, PCD)
		재현율	100	<ul style="list-style-type: none"> 2.5D 바운딩박스 라벨 정확도 큐보이드 라벨 정확도 (LiDAR, PCD)
	Polygon	F β -점수	99.6	<ul style="list-style-type: none"> 폴리곤 정확도

Task 분야	라벨 유형	측정지표	정량목표 (예시)	품질 검사항목(예시)
3D 포인트클라우드 분류 (3D Point Cloud Classification)	Cuboid	정밀도	97.33	• 폴리곤 라벨 정확도
		재현율	98.04	• 폴리곤 라벨 정확도
		정밀도	96.92	• 큐보이드 라벨 정확도
		재현율	96.92	• 큐보이드 라벨 정확도
3D 자세 추정 (3D Human Pose Estimation)	Cuboid	F β -점수	94.04	• 큐보이드(3D 바운딩박스) 정확도
3D 키포인트 검출 (3D Keypoint Detection)	Keypoint	F β -점수	99.32	• 3D 키포인트 정확도

※ 단, '21년도 인공지능 학습용 데이터 구축 사업 중 3D 객체 복원(3D Object Reconstruction), 2D 이미지를 통한 시점 합성(View Synthetic) 모델에 해당되는 과제가 없음

〈표 III-158〉 3D 학습 Task 분야별 품질측정항목(의미 정확성)

2.4 유효성

2.4.1 개요

- 유효성은 데이터 라벨(정답 어노테이션)이 기재된 학습용 데이터로 임무에 적합한 인공지능 모델을 훈련하고 평가해서 드러나는 성능을 확인하는 것으로 예를 들어, 인공지능 모델 예측한 인스턴스 바운딩박스 위치 정보와 분류 태그를 정답과 비교해서 성능을 측정
- 따라서, 인공지능 모델은 분류 태그로 어느 한 클래스 명(코드)만 출력하지 않고 클래스별로 확률(또는 신뢰수준)을 출력하므로 유효성 지표에는 신뢰수준 임계치(Threshold)를 변수로 두어 산정하는 지표(AUC, AP 등)가 존재할 수 있음

2.4.2 유효성 목표 설정 원칙

- 구축사업 추진 시 유효성에 대한 목표 설정 원칙을 준수해야 함

〈유효성 목표 설정 원칙〉

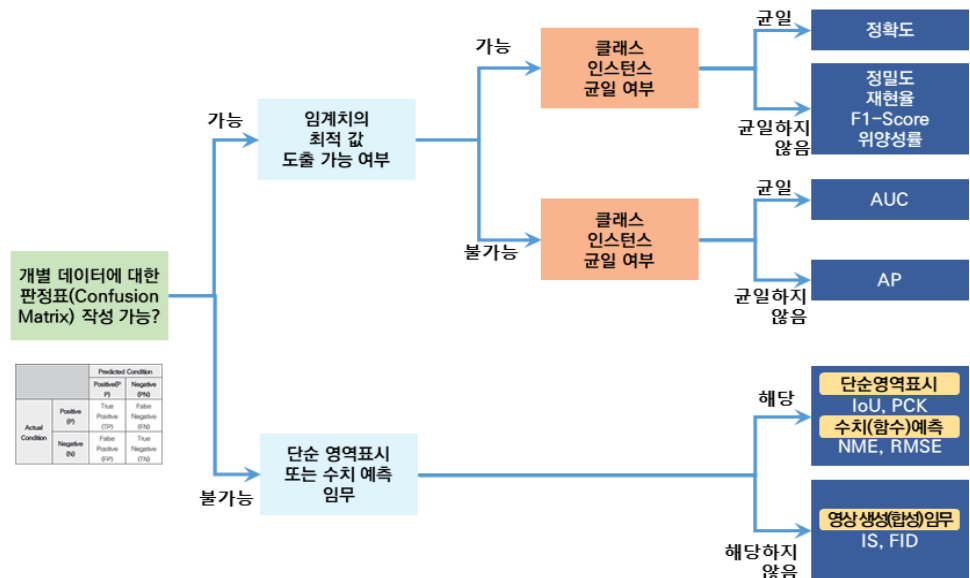
- 최신의 논문, 기고서, 공식 블로그 등에서 공식 보고된 값을 참고하여 유효성 목표치 기준 수립
 - 해당 구축사업과 가장 유사한 주제의 SoTA(State of The Art) 자료(논문, 기고서 등)에서 공식 보고되는 값을 기준으로 유효성 목표치를 설정하는 것이 원칙
 - ※ 단, 지적재산권에 위배됨이 없는 SoTA 모델 선정
 - 해당 구축사업의 이해관계자 간에 목표치에 대해 논의할 때, 반드시 상기 자료를 근거로 하고 출처를 명확히 제시
 - 다만, 선행 연구 분야에 해당하는 임무이기 때문에 SoTA에 해당 자료를 참고하기 힘든 경우 이해관계자 및 해당 임무·도메인 관련 전문가 의견을 반영하여 설정

〈표 III-159〉 유효성 목표 설정 원칙

2.4.3 유효성 지표 선택 방법

- 구축사업에서는 인공지능 학습데이터 구축 목표 또는 임무로 적합한 유효성 지표를 선택하기 어렵다. 따라서 ① 데이터 특성 등에 따라 지표를 선택하는 방법은 다음을 참고

〈데이터 특성에 따른 지표 선택 절차〉



〈표 III-160〉 데이터 특성에 따른 유효성 지표 선택

- ‘데이터 특성에 따른 지표 선택 절차’ 중 ② 판정표(Confusion Matrix) 작성이 가능한 경우에 선택할 수 있는 대표적인 유효성 지표는 다음과 같음

유효성 지표 선택 기준	유효성 지표	수식
<ul style="list-style-type: none"> 클래스별 인스턴스 수의 비율이 균일한 데이터셋을 다룰 때 	정확도 (Accuracy)	$\frac{TP + TN}{TP + FP + TN + FN}$
<ul style="list-style-type: none"> FP가 매우 중요한 경우 선택 〈예시〉 스팸메일 여부 판단 모델: 악성 코드는 없는 단순한 스팸메일이라고 가정했을 때, FP는 정상 메일을 스팸으로 분류해서 업무 차질을 초래하므로 반드시 포함 	정밀도 (Precision)	$\frac{TP}{TP + FP}$
<ul style="list-style-type: none"> FN이 매우 중요한 경우 선택 〈예시〉 코로나 여부 판단 모델: 진단키트가 아닌 순수 AI 모델이 있다고 가정했을 때, FN은 코로나 환자에게 코로나가 아니라고 분류해 방역 문제가 발생하므로 반드시 포함 	재현율 (Recall),	$\frac{TP}{TP + FN}$
<ul style="list-style-type: none"> FP와 TN이 매우 중요한 경우 선택 	위양성율	$\frac{FP}{FP + TN}$
<ul style="list-style-type: none"> 정밀도와 재현율 중 하나를 선택하기 어려운 경우 (두 값의 가중치 조화평균으로, 단순 산술 평균과 달리 어느 하나가 낮은 값을 갖는 것을 비교적 잘 반영함) 	Fβ-점수	$(1+\beta^{-2}) \times \frac{\text{정밀도} \times \text{재현율}}{(\beta^{-2} \times \text{정밀도}) + \text{재현율}}$
<ul style="list-style-type: none"> 최적의 임계치를 선택하기 어렵고 균일한 데이터셋의 경우 F1-Score의 대안으로 선택(F1-Score는 특정 판단 기준이 되는 임계치를 설정해 두고 결과 판별) 단, 양성과 음성 클래스의 비율이 균형을 이룰 경우 선택 ※ 탐지에서는 사용하지 않음(탐지에서는 TN이 존재하지 않음) 	AUC	ROC 커브 면적
<ul style="list-style-type: none"> 최적의 임계치를 선택하기 어렵고 불균일한 데이터셋의 경우 AUC의 대안으로 선택 양성과 음성 클래스의 비율의 불균형이 클 경우 선택 〈예〉 찾고자 하는 양성 클래스의 비율이 매우 작을 경우(예: fraud detection 문제) AUC 대신 선택 	AP	PR 커브 면적

〈표 III-161〉 판정표 작성이 가능한 경우 유효성 지표 선택 기준

- 또한, ③ 인공지능 학습데이터 구축 목표 또는 임무별 주요 유효성 지표를 참고하여 적합한 지표를 선택할 수 있음

임무	세부 임무	주요 성능 지표
분류	이미지/객체/감정 분류	<ul style="list-style-type: none"> 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), $F\beta$-점수, AUC
탐지, 인식	객체 탐지/추적/인식, 공간 탐지, 화자 탐지, 장소/얼굴/표정 인식	<ul style="list-style-type: none"> IoU(Intersection over Union) <ul style="list-style-type: none"> 임무에 따라 최종 지표이거나 개별 데이터 참/거짓 판정의 수단 정밀도(Precision), 재현율(Recall), $F\beta$-점수 AP(Average Precision) <ul style="list-style-type: none"> PR(Precision-Recall) Curve의 면적
(언어적) 이해	키워드 추출	<ul style="list-style-type: none"> 정밀도(Precision), 재현율(Recall), $F\beta$-점수 mAP(mean Average Precision), nDCG(normalized Discounted Cumulative Gain)
	이미지/비디오/음성/상황이해 등	<ul style="list-style-type: none"> $F\beta$-점수, AP
	기계 독해, QA	<ul style="list-style-type: none"> $F\beta$-점수, EM(Exact Matching), 정확도(Accuracy)
	번역	<ul style="list-style-type: none"> BLEU(Bilingual Evaluation Understudy)
요약	OCR, 음성인식	<ul style="list-style-type: none"> WER(Word Error Rate) <ul style="list-style-type: none"> 단어 오류율 측정(라틴어 계열에 많이 사용) CER(Character Error Rate) <ul style="list-style-type: none"> 글자 오류율 측정(한글과 같이 조사가 있는 언어에 주로 사용) $F\beta$-점수, AP
	텍스트 요약, 비디오 요약, 대화 요약 등	<ul style="list-style-type: none"> 전문가 평가 ROUGE(Recall Oriented Understudy of Gisting Evaluation)
합성 (생성)	이미지/비디오/음성 합성(생성), 텍스트 합성(생성) 등	<ul style="list-style-type: none"> 설문 조사 Inception Score(IS) Pitch Accuracy(PA) Pitch Entropy(PE) Fréchet Inception Distance(FID)
예측	이미지 예측, 상황 예측	<ul style="list-style-type: none"> $F\beta$-점수
행동/행위 분석		<ul style="list-style-type: none"> PCK(Percentage of Correct Keypoint) <ul style="list-style-type: none"> 최종 지표로 사용 가능 OKS(Object Keypoint Similarity), PDJ(Percentage of Detected Joints), MPJPE(Mean Per Joint Position Error) <ul style="list-style-type: none"> 주로 개별 데이터 참/거짓 판정의 수단

〈표 III-162〉 임무별 유효성 지표 선택 기준

2.4.3.1 텍스트

- ① 데이터 특성, ② 판정표(Confusion Matrix) 작성이 가능한 경우, ③ 인공지능 학습데이터 구축 목표 또는 임무와 2.3 의미 정확성의 텍스트 데이터에 해당하는 학습 Task를 참조하여 적합한 유효성 측정지표를 선택할 수 있음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
텍스트 분류 (Text Classification)	AUROC	99	• VGG
	MAE	59	• KorBERT
	MSE	71	• KorBERT
	PCC	82.28	• BERT, LSTM
	QWK	81.16	• BERT, LSTM
	정밀도	86.1	• BERT
	정확도	98.68	• ELECTRA
텍스트 요약 (Text Summary)	$F\beta$ -점수	90.8	• BERT • ELECTRA-base, KoBERT • KoBERT • KorBERT • NBC + FastText • VGG
	ROUGE-L	37.55	• T5 • Transformer
	Top-K 정확도	95	• BERT • BERT
기계 번역 (Machine Translation)	$F\beta$ -점수	89.85	• ELECTRA-base, KoBERT • KorBERT, BERT
	BLEU	77.22	• Transformer
순차적 레이블링 (Sequence Labeling)	$F\beta$ -점수	91.3	• ELECTRA-base, KoBERT
질의 응답 (Question Answering)	$F\beta$ -점수	81.17	• ELECTRA-base, KoBERT
	정확도	81.17	• Hierarchical Conditional Relation Networks for Video QA
질의 응답 (Question Answering)	$F\beta$ -점수	85.88	• ELECTRA
	정확도	77.78	• Hierarchical Conditional Relation Networks for Video QA
음성 인식 (Speech Recognition)	$F\beta$ -점수	76.05	• SVM

〈표 III-163〉 텍스트 Task 분야별 유효성 측정지표

2.4.3.2 이미지

- ① 데이터 특성, ② 판정표(Confusion Matrix) 작성이 가능한 경우, ③ 인공지능 학습데이터 구축 목표 또는 임무와 2.3 의미 정확성의 이미지 데이터에 해당하는 학습 Task를 참조하여 적합한 유효성 측정지표를 선택할 수 있음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
객체 인식 (Object Detection)	AP	38.78	<ul style="list-style-type: none"> EfficientNet B0 Scaled-YOLOv4
	AR	72.24	<ul style="list-style-type: none"> Scaled-YOLOv4
	BLEU	81.92	<ul style="list-style-type: none"> ConvMath_VIT
	Correlation	71	<ul style="list-style-type: none"> U-net
	EAO	73.24	<ul style="list-style-type: none"> -
	FID	10.44	<ul style="list-style-type: none"> StyleGAN2 ADA
	FPS	20.93	<ul style="list-style-type: none"> YOLOv5
	IoU	79.98	<ul style="list-style-type: none"> -
	mAP	96.87	<ul style="list-style-type: none"> EfficientDet Fast R-CNN Mask R-CNN RNN, DNN SIMPLE-HRNET YOLACT YOLO YOLOv4 YOLOv4, KNN algorithm YOLOv5
	mAP@IoU	93.6	<ul style="list-style-type: none"> Mask R-CNN YOLOv4 YOLOv5
	mIoU	74.74	<ul style="list-style-type: none"> Swin Transformer
	Top-K 정확도	100	<ul style="list-style-type: none"> EfficientNet B0 SlowFast ++
	재현율	80.48	<ul style="list-style-type: none"> Inception-v4 U-net
	정확도	90.81	<ul style="list-style-type: none"> Inception-v4 ResNet U-net

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
	$F\beta$ -점수	96	<ul style="list-style-type: none"> ResNet U-net YOLO YOLOv4
키포인트 검출 (Keypoint Detection)	mAP	99.29	<ul style="list-style-type: none"> AlphaPose YOLOv5
	정확도	94.48	<ul style="list-style-type: none"> Modified VGGNet
얼굴 인식 (Face Recognition)	AP	98.76	<ul style="list-style-type: none"> RetinaFace
	FID	7.94	<ul style="list-style-type: none"> StyleGAN2
	Kinship verification 정확도	70.66	<ul style="list-style-type: none"> ArcFace
	MAE	3.11	<ul style="list-style-type: none"> CNN network for feature extraction
	Rank-K 정확도	81.9	<ul style="list-style-type: none"> ArcFace
	정확도	99.31	<ul style="list-style-type: none"> ArcFace CNN network for feature extraction
이미지 분류 (Image Classification)	AUROC	98.24	<ul style="list-style-type: none"> FCDD
	Grid Cell 정확도	97.47	<ul style="list-style-type: none"> Mask R-CNN
	mAP	99.1	<ul style="list-style-type: none"> Mask R-CNN YOLOv5
	mIoU	85.93	<ul style="list-style-type: none"> Swin Transformer U-net
	Overall Pixel 정확도	95.95	<ul style="list-style-type: none"> U-net
	Top-K 정확도	95.3	<ul style="list-style-type: none"> ResNet mobileNet
	정확도	91.73	<ul style="list-style-type: none"> U-net
	$F\beta$ -점수	99.6	<ul style="list-style-type: none"> CRAFT Customized Mask R CNN HRCenterNet Mask R-CNN VGG Yolo V5

〈표 III-164〉 이미지 Task 분야별 유효성 측정지표

2.4.3.3 비디오

- ① 데이터 특성, ② 판정표(Confusion Matrix) 작성이 가능한 경우, ③ 인공지능 학습데이터 구축 목표 또는 임무와 2.3 의미 정확성의 비디오 데이터에 해당하는 학습 Task를 참조하여 적합한 유효성 측정지표를 선택할 수 있음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
객체 인식 (Object Detection)	AUROC	95.38	<ul style="list-style-type: none"> CNN based Segmentation model RNN, Transformer WGAN XGBoost, LightGBM, CatBoost, NeuralnetMXnet, NetFastAI
	DSC	94.14	<ul style="list-style-type: none"> RetinaNet, U-Net U-Net, FCN
	MAE	99.6	<ul style="list-style-type: none"> Random Forest
	mAP	98.8	<ul style="list-style-type: none"> DarkNet53 DetectoRS EfficientDet EfficientDet D3 General (A1) Mask R-CNN Scaled-YOLOv4 Swin-Transformer UNET YOLACT YOLOv5 YOLO, EfficientDet YOLOv3 YOLOv3-tiny, YOLOv4 YOLOv5
	mAP@IoU	99.835	<ul style="list-style-type: none"> darknet YOLOv4 DetectoRS Faster R-CNN Mask R-CNN RetinaNet segformer YOLOv4 YOLOv5

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
			<ul style="list-style-type: none"> DeepLab v3+
	mIoU	98.9	<ul style="list-style-type: none"> 3D CNN HRNet Mask R-CNN ResNet UNET
	NRMSE	92.5	<ul style="list-style-type: none"> Polynomial Regression
	RE	95.7	<ul style="list-style-type: none"> Deep Encoder-based Predictor
	RMSE	0.14	<ul style="list-style-type: none"> CNN
	Sensitivity	92	<ul style="list-style-type: none"> RetinaNet, U-Net WGAN
	Specificity	99	<ul style="list-style-type: none"> WGAN
	정확도	91.53	<ul style="list-style-type: none"> ResNet 50 VGG, ResNet, EfficientNet
	$F\beta$ -점수	99.29	<ul style="list-style-type: none"> Resnet50 TransFG U-net VTN XGBoost XGBoost, LightGBM, CatBoost, NeuralnetMXnet, NetFastAI
이미지 분류 (Image Classification)	AUROC	100	<ul style="list-style-type: none"> Densenet-161 DenseNet201 EfficientNet-B0 기반 CNN 모델 Multi-Resolution Input CNN
	Correlation	83	<ul style="list-style-type: none"> U-net
	DSC	99.47	<ul style="list-style-type: none"> CNN based Segmentation model EfficientNet B0모델을 encoder로 하는 FPN U-net
	IoU	89	<ul style="list-style-type: none"> U-net
	mAP	90.1	<ul style="list-style-type: none"> YOLOv5
	MRE	1.737	<ul style="list-style-type: none"> ResNet-50
	RMSE	99.944	<ul style="list-style-type: none"> LSTM

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
	Sensitivity	90.32	<ul style="list-style-type: none"> RetinaNet, YOLO YOLO
	SMAPE	85.91	<ul style="list-style-type: none"> XGBoost
	Top-K 정확도	100	<ul style="list-style-type: none"> EfficientNet-B0 Resnet152 Inception V3+Triplet Loss
	정확도	100	<ul style="list-style-type: none"> DensNet121 InceptionresNet_v2 ResNet 50 SeNet-154
	$F\beta$ -점수	99.89	<ul style="list-style-type: none"> CRNN Deeplab_v3+ grad-CAM, GAN LightGBM ResNet, DenseNet ResNet50 U-net Xception
행동 인식 (Action Recognition)	1DF1	30.5	<ul style="list-style-type: none"> Faster RCNN+Hierarchical clustering
	AP	77	<ul style="list-style-type: none"> P-GCN
	BLEU	16.33	<ul style="list-style-type: none"> Gradient descent fine tuning of Korean GPT model with custom language model head
	mAP	82.2	<ul style="list-style-type: none"> Faster R-CNN
	Mean Gaze Tracking Error	1.64	<ul style="list-style-type: none"> Linear Regression
	Mean Position Error(pixel)	3.26	<ul style="list-style-type: none"> YOLOv3
	mIoU	89.2	<ul style="list-style-type: none"> SOLOv2 RITNet
	MOTA	60.69	<ul style="list-style-type: none"> Faster RCNN+Kalman Filter+ Hungarian Algorithm

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
	PCK	97.6	<ul style="list-style-type: none"> AlphaPose GlobalNet
	Top-K 정확도	96.99	<ul style="list-style-type: none"> PoseC3D R(2+1D)
	정확도	97.21	<ul style="list-style-type: none"> MS-G3D PoseC3D
	$F\beta$ -점수	98.92	<ul style="list-style-type: none"> GCN GRU-D MS-G3D ResNet ST-GCN
[비디오] 비디오 분류 (Video Classification)	AUROC	100	<ul style="list-style-type: none"> 3D CNN
	mA	84.59	<ul style="list-style-type: none"> PAR
	micro F1	51.38	<ul style="list-style-type: none"> Video Transformer
	Sensitivity	100	<ul style="list-style-type: none"> 3D CNN
	Specificity	100	<ul style="list-style-type: none"> 3D CNN
	정밀도	88	<ul style="list-style-type: none"> InceptionResNetV2 MobileNetV2
	정확도	91.53	<ul style="list-style-type: none"> EfficientNet GRU Inception v3
	$F\beta$ -점수	77.41	<ul style="list-style-type: none"> Ensemble Multi-Modal Video Transformer
[비디오] 자세 추정 (Pose Estimation)	mASS(meanAverage ShapeSimilarity)	68.82	<ul style="list-style-type: none"> -
	정확도	88.51	<ul style="list-style-type: none"> C3D, CNN_RNN, SVM
	$F\beta$ -점수	99.32	<ul style="list-style-type: none"> A2J
[비디오] 비디오 인식 (Video Recognition)	PSNR	41.3	<ul style="list-style-type: none"> MPRnet PEPSI
	SSIM		<ul style="list-style-type: none">
[비디오] 키포인트 검출 (Keypoint Detection)	mAP	92.98	<ul style="list-style-type: none"> Mask R-CNN

〈표 Ⅲ-165〉 비디오 Task 분야별 유효성 측정지표

2.4.3.4 오디오

- ① 데이터 특성, ② 판정표(Confusion Matrix) 작성이 가능한 경우, ③ 인공지능 학습데이터 구축 목표 또는 임무와 2.3 의미 정확성의 오디오 데이터에 해당하는 학습 Task를 참조하여 적합한 유효성 측정지표를 선택할 수 있음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
오디오 분류 (Audio Classification)	$F\beta$ -점수	99.35	<ul style="list-style-type: none"> • AE-CNN • CapsuleNet • ResNet
	재현율	99.15	<ul style="list-style-type: none"> • AE-CNN
	정밀도	99.56	<ul style="list-style-type: none"> • AE-CNN
	정확도	98.99	<ul style="list-style-type: none"> • AE-CNN • Convolution 2D • CNN 기반 자체 개발 알고리즘
음성 인식 (Speech Recognition)	BLEU	23.7	<ul style="list-style-type: none"> • Transformer
	CER(문자)	96.6	<ul style="list-style-type: none"> • bi-LSTM • Hybrid HMM / DNN • kospeech(kosr) / CNN-biLSTM-CTC • TDNN, TDNN-F • TDNN-LSTM • Conformer, Jasper • Transformer • Transformer Conformer
	EER	97.12	<ul style="list-style-type: none"> • X-vector
	ESTOI	82	<ul style="list-style-type: none"> • TSCN
	$F\beta$ -점수	34	<ul style="list-style-type: none"> • TSCN
	WER(단어)	88.99	<ul style="list-style-type: none"> • bi-LSTM • TDNN, TDNN-F • TDNN-LSTM • Conformer, Jasper • Transformer • Transformer Conformer
텍스트 분류 (Text Classification)	CER(문자)	99.22	<ul style="list-style-type: none"> • Seq2Seq, LSTM • WAV2VEC 2.0

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
음성 합성 (Speech Synthesis)	MOS	95.91	<ul style="list-style-type: none"> FastSpeech + WaveGlow HiFiSinger+PWGAN Tacotron2 + GST Tacotron2 + Waveglow
	벤치마크 데이터 리커트 스케일 비교	83.13	<ul style="list-style-type: none"> LSTM

〈표 III-166〉 오디오 Task 분야별 유효성 측정지표

2.4.3.5 3D

- ① 데이터 특성, ② 판정표(Confusion Matrix) 작성이 가능한 경우, ③ 인공지능 학습데이터 구축 목표 또는 임무와 2.3 의미 정확성의 3D 데이터에 해당하는 학습 Task를 참조하여 적합한 유효성 측정지표를 선택할 수 있음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
3D 객체 인식 (3D Object Detection)	mAP	99.99	<ul style="list-style-type: none"> SECOND Votenet Voxel R-CNN YOLOv5
	mAP@IoU	95.72	<ul style="list-style-type: none"> darknet YOLOv4 PointPillars segformer
	mIoU	97.01	<ul style="list-style-type: none"> DeepVPCalib Mask R-CNN(Detectron2) ResNet
	MOTA	64.59	<ul style="list-style-type: none"> Point RCNN+Kalman Filter+Hungarian Algorithm
	정확도	99.25	<ul style="list-style-type: none"> Real-Time Seamless Single Shot 6D Object Pose Prediction
3D 포인트클라우드 분류 (3D Point Cloud Classification)	mAcc	84.2	<ul style="list-style-type: none"> KPConv
	mAP	96.11	<ul style="list-style-type: none"> YOLOv4

Task 분야	측정지표	정량목표 (예시)	학습모델(알고리즘 예시)
	mIoU	76.45	• KPConv
	OA	98.7	• KPConv
3D 자세 추정 (3D Pose Estimation)	2D Reprojection 정확도	90.23	• Real-Time Seamless Single Shot 6D Object Pose Prediction
	Top-K 정확도	82.31	• ST-GCN
3D 객체 복원 (3D Object Reconstruction)	mAP	50	• MVXNet
2D 이미지를 통한 시점 합성 (2D View Synthetic)	mAP	60	• SECOND
	F β -점수	91.3	• Volo-D1
3D 키포인트 검출 (3D Keypoint Detection)	AP@oks	99	• Simple baseline
	mAP	97.55	• Point-LSTM

〈표 III-167〉 3D Task 분야별 유효성 측정지표

붙임

1. 품질지표·기준 선택 예시

1. 전제조건

- 이 문서에서 선택 예시로 기술되는 내용은 한국지능정보사회진흥원에서 수행한 인공지능 학습용 데이터 사업결과를 바탕으로 작성되었기 때문에 다양한 도메인의 인공지능 학습용 데이터 구축사업에 모두 적용하기는 어려우며, 사업목표 또는 임무별로 품질에 대한 측정지표와 기준이 변경될 수 있음
- 예시가 없는 경우는 해당 도메인의 전문가 자문 및 근거자료 등을 통해 품질에 대한 적절한 측정지표와 기준을 선택하는 것이 필요함

2. 다양성

- 다양성의 통계, 요건에 대한 항목은 다양성 지표 구성이 지도학습 계열 인공지능이 처리해야 하는 실제 세상의 데이터와 유사한 특성으로, 다양성(통계), 다양성(요건)에 대한 항목을 명확하게 분류하기 어렵기 때문에 별도의 예시를 기술하지 않음
- 인공지능 학습용 데이터 구축 사업 계획 시 도출된 다양성 항목 중 사업의 목표, 목적, 임무 달성에 필요한 요구사항으로 분류되는 경우 다양성(요건)으로 분류하는 것을 원칙으로 함

3. 의미 정확성 예시

3.1 텍스트 데이터 예시

- 텍스트 데이터 품질에 대한 검사항목, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
텍스트 분류 (Text Classification)	Class Labeling	KSIC 분류	정확도	91.33
		관찰기록에 대한 역량평가		79.1
		기술용어 정확도		99
		문법적 정확성		82.75
		분류 라벨 및 영역 정확도		97.84

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
		분류 체계 정확도		100
		비윤리 어휘 적합성		91.92
		비윤리문장의 문형 매핑		92.21
		비윤리성 유형 분류 정확도		90.37
		에세이 난이도		90.5
		에세이 타입		90.38
		우울한 기분 진단 분류		100
		유·불리 판단 정확성		60.38
		의도분류(민원분야/술어)		94.06
		주제 분류 정확도		98.3
		학생 학년		100
		화형 분류		98.02
	Speech to Text	우울한 기분 진단 분류 정확도	정확도	100
	Tagging	개인정보 비식별화 및 민감 이슈 처리	정확도	100
		개체명		94.69
		신조어 라벨링		99
텍스트 요약 (Text Summary)	Class Labeling	Scoring	정확도	89.19
		에세이 주제		91.76
		요약문 분류 정확성		99.1
		키워드 정확도		100
	Sequence Labeling	원문-요약문 주제(의미) 일치도	정확도	99.53
	Text Summary	영상에 대한 요약문 연관성	정확도	99.4
		요약문 품질 정확도		98.93
기계 번역 (Machine Translation)	Machine Translation	번역 정확도	PTQA	88
		영-한 번역 정확도	정확도	95.03
		영-한 번역 점수		100

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
순차적 레이블링 (Sequence Labeling)	Class Labeling	“비윤리 문장”, “비도덕(T/F)분류”, “비도덕 강도” 간의 관계적합성	정확도	86.46
	Sequence Labeling	음성-스크립트(대본) 일치도	정확도	99.99
	Tagging	단락-문장 구조 정확도	정확도	100
		단락-문장 구조 정확도(문장 흐름)		99.06
		판결문 주장/사실/판단 태깅 정확성		73.33
질의 응답 (Question Answering)	Class Labeling	유사질의 다양성	리커트 척도	63.4
	Sequence Labeling	답변 문장 품질	정확도	99.67
		영상에 대한 질의응답 정확성		99.25
		지문-질의-응답 간 내용 일치성		99.3
	Speech to Text	질의 문장 품질	정확도	99.16
음성 인식 (Speech Recognition)	Tagging	화자 라벨링 정확도	정확도	94.66

〈표 III-168〉 텍스트 데이터 예시(의미 정확성)

3.2 이미지 데이터 예시

- 이미지 데이터 품질에 대한 검사항목, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
객체 인식 (Object Detection)	Bounding Box	바운딩박스 정확도	$F\beta$ -점수	99.6
		분류 라벨 정확도		100
		사건/사고 이벤트 분류 정확도		99.09
	Keypoint	2D 키포인트 정확도	$F\beta$ -점수	99.56
	Polygon	폴리곤 정확도	$F\beta$ -점수	99.4
		세그멘테이션 정확도		99.95

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
키포인트 검출 (Keypoint Detection)	Keypoint	분류 라벨 정확도(키포인트)	$F\beta$ -점수	99.92
		키포인트 정확도		99.73
얼굴 인식 (Face Recognition)	Bounding Box	바운딩박스 정확도	정밀도	98.78
	Keypoint	가상인물 기반 인물 스케치 작성	정확도	99.04
		설명문 기반 몽타주 스케치 작성		100
	Polygon	폴리곤 정확도	정밀도	97.99
이미지 분류 (Image Classification)	Bounding Box	객체 라벨링 정확도	$F\beta$ -점수	95.09
		기상특보별 어노테이션 정확도		95.85
		바운딩박스 정확도		99.66
		바운딩박스 정확도_객체		98.24
		바운딩박스 정확도_행동		95.1
		바운딩박스 정확도	정확도	99.99
		바운딩박스, 키포인트 정확도		98.48
	Keypoint	분류 라벨 정확도(동작)	$F\beta$ -점수	100
		바운딩박스, 키포인트 정확도	정확도	98.48
	Polygon	4 포인트 폴리곤 정확도	$F\beta$ -점수	99.5
		폴리곤 정확도		98.02
		폴리곤 라벨 정확도	정밀도	94.69
		폴리곤 정확도	정확도	83.01
	Polyline	폴리라인 정확도	$F\beta$ -점수	97.77
	Tagging	가상인물 기반 설명문 작성	정확도	99.78
		편집정확도		98.9

〈표 III-169〉 이미지 데이터 예시(의미 정확성)

3.3 비디오 데이터 예시

- 비디오 데이터 품질에 대한 품질측정항목, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
객체 인식 (Object Detection)		라벨링 정확도	$F\beta$ -점수	99.45
		바운딩박스 라벨 정확도		100
		바운딩박스 라벨 정확도(추적 대상)		96.36
		바운딩박스 라벨 정확도(추적 비대상)		98.9
		바운딩박스 정확도		100
		분류 라벨 정확도		99.82
	Bounding Box	2D 바운딩박스 라벨 정확도	재현율	96.81
		LiDAR		100
		다각도 RGB		95.53
		바운딩박스 라벨 정확도		100
		바운딩박스 라벨 정확도(얼굴)		99.64
		바운딩박스 라벨 정확도(객체)		100
		바운딩박스 라벨 정확도(실화상)		93.24
		바운딩박스 라벨 정확도(인물 전신)		100
		바운딩박스 정확도		100
		분류 라벨 정확도		100
		열화상		90.77
		엽록소형광		100
		탐뷰 RGB		100
		통과차량 라벨 정확도		99.28
		2D 바운딩박스 라벨 정확도	정밀도	79.78
		LiDAR		98.51
		다각도 RGB		95.08
		바운딩박스 라벨 정확도		99.63
		바운딩박스 라벨 정확도(얼굴)		23.19
		바운딩박스 라벨 정확도(객체)		95.98
		바운딩박스 라벨 정확도(실화상)		95.56
		바운딩박스 라벨 정확도(인물 전신)		47.13

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
		바운딩박스 정확도		99.69
		분류 라벨 정확도		99.88
		열화상		100
		염록소형광		100
		탑뷰 RGB		98.25
		통과차량 라벨 정확도		99.56
		바운딩박스 정확도	정확도	95.47
		분류 라벨 및 영역 정확도		95.45
		분류 라벨 정확도		100
		얼굴 영역 바운딩 박스		88.93
	Cuboid	포인트 클라우드 정확도	$F\beta$ -점수	94.86
	Keypoint	도트(dot) 라벨 정확도	정밀도	99.74
		얼굴마커 프로젝션 정확도	정확도	96.85
	Polygon	폴리곤 라벨 정확도	$F\beta$ -점수	99.85
		폴리곤 라벨 정확도(RGB, Depth)		94
		폴리곤 정확도		98.55
		폴리라인 라벨 정확도	MOTA	98.75
		바운딩박스 라벨 정확도	재현율	79.37
		폴리곤 라벨 정확도		100
		폴리곤 라벨 정확도(RGB)		99.64
		폴리곤 라벨 정확도(당도)		100
		폴리곤 라벨 정확도(카메라)		96.8
		폴리곤 라벨 정확도(NIR)		98.47
		폴리곤 정확도		57.69
		바운딩박스 라벨 정확도	정밀도	32.63
		바운딩박스 라벨 정확도(RGB)		89.38
		폴리곤 라벨 정확도		100
		폴리곤 라벨 정확도(RGB)		88.97
		폴리곤 라벨 정확도(당도)		85.16
		폴리곤 라벨 정확도(카메라)		92.99
		폴리곤 라벨 정확도(NIR)		82.53
		폴리곤 정확도		98.82

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
		경동맥 세그멘테이션 정확도	정확도	100
		분류 라벨 및 영역 정확도		95.45
		분류 라벨 정확도		98.86
		분할 라벨 정확도		100
		뼈 세그멘테이션 정확도		99.6
		척추체 세그멘테이션 정확도		96.66
		퇴행성 경추 협착증 라벨링 정확도		98.04
		퇴행성 요추 디스크 라벨링 정확도		96.19
		흉추 척수증 라벨링 정확도		98.04
	Polyline	폴리라인 라벨 정확도	재현율	88.64
		폴리라인/포인트 라벨 정확도		95.89
		폴리라인 라벨 정확도	정밀도	68.42
		폴리라인/포인트 라벨 정확도		97.21
		경동맥 세그멘테이션 정확도	정확도	100
		골다공증성 골절 통증 라벨링 정확도		100
		만성신경통증 라벨링 정확도		95.72
		분류 라벨 정확도		98.86
		뼈 라인 정확도		98.65
		퇴행성 경추 협착증 라벨링 정확도		98.04
		퇴행성 슬관절 라벨링 정확도		96.04
		퇴행성 요추 디스크 라벨링 정확도		96.19
		퇴행성 요추 협착증 라벨링 정확도		95
		회전근개 파열 라벨링 정확도		95.1
	Segmentation	세그멘테이션 정확도	재현율	100
		세그멘테이션 정확도	정밀도	98.56
이미지 분류 (Image Classification)	Bounding Box	분류 라벨 정확도	$F\beta$ -점수	99.05
		분류 라벨 정확도	정밀도	100
		분류 라벨 정확도	정확도	92
		수술도구 분류 라벨 정확도		99.74
		수술도구 세그멘테이션 정확도		99.58
		약제 이미지 바운딩박스 정확도		100
		약제 이미지와 라벨 정확도		100

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
	Class Labeling	한국어 문장의 재난 정보 분류정확도	정확도	99.6
	Keypoint	키포인트 정확도	$F\beta$ -점수	96.1
	Polygon	폴리곤 정확도	$F\beta$ -점수	99.2
		OCR 라벨 정확도	재현율	99.96
		폴리곤 라벨 정확도(전체 이미지)		99.67
		폴리곤 라벨 정확도(부분 이미지)		100
		폴리곤 라벨 정확도	정밀도	96.98
		폴리곤 라벨 정확도(전체 이미지)		96.14
		폴리곤 라벨 정확도(부분 이미지)		100
		OCR 라벨 정확도	정확도	99.91
행동 인식 (Action Recognition)	Bounding Box	보행자 라벨 정확도	재현율	97.99
		보행자 라벨 정확도	정밀도	99.24
		구간 라벨 정확도	정확도	92
	Keypoint	수어 영상 전사 정확도 (형태소/비수지 가공)	정확도	95.47
	Sequence Labeling	주제 제스처 간의 의미 및 행동 일치성	정확도	97.21
	Tagging	태깅 정확도	$F\beta$ -점수	98.59
비디오 분류 (Video Classification)	Bounding Box	바운딩박스 라벨 정확도	재현율	96.85
		바운딩박스 라벨 정확도	정밀도	97.49
		BMD class 분류 라벨링 정확도	정확도	100
		부위별 구조 어노테이션 정확도		99.4
		분류 라벨 및 영역 정확도		97.58
		분류 라벨 정확도		100
		소장 병변 분류 라벨 정확도		99.2
		수면 영상과 라벨 데이터의 시간 동기화		99.09
		수술단계 분류 라벨 정확도		100
		폐/병변 어노테이션 정확도		98.1
		해부학적 구조물 및 병변 세그멘 테이션 정확도		100
	Class Labeling	분류 라벨 정확도	정확도	99.69

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
	Keypoint	어노테이션 정확도	정확도	99.13
	Polygon	폴리곤 라벨 정확도	재현율	98.24
		폴리곤 라벨 정확도	정밀도	98.55
		뇌전이암 세그멘테이션 정확도	정확도	97.14
		부위별 구조 분류 라벨 정확도		99.6
		분류 라벨 및 분할 정확도		97.49
		분류 라벨 및 영역 정확도		97.58
		소아 흉부질환 분류 라벨 정확도		99.2
		질환 분류 라벨 정확도		90.71
	Polyline	폴리라인 라벨 정확도	재현율	97.36
		폴리라인 라벨 정확도	정밀도	94.98
	Tagging	하이라이트 구간 라벨링 결과 간 유사도	동일응답 비율	79.67
		하이라이트 구간 태깅 결과 간 코사인 유사도	평균 코사인 유사도	82.28
자세 추정 (Pose Estimation)	Keypoint	분류 라벨 정확도	정확도	96.88
비디오 인식 (Video Recognition)	Bounding Box	실화상-열화상 동기화 정합도	정합도	95.13
	Sequence Labeling	멀티모달 데이터 시점 일치	정확도	100
키포인트 검출 (Keypoint Detection)	Keypoint	키포인트 정확도	AP@oks	98.68
		키포인트 정확도	F β -점수	99.3
		키포인트 정확도	OKS	99.96
		키포인트 속성 분류 정확도	라벨 분류 정확도	98.89
		키포인트 라벨 정확도	재현율	50.86
		키포인트 라벨 정확도(객체)		100
		키포인트 정확도		100
		키포인트 라벨 정확도	정밀도	77.08
		키포인트 라벨 정확도(객체)		71.06
		키포인트 정확도		97.6
		분류 라벨 정확도(제스처)	정확도	100

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
		스켈레톤 투사 정확도		96.85
		키포인트 정확도		98.65
		키포인트 정확도(손)		59.38

〈표 III-170〉 비디오 데이터 예시(의미 정확성)

3.4 오디오 데이터 예시

- 오디오 데이터 품질에 대한 검사항목, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
오디오 분류 (Audio Classification)	Class Labeling	구문 정확도	정확도	95.38
		발화 라벨링 정확도		98.33
		발화 주제 분류정확도		98.44
		분류 라벨 정확도(음성)		95.95
		소리 분류 정확도		98.65
		소음클래스 일치도		99.82
		전사 정확도		99.99
		주제 분류 정확도		90.14
		주제 분류 정확도(도메인)		98.7
		주제별 분포 정확도		91.66
		카테고리 분류 정확도		95.62
		화자 규모 정확도		98.12
	Sequence Labeling	노트 가사 / 뱀당 / 바이브레이션 / 호흡 일치성	정확도	99.7
		노트 시작 시간과 지속 시간 일치성		99.97
		발화 시작 / 끝 위치 정확도		78.6
		음성 시작-끝 발화구간 정확도		99.1
	Tagging	발화 시작 / 끝 시간 라벨링	정확도	78.28
		발화 시작 시간과 끝 시간 정확도		91.07
		발화시간 및 종료시간 일치성		100
		발화시작/끝시간 태깅 확인		99.68

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
음성 인식 (Speech Recognition)	Class Labeling	태깅 정보 정확도		89.37
		감정표현 정확도	정확도	93.01
		음성품질 정확도		99.31
	Machine Translation	감정 태깅 정보 정확도	정확도	71.29
	Sequence Labeling	노트 가사 / 밴딩 / 바이브레이션 / 호흡 일치성	정확도	99.02
		노트 시작 시간과 지속 시간 일치성		99.99
		노트 피치 일치성		99.99
	Speech to Text	번역 정확도	리커트 척도	82.2
		전사 정확도	정확도	99.99
		전사규칙 부합도		99.53
		전사품질 정확도		99.54
		철자 전사 정확성 (발화자 의도 정확성) (단어, 문장, 문단)		96.94
		철자 전사 정확성 (발화자 의도 정확성) (준자유, 자유)		99.07
	Tagging	발화 시작 / 끝 위치 정확도	정확도	84.68
		발화 시작 시간, 발화 종료 시간 라벨 정확도		94.51
		음성 구간 어노테이션 정확도		100
		전사 정확도		71.41
텍스트 분류 (Text Classification)	Class Labeling	문장 품질	정확도	99.23
		상담유형 및 주제분류 정확도		93.59
		카테고리 분류 정확도		99.14
	Sequence Labeling	발화스크립트 중복 확인	정확도	100
음성 합성 (Speech Synthesis)	Class Labeling	감성 상태 분류 정확도	정확도	94.42
		감정 상태 분류 정확도		94.46
		발화스타일 분류 정확도		96.98
		캐릭터 감정 분류 정확도		90.52

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
	Sequence Labeling	발화 시작 시간, 발화 종료 시간 라벨 정확도	정확도	99.2
	Speech to Text	전사 정확도	정확도	99.99
	Tagging	발화 주제와 발화 음성 간의 적합성	정확도	94.53

〈표 III-171〉 오디오 데이터 예시(의미 정확성)

3.5 3D 데이터 예시

- 3D 데이터 품질에 대한 검사항목, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
3D 객체 인식 (3D Object Detection)	Bounding Box	바운딩박스 정확도	$F\beta$ -점수	99.4
		소방시설 객체 바운딩박스 정확도		98.16
		바운딩박스 라벨 정확도	재현율	98.08
		바운딩박스 라벨 정확도	정밀도	94.74
	Cuboid	큐보이드 정확도	$F\beta$ -점수	99.01
		큐보이드 라벨 정확도(추적 대상)	MOTA	96.13
		2.5D 바운딩박스 라벨 정확도	재현율	93.31
		큐보이드 라벨 정확도		100
		큐보이드 라벨 정확도(LiDAR)		94.96
		큐보이드 라벨 정확도_PCD		100
		2.5D 바운딩박스 라벨 정확도	정밀도	98.79
		큐보이드 라벨 정확도		97.55
		큐보이드 라벨 정확도(LiDAR)		97.95
		큐보이드 라벨 정확도_PCD		96.66
	Polygon	폴리곤 정확도	$F\beta$ -점수	99.6
		폴리곤 라벨 정확도	재현율	98.04
		폴리곤 라벨 정확도	정밀도	97.33
3D 포인트클라우드 분류	Cuboid	큐보이드 라벨 정확도	재현율	96.92
		큐보이드 라벨 정확도	정밀도	96.92

Task 분야	라벨 유형	검사항목	측정지표	정량목표 (예시)
(3D Point Cloud Classification)				
3D 자세 추정 (3D Human Pose Estimation)	Cuboid	큐보이드(3D 바운딩박스) 정확도	$F\beta$ -점수	94.04
3D 키포인트 검출 (3D Keypoint Detection)	Keypoint	3D 키포인트 정확도	$F\beta$ -점수	99.32

〈표 III-172〉 3D 데이터 예시(의미 정확성)

4. 유효성 예시

4.1 텍스트 데이터 예시

- 텍스트 데이터에 대한 학습모델, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
텍스트 분류 (Text Classification)	VGG	AUROC	99
	KorBERT	MAE	59
	KorBERT	MSE	71
	BERT, LSTM	PCC	82.28
	BERT, LSTM	QWK	81.16
	BERT	정밀도	86.1
	ELECTRA	정확도	98.68
	BERT	$F\beta$ -점수	90.8
	ELECTRA-base, KoBERT		
	KoBERT		
	KorBERT		
	NBC + FastText		
	VGG		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
텍스트 요약 (Text Summary)	T5	ROUGE-L	37.55
	Transfomer		
	BERT	Top-K 정확도	95
	BERT		
	ELECTRA-base, KoBERT	F β -점수	89.85
	KorBERT, BERT		
기계 번역 (Machine Translation)	Transfomer	BLEU	77.22
	ELECTRA-base, KoBERT	F β -점수	91.3
순차적 레이블링 (Sequence Labeling)	ELECTRA-base, KoBERT	F β -점수	81.17
질의 응답 (Question Answering)	Hierarchical Conditional Relation Networks for Video QA	정확도	77.78
	ELECTRA	F β -점수	85.88
음성 인식 (Speech Recognition)	SVM	F β -점수	76.05

〈표 III-173〉 텍스트 데이터 예시(유효성)

4.2 이미지 데이터 예시

- 이미지 데이터에 대한 학습모델, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
객체 인식 (Object Detection)	EfficientNet B0	AP	38.78
	Scaled-Yolov4		
	Scaled-Yolov4	AR	72.24
	ConvMath_VIT	BLEU	81.92
	U-net	Correlation	71
	StyleGAN2 ADA	FID	10.44
	YOLOv5	FPS	20.93
	EfficientDet	mAP	96.87
	Fast R-CNN		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	Mask R-CNN		
	RNN, DNN		
	SIMPLE-HRNET		
	YOLACT		
	YOLO		
	YOLOv4		
	YOLOv4, KNN algorithm		
	YOLOv5		
	Mask R-CNN	mAP@IoU	93.6
	YOLOv4		
	YOLOv5		
	Swin Transformer	mIoU	74.74
	EfficientNet B0	Top-K 정확도	100
	SlowFast ++		
	Inception-v4	재현율	80.48
	U-net		
	Inception-v4	정확도	90.81
	ResNet		
	U-net		
	ResNet	F β -점수	96
	U-net		
	YOLO		
	YOLOv4		
키포인트 검출 (Keypoint Detection)	AlphaPose	mAP	99.29
	YOLOv5		
	Modified VGGNet	정확도	94.48
얼굴 인식 (Face Recognition)	RetinaFace	AP	98.76
	StyleGAN2	FID	7.94

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	ArcFace	Kinship Verification 정확도	70.66
	CNN network for feature extraction	MAE	3.11
	ArcFace	Rank-K 정확도	81.9
	ArcFace	정확도	99.31
	CNN network for feature extraction		
이미지 분류 (Image Classification)	FCDD	AUROC	98.24
	Mask R-CNN	Grid Cell 정확도	97.47
	Mask R-CNN	mAP	99.1
	YOLOv5		
	Swin Transformer	mIoU	85.93
	U-net		
	U-net	Overall Pixel 정확도	95.95
	ResNet	Top-K 정확도	95.3
	mobileNet		
	U-net	정확도	91.73
	CRAFT	$F\beta$ -점수	99.6
	Customized Mask R CNN		
	HRCenterNet		
	Mask R-CNN		
	VGG		
	YOLOv5		

〈표 III-174〉 이미지 데이터 예시(유효성)

4.3 비디오 데이터 예시

- 비디오 데이터에 대한 학습모델, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
객체 인식 (Object Detection)	CNN based Segmentation model	AUROC	95.38
	RNN,Transformer		
	WGAN		
	XGBoost, LightGBM, CatBoost, NeuralnetMXnet, NetFastAI		
	RetinaNet, U-Net	DSC	94.14
	U-Net, FCN		
	Random Forest	MAE	99.6
	DarkNet53		
	DetectoRS	mAP	98.8
	EfficientDet		
	EfficientDet D3		
	General (A1)		
	Mask R-CNN		
	Scaled-YOLOv4		
	Swin-Transformer		
	UNET		
	YOLACT		
	YOLO, EfficientDet		
	YOLOv3		
	YOLOv3-tiny, YOLOv4		
	YOLOv5		
	darknet YOLOv4	mAP@IoU	99.835
	DetectoRS		
	Faster R-CNN		
	Mask R-CNN		
	RetinaNet		
	segformer		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	YOLOv4		
	YOLOv5		
	DeepLab v3+		
	3D CNN	mIoU	98.9
	HRNet		
	Mask R-CNN		
	ResNet		
	UNET		
	Polynomail Regression	NRMSE	92.5
	Deep Encoder-based Predictor	RE	95.7
	CNN	RMSE	0.14
	RetinaNet, U-Net	Sensitivity	92
	WGAN		
	WGAN	Specificity	99
	ResNet 50	정확도	91.53
	VGG, ResNet, EfficientNet		
	Resnet50	$F\beta$ -점수	99.29
	TransFG		
	U-net		
	VTN		
	XG,Boost		
	XGBoost, LightGBM, CatBoost, NeuralnetMXnet, NetFastAI		
이미지 분류 (Image Classification)	Densenet-161	AUROC	100
	DenseNet201		
	EfficientNet-B0 기반 CNN 모델		
	Multi-Resolution Input CNN		
	U-net	Correlation	83
	CNN based Segmentation model	DSC	99.47
	EfficientNet B0모델을 encoder로 하는 FPN		
	U-net		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	U-net	IoU	89
	YOLOv5	mAP	90.1
	ResNet-50	MRE	1.737
	LSTM	RMSE	99.944
	RetinaNet, YOLO	Sensitivity	90.32
	YOLO		
	XGBoost	SMAPE	85.91
	EfficientNet-B0	Top-K 정확도	100
	Resnet152		
	Inception V3+Triplet Loss		
	DensNet121	정확도	100
	InceptionresNet_v2		
	ResNet 50		
	SeNet-154		
	CRNN	$F\beta$ -점수	99.89
	Deeplab_v3+		
	grad-CAM, GAN		
	LightGBM		
	ResNet, DenseNet		
	ResNet 50		
	U-net		
	Xception		
행동 인식 (Action Recognition)	Faster RCNN+Hierarchical clustering	1DF1	30.5
	P-GCN	AP	77
	Gradient descent fine tuning of Korean GPT model with custom language model head	BLEU	16.33
	Faster R-CNN	mAP	82.2
	Linear Regression	Mean Gaze Tracking Error	1.64
	Yolov3	Mean Position Error(pixel)	3.26
	SOLOv2	mIoU	89.2

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	RITNet		
	Faster RCNN+Kalman Filter+Hungarian Algorithm	MOTA	60.69
	AlphaPose	PCK	97.6
	GlobalNet		
	PoseC3D	Top-K 정확도	96.99
	R(2+1D)		
	MS-G3D	정확도	97.21
	PoseC3D		
	GCN	$F\beta$ -점수	98.92
	GRU-D		
	MS-G3D		
	ResNet		
	ST-GCN		
비디오 분류 (Video Classification)	3D CNN	AUROC	100
	PAR	mA	84.59
	Video Transformer	micro F1	51.38
	3D CNN	Sensitivity	100
	3D CNN	Specificity	100
	InceptionResNetV2	정밀도	88
	MobileNetV2		
	EfficientNet	정확도	91.53
	GRU		
	Inception v3		
	Ensemble	$F\beta$ -점수	77.41
	Multi-Modal		
	Video Transformer		
자세 추정 (Pose Estimation)	C3D, CNN_RNN, SVM	정확도	88.51
	A2J	$F\beta$ -점수	99.32
비디오 인식 (Video Recognition)	MPRnet	PSNR	41.3
	PEPSI		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
키포인트 검출 (Keypoint Detection)	Mask R-CNN	mAP	92.98

〈표 III-175〉 비디오 데이터 예시(유효성)

4.4 오디오 데이터 예시

- 오디오 데이터에 대한 학습모델, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
오디오 분류 (Audio Classification)	AE-CNN	재현율	99.15
	AE-CNN	정밀도	99.56
	Convolution 2D	정확도	98.99
	AE-CNN		
	CNN 기반 자체 개발 알고리즘		
	AE-CNN	F β -점수	99.35
	CapsuleNet		
	ResNet		
음성 인식 (Speech Recognition)	Transformer	BLEU	23.7
	bi-LSTM	CER(문자)	96.6
	Hybrid HMM / DNN		
	kospeech(kosr)/CNN-biLSTM-CTC		
	TDNN, TDNN-F		
	TDNN-LSTM		
	Conformer, Jasper		
	Transformer		
	Transformer Conformer		
	X-vector	EER	97.12
	TSCN	ESTOI	82
	bi-LSTM	WER(단어)	88.99
	TDNN, TDNN-F		

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	TDNN-LSTM		
	Conformer, Jasper		
	Transformer		
	Transformer Conformer		
	TSCN		
텍스트 분류 (Text Classification)	Seq2Seq, LSTM	F β -점수	34
	WAV2VEC 2.0		
음성 합성 (Speech Synthesis)	FastSpeech + WaveGlow	MOS	95.91
	HiFiSinger+PWGAN		
	Tacotron2 + GST		
	Tacotron2 + Waveglow		
	LSTM	벤치마크 데이터 리커트 스케일 비교	83.13

〈표 III-176〉 오디오 데이터 예시(유효성)

4.5 3D 데이터 예시

- 3D 데이터에 대한 학습모델, 측정지표, 정량목표 예시는 다음과 같음

※ 제시한 정량목표(예시)는 과제 특성에 따라 다르게 적용될 수 있으므로 참고용으로 활용

Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
3D 객체 인식 (3D Object Detection)	SECOND	mAP	99.99
	Votenet		
	Voxel RCNN		
	YOLOv5		
	darknet YOLOv4	mAP@IoU	95.72
	PointPillars		
	segformer		
	DeepVPCalib	mIoU	97.01

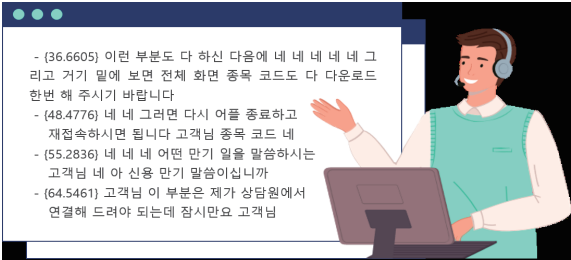
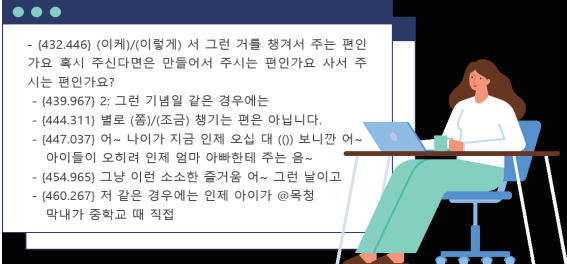
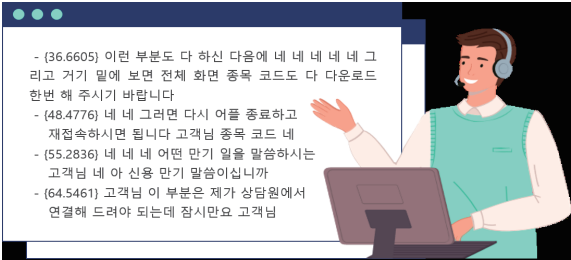
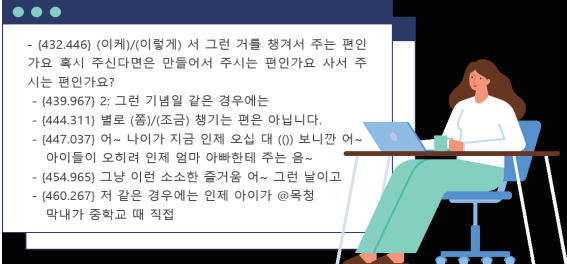
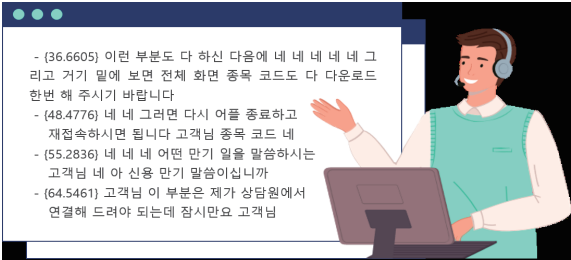
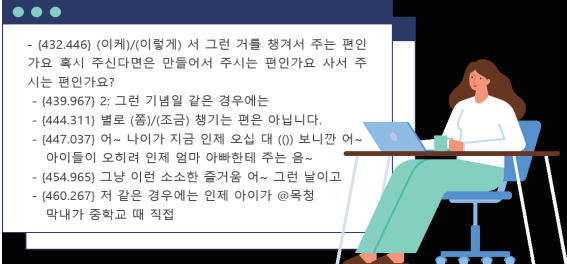
Task 분야	학습모델(알고리즘)	측정지표	정량목표 (예시)
	Mask R-CNN		
	ResNet		
	Point RCNN+Kalman Filter+Hungarian Algorithm	MOTA	64.59
	Real-Time Seamless Single Shot 6D Object Pose Prediction	정확도	99.25
3D 포인트클라우드 분류 (3D Point Cloud Classification)	KPConv	mAcc	84.2
	YOLOv4	mAP	96.11
	KPConv	mIoU	76.45
	KPConv	OA	98.7
3D 자세 추정 (3D Pose Estimation)	Real-Time Seamless Single Shot 6D Object Pose Prediction	2D Reprojection 정확도	90.23
	ST-GCN	Top-K 정확도	82.31
3D 객체 복원 (3D Object Reconstruction)	MVXNet	mAP	50
2D 이미지를 통한 시점 합성 (2D View Synthetic)	SECOND	mAP	60
	Volo-D1	$F\beta$ -점수	91.3
3D 키포인트 검출 (3D Keypoint Detection)	Simple baseline	AP@oks	99
	Point-LSTM	mAP	97.55



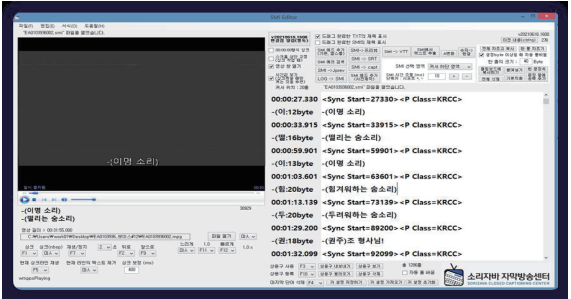


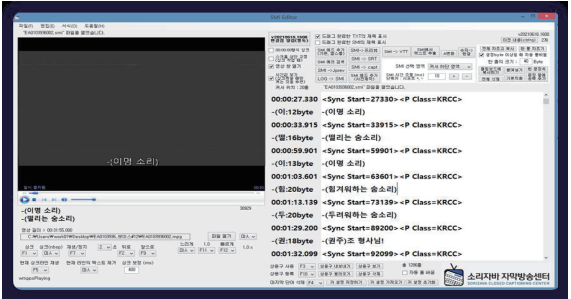


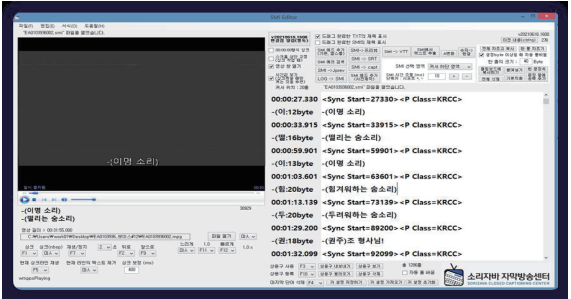
〈표 III-177〉 3D 데이터 예시(유효성)

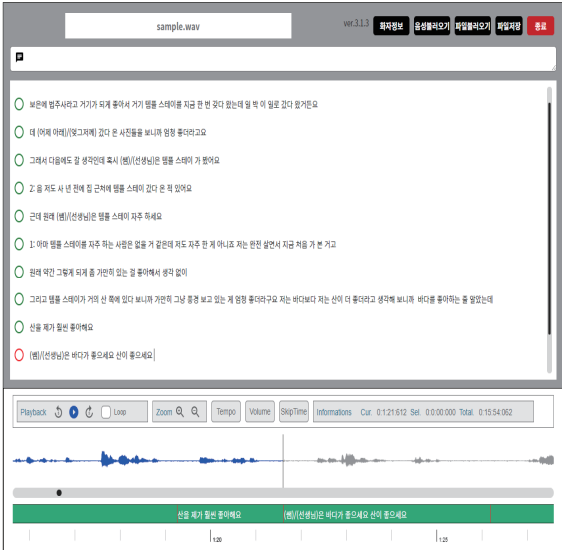
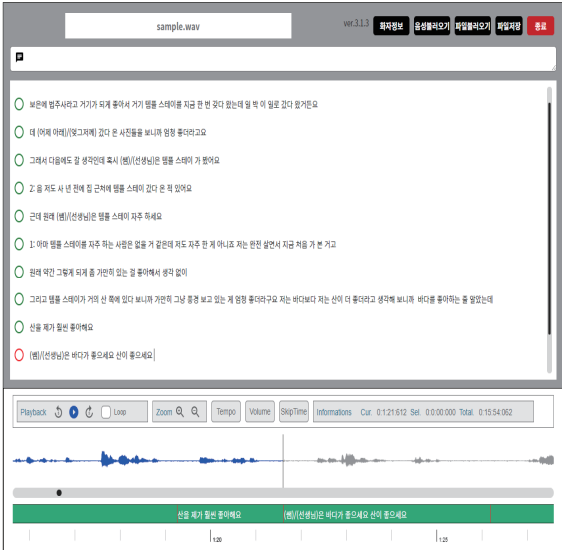
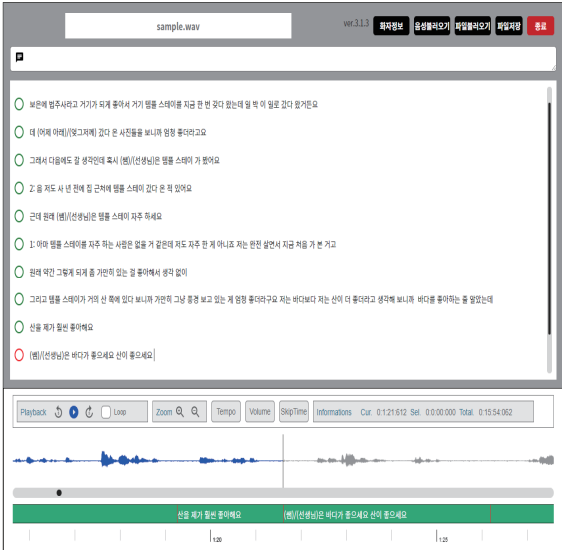
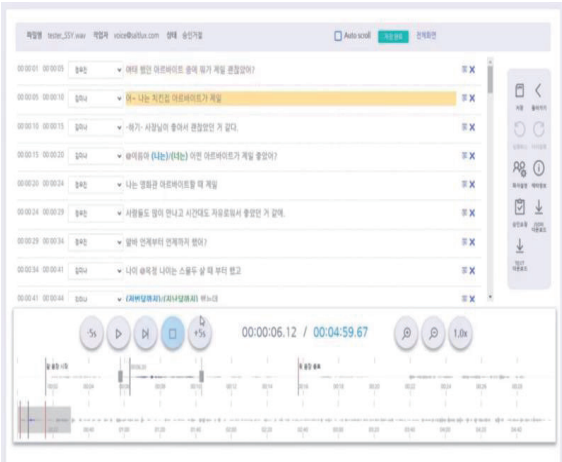
붙임

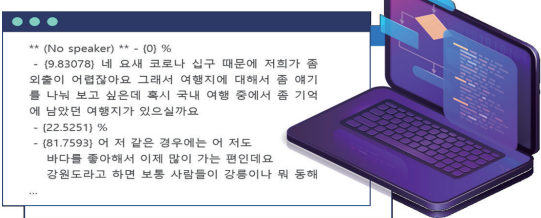
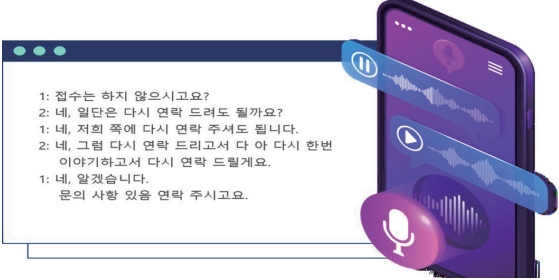
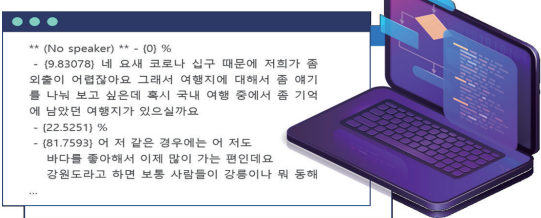
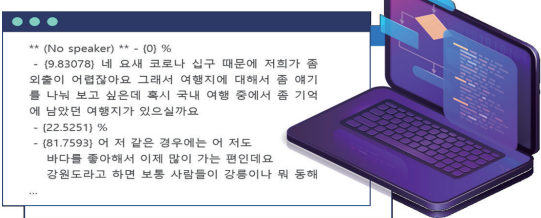
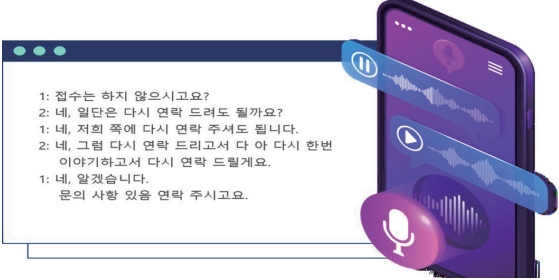

2. 데이터 라벨링 방법

1. 텍스트, 음성 데이터 라벨링

데이터 라벨링	설명							
전사 (Transcription)	<ul style="list-style-type: none"> • 말소리를 음성 문자로 옮겨 적는 방법 • 전사는 전사 방법, 전사 주체 및 전사 기관에 따라 분류할 수 있음 <ul style="list-style-type: none"> ① 전사 방법에 따른 분류 							
	<table border="1"> <thead> <tr> <th>전사방법</th><th>최소 기준치(선택)</th></tr> </thead> <tbody> <tr> <td>일반 전사</td><td> <ul style="list-style-type: none"> • 발음전사 • 사람이 말한 그대로 문자화하여 전사 • 말을 더듬거나 반복해도 따로 표기하지 않고 그대로 전사 • 전사방법이 까다롭지 않으며 작업시간도 비교적 짧음  <p> - (36.6605) 이런 부분도 다 하신 다음에 네 네 네 네 그 리고 거기 밑에 보면 전체 화면 종속 코드도 다 다운로드 한번 해 주시기 바랍니다 - (48.4776) 네 네 그러면 다시 어플 종료하고 재접속하시면 됩니다 고객님의 종속 코드 네 - (55.2836) 네 네 네 어떤 만기 일을 말씀하시는 고객님의 네 아 신용 만기 말씀이십니까 - (64.5461) 고객님의 이 부분은 제가 상담원에서 연결해 드려야 되는데 잠시만요 고객님의 </p> </td></tr> <tr> <td>이중 전사</td><td> <ul style="list-style-type: none"> • 말 그대로 한 단어 또는 어절에 대해 두 번 전사 • 발음 전사와 철자 전사(표준어 표기)를 병기 • (철자 전사)/(발음 전사)의 형태이나 표기 순서는 작업에 따라 변경 가능 • 방언 전사에도 적용  <p> - (432.446) (이케)/(이렇게) 서 그런 거를 챙겨서 주는 편인 가요 혹시 주신다면은 만들어서 주시는 편인가요 사서 주시는 편인가요? - (439.967) 2: 그런 기념일 같은 경우에는 - (444.311) 별로 (쫘)/(조금) 챙기는 편은 아닙니다. - (447.037) 어~ 나이가 지금 인제 오십 대 (()) 보니까 어~ 아이들이 오히려 인제 얼마 아빠한테 주는 음~ - (454.965) 그냥 이런 소소한 즐거움 어~ 그런 날이고 - (460.267) 저 같은 경우에는 인제 아이가 @육점 막내가 중학교 때 직접 </p> </td></tr> <tr> <td>화자 전사</td><td> <ul style="list-style-type: none"> • 음성 데이터에 등장하는 화자가 여러 명일 때 화자를 구분 • 화자와 소음에 대한 태깅 • 발화 내용은 전사하지 않음 </td></tr> </tbody> </table>	전사방법	최소 기준치(선택)	일반 전사	<ul style="list-style-type: none"> • 발음전사 • 사람이 말한 그대로 문자화하여 전사 • 말을 더듬거나 반복해도 따로 표기하지 않고 그대로 전사 • 전사방법이 까다롭지 않으며 작업시간도 비교적 짧음  <p> - (36.6605) 이런 부분도 다 하신 다음에 네 네 네 네 그 리고 거기 밑에 보면 전체 화면 종속 코드도 다 다운로드 한번 해 주시기 바랍니다 - (48.4776) 네 네 그러면 다시 어플 종료하고 재접속하시면 됩니다 고객님의 종속 코드 네 - (55.2836) 네 네 네 어떤 만기 일을 말씀하시는 고객님의 네 아 신용 만기 말씀이십니까 - (64.5461) 고객님의 이 부분은 제가 상담원에서 연결해 드려야 되는데 잠시만요 고객님의 </p>	이중 전사	<ul style="list-style-type: none"> • 말 그대로 한 단어 또는 어절에 대해 두 번 전사 • 발음 전사와 철자 전사(표준어 표기)를 병기 • (철자 전사)/(발음 전사)의 형태이나 표기 순서는 작업에 따라 변경 가능 • 방언 전사에도 적용  <p> - (432.446) (이케)/(이렇게) 서 그런 거를 챙겨서 주는 편인 가요 혹시 주신다면은 만들어서 주시는 편인가요 사서 주시는 편인가요? - (439.967) 2: 그런 기념일 같은 경우에는 - (444.311) 별로 (쫘)/(조금) 챙기는 편은 아닙니다. - (447.037) 어~ 나이가 지금 인제 오십 대 (()) 보니까 어~ 아이들이 오히려 인제 얼마 아빠한테 주는 음~ - (454.965) 그냥 이런 소소한 즐거움 어~ 그런 날이고 - (460.267) 저 같은 경우에는 인제 아이가 @육점 막내가 중학교 때 직접 </p>	화자 전사
전사방법	최소 기준치(선택)							
일반 전사	<ul style="list-style-type: none"> • 발음전사 • 사람이 말한 그대로 문자화하여 전사 • 말을 더듬거나 반복해도 따로 표기하지 않고 그대로 전사 • 전사방법이 까다롭지 않으며 작업시간도 비교적 짧음  <p> - (36.6605) 이런 부분도 다 하신 다음에 네 네 네 네 그 리고 거기 밑에 보면 전체 화면 종속 코드도 다 다운로드 한번 해 주시기 바랍니다 - (48.4776) 네 네 그러면 다시 어플 종료하고 재접속하시면 됩니다 고객님의 종속 코드 네 - (55.2836) 네 네 네 어떤 만기 일을 말씀하시는 고객님의 네 아 신용 만기 말씀이십니까 - (64.5461) 고객님의 이 부분은 제가 상담원에서 연결해 드려야 되는데 잠시만요 고객님의 </p>							
이중 전사	<ul style="list-style-type: none"> • 말 그대로 한 단어 또는 어절에 대해 두 번 전사 • 발음 전사와 철자 전사(표준어 표기)를 병기 • (철자 전사)/(발음 전사)의 형태이나 표기 순서는 작업에 따라 변경 가능 • 방언 전사에도 적용  <p> - (432.446) (이케)/(이렇게) 서 그런 거를 챙겨서 주는 편인 가요 혹시 주신다면은 만들어서 주시는 편인가요 사서 주시는 편인가요? - (439.967) 2: 그런 기념일 같은 경우에는 - (444.311) 별로 (쫘)/(조금) 챙기는 편은 아닙니다. - (447.037) 어~ 나이가 지금 인제 오십 대 (()) 보니까 어~ 아이들이 오히려 인제 얼마 아빠한테 주는 음~ - (454.965) 그냥 이런 소소한 즐거움 어~ 그런 날이고 - (460.267) 저 같은 경우에는 인제 아이가 @육점 막내가 중학교 때 직접 </p>							
화자 전사	<ul style="list-style-type: none"> • 음성 데이터에 등장하는 화자가 여러 명일 때 화자를 구분 • 화자와 소음에 대한 태깅 • 발화 내용은 전사하지 않음 							

데이터 라벨링	설명								
	<table border="1"> <thead> <tr> <th data-bbox="358 293 490 342">전사방법</th><th data-bbox="490 293 1138 342">최소 기준치(선택)</th></tr> </thead> <tbody> <tr> <td data-bbox="358 342 490 706"></td><td data-bbox="490 342 1138 706">  </td></tr> <tr> <td data-bbox="358 706 490 1179">배경음 및 화자 감정 태깅</td><td data-bbox="490 706 1138 1179"> <ul style="list-style-type: none"> • 드라마 등에서 배경음이 나오는 구간을 음성 싱크를 설정하고 태깅 • 화자의 발화 내용에 감정이 나타났을 경우 해당 구간만 음성 싱크를 설정하고 발화 내용을 전사 • 전사한 문장 끝에 화자의 감정을 태깅 • 감정 표기는 중립, 분노, 기쁨, 슬픔으로 주로 네 가지를 사용  </td></tr> <tr> <td data-bbox="358 1179 490 1667">방송 영상 자막 사전 제작</td><td data-bbox="490 1179 1138 1667"> <ul style="list-style-type: none"> • 청각 장애인의 시청권 보장을 위해 영화와 같은 프로그램의 내용을 자막으로 사전에 제작 • 등장 인물의 말뿐만 아니라 상황과 배경 음악에 대한 설명도 전사해야하므로 작업이 까다로운 편  </td></tr> </tbody> </table>	전사방법	최소 기준치(선택)			배경음 및 화자 감정 태깅	<ul style="list-style-type: none"> • 드라마 등에서 배경음이 나오는 구간을 음성 싱크를 설정하고 태깅 • 화자의 발화 내용에 감정이 나타났을 경우 해당 구간만 음성 싱크를 설정하고 발화 내용을 전사 • 전사한 문장 끝에 화자의 감정을 태깅 • 감정 표기는 중립, 분노, 기쁨, 슬픔으로 주로 네 가지를 사용 	방송 영상 자막 사전 제작	<ul style="list-style-type: none"> • 청각 장애인의 시청권 보장을 위해 영화와 같은 프로그램의 내용을 자막으로 사전에 제작 • 등장 인물의 말뿐만 아니라 상황과 배경 음악에 대한 설명도 전사해야하므로 작업이 까다로운 편 
전사방법	최소 기준치(선택)								
									
배경음 및 화자 감정 태깅	<ul style="list-style-type: none"> • 드라마 등에서 배경음이 나오는 구간을 음성 싱크를 설정하고 태깅 • 화자의 발화 내용에 감정이 나타났을 경우 해당 구간만 음성 싱크를 설정하고 발화 내용을 전사 • 전사한 문장 끝에 화자의 감정을 태깅 • 감정 표기는 중립, 분노, 기쁨, 슬픔으로 주로 네 가지를 사용 								
방송 영상 자막 사전 제작	<ul style="list-style-type: none"> • 청각 장애인의 시청권 보장을 위해 영화와 같은 프로그램의 내용을 자막으로 사전에 제작 • 등장 인물의 말뿐만 아니라 상황과 배경 음악에 대한 설명도 전사해야하므로 작업이 까다로운 편 								

데이터 라벨링	설명				
<p>② 전사 주체에 따른 분류</p>	<table border="1"> <thead> <tr> <th>전사주체</th><th>최소 기준치(선택)</th></tr> </thead> <tbody> <tr> <td>사람에 의한 전사</td><td> <ul style="list-style-type: none"> 음성 데이터를 가지고 처음부터 끝까지 사람이 작업을 진행하는 것으로 프로그램에 의해 전사했을 때 제대로 인식이 되지 않은 부분을 보다 명확히 알아들을 수 있음  </td></tr> </tbody> </table>	전사주체	최소 기준치(선택)	사람에 의한 전사	<ul style="list-style-type: none"> 음성 데이터를 가지고 처음부터 끝까지 사람이 작업을 진행하는 것으로 프로그램에 의해 전사했을 때 제대로 인식이 되지 않은 부분을 보다 명확히 알아들을 수 있음 
전사주체	최소 기준치(선택)				
사람에 의한 전사	<ul style="list-style-type: none"> 음성 데이터를 가지고 처음부터 끝까지 사람이 작업을 진행하는 것으로 프로그램에 의해 전사했을 때 제대로 인식이 되지 않은 부분을 보다 명확히 알아들을 수 있음 				
<p>STT (Speech To Text)</p>	<ul style="list-style-type: none"> 프로그램에 의해 하나의 음성 데이터를 전체 전사하는 것으로 사람이 전사할 때보다 속도가 빠름 				

데이터 라벨링	설명					
	③ 전사 기관에 따른 분류					
	<table><tr><th>전사기관</th><th>최소 기준치(선택)</th></tr><tr><td>연구 기관의 과제 전사</td><td><ul style="list-style-type: none">인공 지능 학습을 위한 연구 과제에서 요구하는 음성 데이터를 전사특정 조건의 음성 데이터를 필요로 하기 때문에 해당 조건에 맞는 음성 데이터를 수집하여 전사까지 진행하는 경우도 있음</td></tr><tr><td>기업의 콜센터 녹취 전사</td><td><ul style="list-style-type: none">기업에서 보유하고 있는 인공지능 엔진의 학습을 목적으로 함콜센터 상담원과 고객 간 통화가 녹음된 데이터를 전사기업에서 요구하는 부분에 따라 전사 방법은 조금씩 차이가 있음</td></tr></table>	전사기관	최소 기준치(선택)	연구 기관의 과제 전사	<ul style="list-style-type: none">인공 지능 학습을 위한 연구 과제에서 요구하는 음성 데이터를 전사특정 조건의 음성 데이터를 필요로 하기 때문에 해당 조건에 맞는 음성 데이터를 수집하여 전사까지 진행하는 경우도 있음 	기업의 콜센터 녹취 전사
전사기관	최소 기준치(선택)					
연구 기관의 과제 전사	<ul style="list-style-type: none">인공 지능 학습을 위한 연구 과제에서 요구하는 음성 데이터를 전사특정 조건의 음성 데이터를 필요로 하기 때문에 해당 조건에 맞는 음성 데이터를 수집하여 전사까지 진행하는 경우도 있음 					
기업의 콜센터 녹취 전사	<ul style="list-style-type: none">기업에서 보유하고 있는 인공지능 엔진의 학습을 목적으로 함콜센터 상담원과 고객 간 통화가 녹음된 데이터를 전사기업에서 요구하는 부분에 따라 전사 방법은 조금씩 차이가 있음 					
태깅 (Tagging)	<ul style="list-style-type: none">원본 지문(텍스트)에서 질문에 대한 답 생성 					

〈표 III-178〉 텍스트, 음성 데이터 라벨링 방법

2. 이미지, 영상 데이터 라벨링

데이터 라벨링	설명
<p>바운딩 박스 (Bounding Box)</p>	<ul style="list-style-type: none"> 이미지 혹은 영상 안 객체를 직사각형 모양의 박스 안에 포함되도록 그리는 라벨링 방법 데이터 라벨링 작업에서 가장 일반적으로 사용 객체 전체가 커버되도록 하며, 박스 안에 객체 이외의 여백을 최소화하도록 지정 (바운딩 박스 예시) <div data-bbox="368 469 1108 1097">  </div>
<p>키포인트 (Keypoint)</p>	<ul style="list-style-type: none"> 특정 지점(특징점)을 라벨링하는 방법 안면 인식을 통한 감정 분석과 같이 정밀하고 섬세한 작업을 요구하는 기술 객체의 중요 특징점을 지정하여 물체를 추적하고 인식할 수 있음 (키포인트 예시) <div data-bbox="368 1252 1103 1652">  </div>


데이터 라벨링	설명
<p>폴리곤 (Polygon)</p>	<ul style="list-style-type: none"> • 다각형 모양으로 객체의 가시 영역 외곽선을 따라 점을 찍어 그리는 라벨링 방법 • 객체 이외의 포함된 빈공간으로 인해 발생하는 오류에 대응할 수 있는 기능 • 사물의 테두리를 따라 그리는 것을 통해 여백 없이 정확히 물체만을 인식하기 위해 사용 <p>〈폴리곤 예시〉</p> <div>   </div>
<p>폴리라인 (Polyline)</p>	<ul style="list-style-type: none"> • 여러개의 점을 가진 선을 활용하여 특정 영역을 라벨링하는 방법 • 인도와 차선, 경계 등의 선형 데이터들을 구분하고 인식시키기 위해서 사용 <p>〈폴리라인 예시〉</p> <div>   </div>
<p>시멘틱 세그멘테이션 (Semantic segmentation)</p>	<ul style="list-style-type: none"> • 이미지의 모든 픽셀에 클래스를 부여하는 라벨링 방법 • 물체, 배경 등을 의미적으로 분할하여 자율주행, 의료 영상 분석 기술 개발 등에 사용

데이터 라벨링	설명
	<p>〈시멘틱 세그멘테이션 예시〉</p> 
<p>인스턴스 세그멘테이션 (Instance segmentation)</p>	<ul style="list-style-type: none"> • 이미지의 모든 픽셀이 같은 클래스여도 다른 인스턴스를 부여하는 라벨링 방법 • 객체인식 시 중첩된 경우 분할하는 기술 개발 등에 사용 <p>〈인스턴스 세그멘테이션 예시〉</p>  <p>※ 출처 : MS COCO 데이터셋</p>
<p>팬옵틱 세그멘테이션 (panoptic segmentation)</p>	<ul style="list-style-type: none"> • 시멘틱 세그멘테이션과 인스턴스 세그멘테이션을 결합한 방식으로 셀 수 있는 클래스(예, 차, 사람 등)를 Thing 클래스라고 하며, 셀 수 없는 클래스(예, 하늘, 길 등)는 Stuff 클래스라고 하며, Thing 클래스에 대해서는 인스턴스 세그멘테이션, Stuff 클래스에 대해서는 시멘틱 세그멘테이션 작성을 수행

데이터 라벨링	설명
	<p>〈팬옵틱 세그멘테이션 예시〉</p>  <p>※ 출처 : MS COCO 데이터셋</p>

〈표 III-179〉 이미지, 영상 데이터 라벨링 방법

3. 3D 데이터 라벨링

데이터 라벨링	설명
큐보이드 (Cuboid)	<ul style="list-style-type: none"> • 2D로 작업할 수 없는 자동차, 건물 등 입체적인 객체를 직육면체(3D Bounding Box)로 라벨링하는 방식 • 3D 객체 라벨링 작업에서 사용 • 객체 전체가 커버되도록 하며, 직육면체(3D Bounding Box) 직육면체 안에 객체가 들어오도록 지정 <p>출처 : 아이티데일리(http://www.itdaily.kr), 아주일보(https://www.ajunews.com)</p> <p>〈큐보이드 예시〉</p> 

〈표 III-180〉 3D 데이터 라벨링 방법

부록 3 | 개인정보보호 가이드

1 개요

1.1 배경 및 목적

- 인공지능 학습용 데이터 구축 사업을 수행함에 따라 발생 가능한 개인정보 침해 및 유출 가능성 사전 예방·제거

구분	내용
배경	<ul style="list-style-type: none"> • 인공지능 기술이 적용된 새로운 서비스가 도입·확산됨에 따라 개인정보 침해 등 다양한 사회적 문제 발생 우려 • 인공지능 학습용 데이터 구축 시 개인정보 처리의 안전성과 신뢰성 확보 필요 • 크고 작은 개인정보 유출 사고가 지속되어 개인정보보호 정책을 강화해야 한다는 사회적 요구 지속 • 다양한 데이터 활용을 필요로 하는 새로운 산업과 기술발전으로 개인정보 침해 위험도 증가 추세
목적	<ul style="list-style-type: none"> • 인공지능 학습용 데이터 구축 사업 생애주기(Life-Cycle) 과정에서 개인정보의 안전한 활용과 보호를 위한 관리체계 마련 • 인공지능 학습용 데이터 구축 사업의 개인정보보호 관련 지원 및 점검 확대 • 사업에 참여하는 사업수행기관(주관기관, 참여기관) 및 수요기관의 개인정보 보호를 위한 역할과 책임 강화 • 개인정보의 적법·안전한 처리와 침해예방 등을 위한 「개인정보 보호법」 준수 가이드

〈표 III-181〉 개인정보보호의 배경 및 목적

1.2 인공지능 관련 개인정보 처리 특성 및 기본방향

구분	인공지능 학습용 데이터 처리 특성	인공지능 학습용 데이터 처리 기본방향
대규모 데이터의 처리	AI 학습데이터는 규모가 크고 다양한 개인정보와 민감한 사생활에 관한 정보가 포함될 가능성이 높음. 또한 서비스 운영 중에도 이러한 데이터를 지속적으로 활용하려는 수요가 많음	정보주체의 동의, 가명처리 등 적법한 방법으로 개인정보를 수집하여 예측 가능하고 허용된 목적 범위에서 이용하고, 안전하게 관리하는 것이 중요
복잡성 불투명성	AI 학습데이터는 구축하는 과정에서 개인정보 처리방식은 매우 복잡하고 이용자는 자신의 개인정보가 어떻게 처리되는지 알기 어려워 정보주체의 권리행사가 제한될 우려	개인정보 처리내역 등을 투명하게 공개하여 정보주체가 자신의 개인정보 처리에 대한 권리를 행사할 수 있도록 이용자 참여에 대한 보장이 중요
자동화 불확실성	일반적으로 AI 모델은 지식·확률기반의 추론방식을 이용하여 데이터를 분석·처리함. 이를 활용한 자동화된 서비스를 개발·운영하는 과정에서 데이터 처리 결과를 예측하기 어려워 사생활 침해, 사회적 차별·편향 등의 문제가 발생할 수 있음	프라이버시가 보호될 수 있도록 개인정보 처리 전반에 대한 책임 있는 관리가 중요하고, 이용자가 차별받지 않도록 개인정보의 처리 결과의 공정함을 고려할 필요

※ 출처 : 인공지능(AI) 개인정보보호 자율점검표(2021. 05, 개인정보보호위원회)

〈표 Ⅲ-182〉 인공지능 학습용 데이터의 처리 특성 및 기본방향

1.3 인공지능 관련 개인정보보호 원칙 법 및 기준

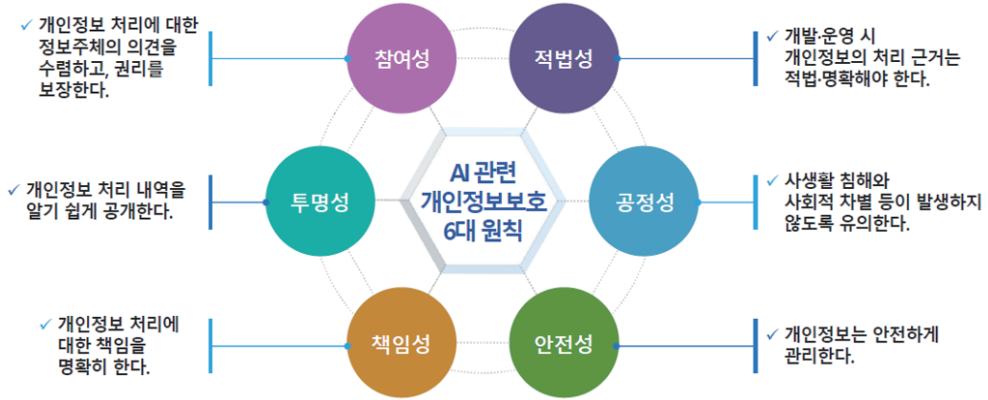
- 인공지능 기술·서비스의 개인정보 처리 특성을 고려한 개인정보보호를 위해서는 현행 법령에 따른 개인정보보호 의무사항 준수뿐 아니라, 자율적 보호 활동 및 윤리적 이슈 대응이 중요

「개인정보 보호법」상 개인정보보호 원칙(§3)	개인정보보호 중심설계(PbD)	AI 윤리기준
① 필요 최소한의 정보 수집 및 처리 목적의 명확화 ② 목적 범위 내에서 적법하게 처리, 목적 외 활용금지 ③ 처리목적 내에서 정확성·완전성·최신성 보장 ④ 권리침해 가능성 등을 고려하여 안전하게 관리 ⑤ 개인정보 처리 내역의 공개 및 정보주체의 권리보장 ⑥ 사생활 침해를 최소화하는 방법으로 처리 ⑦ 익명 및 가명 처리의 원칙 ⑧ 개인정보처리자의 책임준수·신뢰 확보 노력	① 사후 조치가 아닌 사전 예방 ② 초기 설정부터 프라이버시 보호 조치 ③ 프라이버시 보호를 내재한 설계 ④ 프라이버시 보호와 사업기능의 균형 ⑤ 개인정보 생애주기 전체에 대한 보호 ⑥ 개인정보 처리 과정에 대한 투명성 유지 ⑦ 이용자 프라이버시 존중	① 인권보장 ⑥ 연대성 ② 프라이버시 보호 ⑦ 데이터 권리 ③ 다양성 존중 ⑧ 책임성 ④ 침해금지 ⑨ 안전성 ⑤ 공공성 ⑩ 투명성

※ 출처 : 인공지능(AI) 개인정보보호 자율점검표(2021. 05, 개인정보보호위원회)

〔그림 Ⅲ-11〕 개인정보보호 원칙 법 및 기준

1.4 인공지능 관련 개인정보보호 6대 원칙



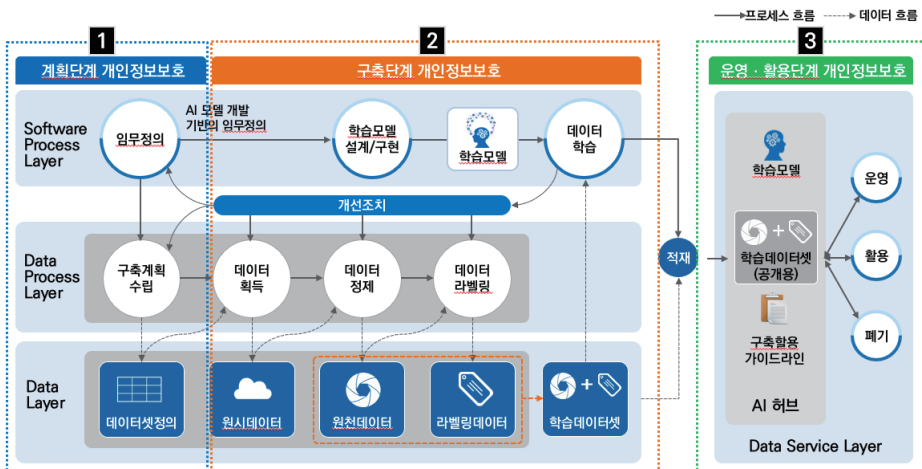
※ 출처 : 인공지능(AI) 개인정보보호 자율점검표(2021. 05, 개인정보보호위원회)

[그림 III-12] 개인정보보호 6대 원칙

2 생애주기별 개인정보보호 가이드라인

2.1 개요

- 인공지능 학습용 데이터의 생애주기별(계획, 구축, 운영·활용 단계) 개인정보보호에 대한 가이드라인 구축



[그림 III-13] 인공지능 학습용 데이터 생애주기

● 가이드라인 구축

단계	내용
계획	<ul style="list-style-type: none"> 사업수행계획서 포함 사항 <ul style="list-style-type: none"> 데이터 구축에 개인정보 포함 여부 및 필수적인 개인정보의 구체적인 사항 개인정보를 활용하는 경우, 개인정보 보호법상의 의무·권고 사항에 대해 사전점검 및 법률 자문 실시* 개인정보보호 및 관리를 위한 이행·점검 및 조치 방안** 수행기관(주관기관, 참여기관) 담당자는 사업착수단계에서 사업관리기관에서 시행하는 개인정보보호 교육 이수 개인정보 보호법의 심각한 위반사항 발생 시 관련 법령 및 규정에 따라 협약 해지, 사업 참여 제한 등의 불이익을 받을 수 있음
구축	<ul style="list-style-type: none"> 데이터 획득(수집) <ul style="list-style-type: none"> 개인정보가 포함된 데이터를 수집하는 경우, 법률에 근거하거나 정보주체의 동의 필수 만14세 미만 아동의 개인정보를 수집하는 경우 법정대리인 동의 필요 정보주체 이외 개인정보 수집에 대한 수집 출처 및 처리목적 고지 데이터 정제/라벨링 <ul style="list-style-type: none"> 데이터의 개인정보(얼굴, 번호판 등), 국가 보안 사항(공간정보, 위치정보 등) 등이 포함된 경우 개인정보·민감정보 비식별화 조치 음성 및 텍스트 데이터의 경우 비속어 및 개인정보, 민감정보 등의 비식별화 처리 개인정보보호 자율 점검 실시 및 보고
운영·활용	<ul style="list-style-type: none"> 과제 수행 산출물(원시데이터 포함) <ul style="list-style-type: none"> 과제 종료 후 5년간 보존 AI 개발·운영종료 등으로 개인정보가 불필요하게 되었을 때에는 지체 없이 파기 사업 종료 이후 인공지능 사업 결과물(모델)에 개인정보 포함 여부 확인 데이터 개방 이후에도 개인정보보호에 대한 주기적 검증 실시

* 인공지능 개인정보보호 자율점검표(2021. 05, 개인정보보호위원회)에 따른 자가 점검을 실시하고 그 결과를 제출 및 과제조정위원회에서 검토

** 개인정보 비식별화를 시행하는 경우 '가명정보 처리 가이드라인(개인정보보호위원회)'을 준수

〈표 Ⅲ-183〉 생애주기별 개인정보보호 절차 및 가이드

2.1 계획 단계 개인정보보호 가이드

● 계획 단계 개인정보보호 검토사항

- 인공지능 학습용 데이터 구축 기획·설계 단계부터 PbD 원칙³⁾에 따라 개인정보 침해 위험을 분석·제거

3) Privacy by Design(PbD) : 제품·서비스 개발 시 기획 단계부터 개인정보 처리의 전체 생애주기에 걸쳐 이용자의 프라이버시를 고려한 기술·정책을 설계에 반영하는 것을 의미하며, 국제적으로 광범위하게 통용되는 개인정보보호 원칙

- 「개인정보 보호법」 등 관련 법에 따른 준수사항, 사생활 보호 방안 등을 AI 학습용 데이터 구축 계획에 반영
- 사업수행계획서에 아래의 사항을 반드시 포함
 - 데이터 구축에 개인정보 포함 여부 및 필수적인 개인정보의 구체적인 사항
 - 개인정보를 활용하는 경우, 개인정보 보호법 상의 의무·권고사항에 대해 사전점검 및 법률 자문 실시 방안
 - 개인정보보호 및 관리를 위한 이행·점검 및 조치 방안

● 계획 단계 개인정보보호 참고사항

- PbD 원칙 적용
 - 수집 이용하려는 전체 데이터 현황 파악
 - 개인정보 항목과 유형(식별자, 속성정보 등) 분석
 - 개인정보 항목별 수집 근거(동의 등)와 활용 방식(가명처리 등) 결정
 - 개인정보 처리 흐름도 작성을 사전에 수행하고, 개인정보 처리에 따른 위험성 침해요인 분석, 대안 마련 등을 순차적으로 수행하여, 인공지능 학습용 데이터 구축 전반의 개인정보보호 관련 사항을 검토
- 사전 위험분석, 안전성 점검, 법령 해석 등을 위해 외부 전문가에게 자문을 받거나 전문기관에서 질의 가능
 - ※ 한국인터넷진흥원(개인정보침해신고센터, 국번없이 118) 등
- 「자동처리 되는 개인정보보호 가이드라인(Privacy by Design 적용사례를 중심으로, '20. 12)」 참고
 - 인공지능 학습용 데이터 구축에 개인정보 포함 여부 및 필수적인 개인정보의 구체적인 사항

2.2 구축 단계 개인정보보호 가이드

2.2.1 데이터 획득(수집) 단계 개인정보보호 가이드

● 개인정보 수집 동의

- 개인정보는 필수·선택 동의로 구분하여 목적에 필요한 최소한의 개인정보를 수집
- 개인정보 처리에 대해 정보주체의 동의를 받을 때에는 각각의 동의 사항을 구분하여 알리고 수집·이용 목적, 수집·이용하려는 항목, 기간 등에 관한 내용을 명확히 알려야함
- 개인정보 수집·이용 동의를 받았더라도 목적 외 이용인 경우에는 정보주체의 추가적 동의 필요
- 만 14세 미만 아동의 개인정보 수집 시 법정대리인의 동의

● 개인정보 수집 동의 참고사항

- ‘인공지능 학습용 데이터 구축’ 개인정보의 수집·이용 동의를 받는 경우 정보주체가 충분히 이해·예측할 수 있도록 ‘인공지능 학습용 데이터’ 의미, 개인정보 이용 범위·목적 등을 구체적으로 알려야 함

단계	내용
개인정보 수집항목	“OO SNS 대화 정보”와 같이 이용 항목을 구체적으로 작성
수집·이용 목적	“OO 서비스의 챗봇 알고리즘 개발(학습용)”과 같이 구체적으로 작성
보유·이용 기간	OO 알고리즘 개발을 위해 O 개월 보관·이용

- 다만, 과학적 연구, 통계작성, 공익적 기록보존 등 목적으로 가명처리하여 필요 최소한의 기간동안 보관·이용 가능
- 동의를 거부할 권리가 있음을 알리고, 거부 시 불이익이 있는 경우 그 내용을 기재

〈표 III-184〉 인공지능 학습용 데이터 수집·이용 동의

● 민감정보와 고유식별정보

구분	내용
민감정보	<ul style="list-style-type: none"> • 민감정보(법 23조, 영 제18조) : ① 사상·신념, ② 노동조합·정당의 가입·탈퇴, ③ 정치적 견해, ④ 건강, ⑤ 성생활 등에 관한 정보, ⑥ 유전정보, ⑦ 범죄경력자료에 해당하는 정보, ⑧ 바이오정보(특정정보), ⑨ 인종이나 민족에 관한 정보 등 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보
고유식별정보	<ul style="list-style-type: none"> • 고유식별정보(법 24조, 영 제19조) : ① 여권번호, ② 운전면허번호, ③ 외국인등록번호 등 • 단, 주민등록번호는 법률·대통령령 또는 헌법기관 규칙에 구체적 근거가 없는 경우 처리가 불가능함(법 24조의2)

〈표 III-185〉 민감정보와 고유식별정보

● 정보주체 동의 없이 개인정보 수집

구분	내용
개인정보 처리자	<ul style="list-style-type: none"> • 법률에 특별한 규정 또는 법령상 의무 준수를 위해 불가피한 경우(법 제5①2) • 공공기관이 법령 등에서 정하는 소관 업무의 수행을 위하여 불가피한 경우(법 제15①3) • 정보 주체와의 계약의 체결·이행을 위하여 불가피하게 필요한 경우(법 제15①4) • 명백히 정보주체 또는 제3자의 급박한 생명, 신체, 재산의 이익을 위하여 필요하다고 인정되는 경우(법 제15①5) • 개인정보 처리자의 정당한 이익을 달성하는 데 필요한 경우로서 명백하게 정보주체의 권리보다 우선하는 경우(법 제15①6)
정보통신 서비스 제공자	<ul style="list-style-type: none"> • 정보통신서비스의 제공에 관한 계약을 이행하는 데 필요한 개인정보로서 경제적·기술적인 사유로 통상적인 동의를 받는 것이 뚜렷하게 곤란한 경우(법 제39의3②1) • 정보통신서비스의 제공에 따른 요금 정산을 위하여 필요한 경우(법 제39의3②2) • 다른 법률에 특별한 규정이 있는 경우(법 제39의3②3)

※ 출처 : 인공지능(AI) 개인정보보호 자율점검표(2021. 05, 개인정보보호위원회)

〈표 III-186〉 정보주체 동의 없이 개인정보 수집이 가능한 경우

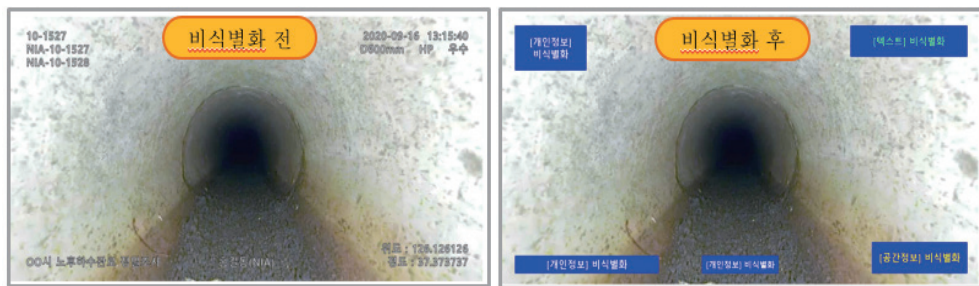
● 공간정보 수집

공간정보	등급	분류기준
항공사진	비공개	<ul style="list-style-type: none"> 일반인 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함)이 노출된 3차원 위성자료
	공개제한	<ul style="list-style-type: none"> 일반인 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함)이 삭제된 흔적이 남아있는 사진 및 영상, 3차원 입체자료 2차원 좌표가 포함된 해상도 30m보다 정밀한 자료 3차원 좌표가 포함된 해상도 90m보다 정밀한 자료
	공개	<ul style="list-style-type: none"> “비공개” 및 “공개제한” 대상 이외의 항공사진 및 영상, 3차원 입체자료(인터넷·내비게이션·휴대폰에는 좌표 표시 불가) 해상도 25cm보다 정밀한 항공사진은 건물·토지의 소유자와 법 제2조 4항의 관리 기관 또는 관리 기관의 장이 승인한 경우에 한하여 제공 또는 판매하고, 인적사항 및 사진 내용을 기록하여야 함 ※ 단, 올림픽 등 국제행사 지역은 관계기관과 협의를 거쳐, 행사 동안 해상도 25cm보다 정밀한 항공사진을 일반인에게 제공 또는 판매할 경우 기록을 생략할 수 있음
위성영상	비공개	<ul style="list-style-type: none"> 일반인 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함)이 노출된 3차원 위성자료
	공개제한	<ul style="list-style-type: none"> 정밀 보정된 2차원 좌표가 포함된 해상도 30m 보다 정밀한 자료 일반인 출입이 통제되는 국가보안시설과 군사시설(휴전선 접경 지역 내 시설 포함)이 노출된 해상도 4m보다 정밀한 자료 3차원 좌표가 포함된 해상도 90m보다 정밀한 자료
	공개	<ul style="list-style-type: none"> “비공개” 및 “공개제한” 대상 이외의 위성영상 및 3차원 위성자료 촬영 당시 위성 자세 정보가 포함된 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함) 이외 지역의 자료(인터넷·내비게이션·휴대폰에는 좌표 표시 불가) ※ 단, 해상도 25m보다 정밀한 위성영상 제공 또는 판매 시 인적사항 및 사진내용 기록유지
전자지도	비공개	<ul style="list-style-type: none"> 축척에 관계없이 일반인의 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경지역내 시설 포함)이 포함된 지도
	공개제한	<ul style="list-style-type: none"> 군사지도 전력·통신·가스 등 공공의 이익 및 안전과 밀접한 관계가 있는 국가 기간 시설이 포함된 지도 ※ 단, 항공기·선박의 안전항행 등에 필요한 전력·통신·가스 등 국가 기간시설은 보안성 검토를 거쳐 지도에 표기 1:1,000 축척 이상 지도의 등고선과 표고점
	공개	<ul style="list-style-type: none"> “비공개” 및 “공개제한” 대상 이외의 지도 인터넷·내비게이션·휴대폰 등을 통해 좌표와 1:5,000 축척 이상 지도의 등고선·표고값 표시 불가
해양 공간 정보	비공개	<ul style="list-style-type: none"> 접경해역 수심 자료 좌표가 표기되어 있고 국가보안시설 및 군사시설 인근 해역 또는 접경 해역 내 해상도 30m보다 정밀한 해저 영상자료

공간정보	등급	분류기준
	공개제한	<ul style="list-style-type: none"> 좌표가 포함된 해상도 120m보다 정밀한 수심 자료 <ul style="list-style-type: none"> ※ 단, 공개제한 대상 수심이라도 항해안전, 해양레저, 해양공간정보 산업화 등에 필수적인 해역의 수심은 관계기관 협의 후 공개 좌표가 포함된 해상도 90m보다 정밀한 해저 영상자료
	공개	<ul style="list-style-type: none"> “비공개” 및 “공개제한” 대상 이외의 수심 및 해저 영상자료 해도, 전자해도 등
기타 공간정보	비공개	<ul style="list-style-type: none"> 일반인의 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함)이 노출된 3차원 공간정보 <ul style="list-style-type: none"> - 국가보안목표시설로 지정된 정수장이 표시된 상하수시스템, 도로명주소 기본도 일반인의 출입이 통제되는 국가보안시설 및 군사시설(휴전선 접경 지역 내 시설 포함)의 명칭 및 속성자료
	공개제한	<ul style="list-style-type: none"> 공공의 이익 및 안전과 밀접한 관계가 있는 국가기간시설의 명칭 및 속성자료 해상도가 90m보다 정밀하고 3차원 좌표가 포함된 3차원 공간정보 <ul style="list-style-type: none"> ※ 단, 해상도 90m보다 정밀하고 3차원 좌표가 포함된 공간정보 중 도로 지역은 보안성 검토를 거쳐 국가안보상 위해 요인이 없는 경우 공개
	공개	<ul style="list-style-type: none"> 좌표가 없는 일반지역 3차원 영상자료 3차원 좌표가 있고 해상도가 90m보다 낮은 입체영상자료 <ul style="list-style-type: none"> - 토양·지질·지반도, 도시·도로 건설계획도 등 도로명사업에 의하여 작성된 자료 중 보안관리 대상 시설을 제외한 도로명주소 안내도

※ 참고 : 행정안전부 공간정보 보안관리규정(개정 2020. 12, 행정안전부훈령 제 173호)

〈표 Ⅲ-187〉 공간정보 수집(예시)



[그림 Ⅲ-14] 공간정보 비식별화 예시

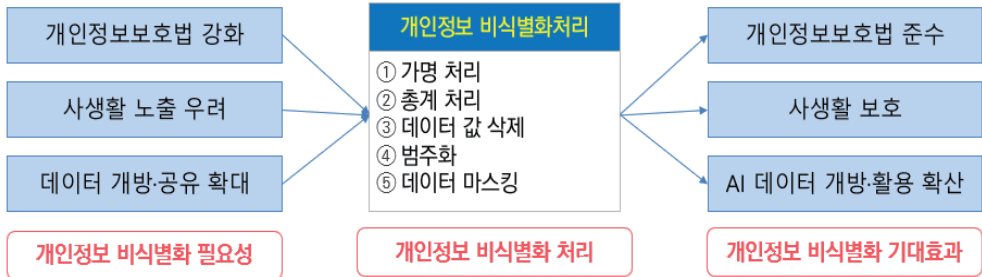
2.2.2 데이터 정제/라벨링 단계 개인정보보호 가이드

● 개인정보 비식별화

- 개인정보 비식별화 조치

- 개인정보에서 개인식별 요소를 제거하여 특정 개인을 알아볼 수 없는 형태로 만드는 조치
- 식별자 삭제 등 일련의 조치를 통한 식별방지와 프라이버시 모델 기반 추론 방지

- 정보에 포함되어 있는 개인정보의 일부 또는 전부를 삭제하거나 다른 정보로 대체함으로써 다른 정보와 결합하여도 특정 개인을 식별하기 어렵도록 하는 일련의 조치



※ 출처 : 개인정보 비식별 조치 가이드라인 (비식별 조치 기준 및 지원·관리체계 안내), 2016. 06 관계부처 합동)

[그림 Ⅲ-15] 개인정보 비식별화 정의

● 개인정보 비식별 조치 및 사후관리 절차

- **(사전검토)** 개인정보에 해당하는지 여부를 검토 후, 개인정보가 아닌 것이 명백한 경우 법적 규제 없이 자유롭게 활용
- **(비식별 조치)** 정보 집합물(데이터셋)에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용, 개인을 알아볼 수 없도록 하는 조치
- **(적정성 평가)** 다른 정보와 쉽게 결합하여 개인을 식별할 수 있는지를 「비식별 조치 적정성 평가단」을 통해 평가
- **(사후관리)** 비식별 정보 안전조치, 재식별 가능성 모니터링 등 비식별 정보 활용 과정에서 재식별 방지를 위해 필요한 조치 수행



※ 출처 : 개인정보 비식별 조치 가이드라인 (비식별 조치 기준 및 지원·관리체계 안내), 2016. 06 관계부처 합동)

[그림 Ⅲ-16] 개인정보 비식별화 조치 및 사후 절차

가명처리

- 인공지능 학습용 데이터의 가명처리⁴⁾ 유의사항

- 인공지능 학습용 데이터를 정보주체의 동의 없이 과학적 연구 등 목적으로 개인정보를 가명처리하여 활용하려는 경우, 가명처리의 적정성, 사생활 침해 가능성에 대한 검토 강화 필요
- 이를 위해 데이터 유형에 따라 현 기술 수준에서 활용 가능한 가명처리 방법을 적용하여 학습데이터를 안전하게 가명처리하고, 적정성 검토 후 문제 발견 시 추가 가명처리하여 정보주체의 권리가 침해되지 않도록 해야 함
→ (예시) SNS 대화 데이터의 경우 대화 발화자의 식별정보⁵⁾ 외에도 대화내용에 포함된 특정 개인의 식별가능정보 또는 사생활 침해 우려 정보도 가명처리 필요

- 인공지능 학습용 데이터의 가명정보의 공개제한

- 가명정보를 불특정 다수에게 공개하는 경우에는 공개의 목적이 통계작성, 과학적 연구, 공익적 기록보존 등의 목적에 해당하는지가 불분명하고, 불특정 다수 중 누군가는 공개하는 정보와 결합하여 특정 개인을 알아볼 수 있는 정보를 가지고 있을 수 있어, 가명정보의 공개는 사실상 제한됨
→ 불특정 제3자(공개 등)에게 제공하는 경우 익명정보로 처리하는 것을 원칙으로 함
→ 가명정보를 특정 제3자에게 제공하는 경우 가명정보 처리목적의 합법성, 재식별 금지 등 책임은 제3자에게 있으며, 이에 대한 별도 계약을 체결할 수 있음

개인정보	가명정보	익명정보
 홍길동 남자 801010-1234567 010-1234-1234 서울특별시 송파구 송파동 백제고분로 41길 40 honggildong@gmail.com 통신요금 12만3천원	 가명정보 *** 남자 80년생 010-***-**** 서울특별시 송파구 *****@gmail.com 통신요금 12만3천원	 익명정보 *** 남자 40대 ***** 서울특별시 *****@***** 통신요금 12만3천원
정보주체의 이익, 안전조치(암호화)여부등에 따라 동의없이 활용가능	통계작성, 과학적 연구, 공익적 기록보존 등에 동의없이 활용가능	언제든 동의없이 활용 가능 다른 정보와 결합해도 개인을 알아 볼 수 없는 상태

※ 출처 : https://blog.testworks.co.kr/portrait-of-the-data-age_how-much-is-your-face/

[그림 III-17] 개인정보/가명정보/익명정보 예시

텍스트 데이터 개인정보 정제 기준

- 데이터 라벨링에 포함하지 않아야 할 개인정보 등을 필터링하는 정제 기준을 마련

4) 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말함

5) 식별정보, 식별가능정보의 예시 등 자세한 사항은 「가명정보 처리 가이드라인」 (P.17) 참고

범주	항목	변환	예시
이름	실명, 특수 애칭, 별명, 대화명, 필명	#@이름#	(수정 전) 소연아 너무 고마워 (수정 후) #@이름# 너무 고마워
	일반 애칭, 별명	불필요	자기야, 여보 등
	공인 실명		김연아, 빌 게이츠 등
온라인	아이디, 이메일 주소, URL	#@계정#	(수정 전) sample@sample.com 으로 보내 (수정 후) #@계정#으로 보내
각종 번호	고유 식별 번호 (주민번호, 학번, 사번 등)	#@신원#	(수정 전) 응 학번은 200101-1234567 (수정 후) 응 학번은 #@신원#
	전화번호	#@전번#	(수정 전) 언니 번호 010-1234-56780이야 (수정 후) 언니 번호 #@전번#이야
	금융 번호(계좌, 카드번호 등)	#@금융#	(수정 전) 신한 110-234-45-67890 (수정 후) #@금융#
	일련번호, 비밀번호 등	#@번호#	(수정 전) 사업자등록번호 123-45-67890 (수정 후) 사업자등록번호 #@번호#
장소	상세 주소(동 이하), 거주 아파트 및 건물명	#@주소#	(수정 전) 배송지는 서구 연희동 123 로요 (수정 후) 배송지는 서구 #@주소# 로요
	거주지 익명 (지하철역, 기차역 등), 상호명	불필요	도곡역 3번출구로 오세요 롯데리아에서 만날래?
출신 및 소속	출신 및 소속 학교, 직장, 부대	#@소속#	(수정 전) 한국대학교에 재학중입니다 (수정 후) #@소속#에 재학중입니다

〈표 Ⅲ-188〉 정제 기준

● 음성데이터 개인정보 익명처리

- 원천데이터에 포함된 개인정보는 라벨링 대상에서 제외하거나, 익명처리 등 비식별화를 통해 개인정보 비식별화 처리
- 개인정보 활용 및 3자 제공 동의를 받은 경우엔 동의 범위 내에서 개인정보를 라벨링 데이터로 활용할 수 있음

개인정보에 대한 라벨링 익명처리 예시

- 대화자들의 신분 보장을 위해 이름, 주민등록번호, 카드 번호, 전화 번호 등 개인 정보와 관련된 사항은 그대로 전사하지 않고 아래와 같이 영문 기호로 적는다.

- 사람 이름은 '&name&'로 전사한다. 정치인, 연예인 등 유명인 이름은 그대로 전사한다.

(발화) 저는 김철수입니다.

(전사) 저는 &name&입니다.

- 여러 이름이 나올 때는 '&name1&', '&name2&' 등 name 뒤에 숫자를 붙여 구별한다.

(발화) 그때 철수랑 민수랑 너랑 나랑 갔잖아. 철수도 알고 있지?

(전사) 그때 &name1&이랑 &name2&이랑 너랑 나랑 갔잖아. &name1&도 알고 있지?

- 집 주소 등 개인정보에 해당되는 주소는 동 단위까지만 전사하고, 그 이하 구체 적인 주소는 영문 기호로 적는다.

(발화) 홍제동 한양아파트 205동 304호로 배달해주세요.

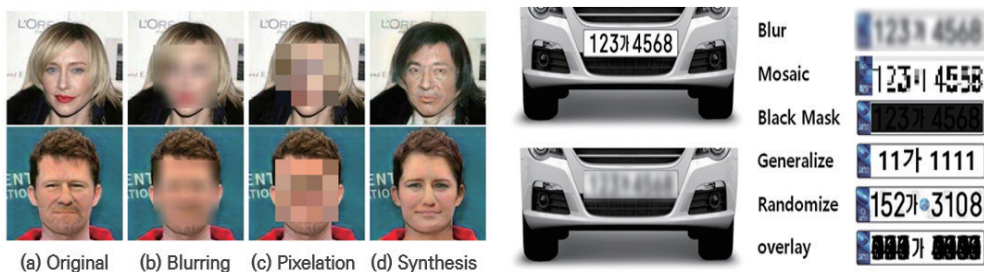
(전사) 홍제동 &address&로 배달해주세요.

※ 출처 : 인공지능 학습용 데이터셋 구축 안내서(2021. 02, 한국지능정보사회진흥원)

〔그림 Ⅲ-18〕 음성데이터 개인정보 익명처리 예시

● 이미지 데이터 개인정보 비식별화 처리

- 비식별화 범위 : 인물 얼굴, 자동차 번호판
- 개인정보에 해당하는 얼굴 부분, 자동차 번호판 흐림 효과(blur)를 통해 비식별화



※ 출처 : https://blog.testworks.co.kr/portrait-of-the-data-age_how-much-is-your-face/

[그림 III-19] 이미지(인물 얼굴 및 자동차 번호판) 비식별화 예시

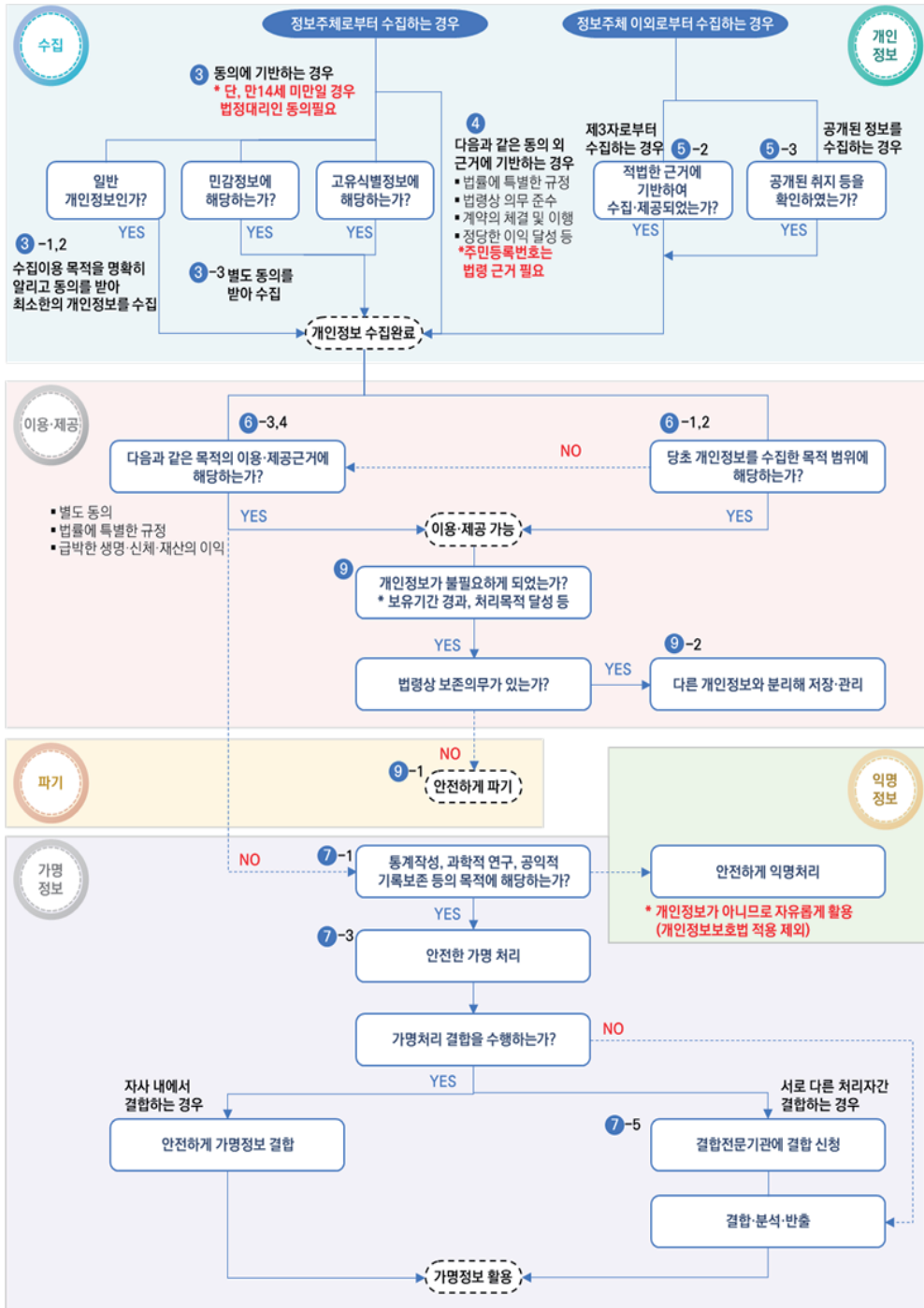
2.3 운영·활용 단계 개인정보보호 가이드

- 원시데이터를 포함한 과제 수행 산출물은 과제 종료 후 5년간 보존
- 인공지능 개발 및 운영종료 등으로 개인정보가 불필요하게 되었을 때에는 지체없이 파기
- 데이터 운영·활용 단계 개인정보보호 가이드
 - 인공지능 모델 개발이 완료되면 반복적인 테스트를 통해 사회적 차별·편향 등 이슈 발생 가능성 등을 검토하고, 발견된 문제점은 서비스를 출시하기 이전에 개선해야 함
 - 사업 종료 이후 인공지능 학습용 데이터 구축사업 결과물에 개인정보가 포함되어 있는지 등을 확인
 - 데이터를 개방하는 경우에는 개방 이후에도 개인정보탐지 SW 등을 활용한 주기적 검증 등 실시
- 기타 보안관리 강화
 - (물리적 보안) 개인정보를 무단으로 반출하지 못하도록 입·출입통제, 작업 승인, 개발 PC 분리, 작업 PC 반출 불가, USB 봉인 등
 - (기술적 보안) 권한, 접근제어, 암호화, 네트워크, 콘텐츠 보안 등 기술적 보안
 - (관리적 보안) 실증랩 이용 이력 관리, 개인정보 취급자 지정 및 교육, 정기적 보안 취약 점검 등

[참고자료]

구분	법	시행령	시행규칙 등
개인정보보호 관계법령	개인정보 보호법	개인정보 보호법 시행령	개인정보의 안전성 확보조치 기준
			표준 개인정보 보호 지침
정보보호 관계법령	정보통신망 이용촉진 및 정보보호 등에 관한 법	정보통신망 이용촉진 및 정보보호 등에 관한 법률 시행령	정보통신망 이용촉진 및 정보보호 등에 관한 법률 시행규칙
	정보통신기반 보호법	정보통신기반 보호법 시행령	정보통신기반 보호법 시행규칙
	국가정보화 기본법	국가정보화 기본법 시행령	국가정보화 기본법 시행규칙
	전자정부법	전자정부법 시행령	-
	위치정보의 보호 및 이용 등에 관한 법률	위치정보의 보호 및 이용 등에 관한 법률 시행령	-
인공지능 참고법령	데이터 산업진흥 및 이용촉진에 관한 기본법(2022. 04. 20)	-	-
	산업 디지털 전환 촉진법 (2022. 01. 04)	-	-
	전자문서 및 전자거래 기본법	전자문서 및 전자거래 기본법 시행령	전자문서 및 전자거래 기본법 시행규칙
	전자서명법	전자서명법 시행령	전자서명법 시행규칙
	통신비밀보호법	통신비밀보호법 시행령	통신제한조치 등 허가 규칙

〈표 III-189〉 개인정보보호 관련 법령



[그림 III-20] 개인정보 처리 흐름도



참고자료

제1장 용어 정의

제2장 참고자료

제1장 | 용어 정의

7

- 가명처리
 - 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없는 특정 개인을 알아볼 수 없도록 처리하는 것을 말함
- 개방데이터
 - AI 통합 플랫폼인 AI허브(<https://aihub.or.kr/>)에 공개되는 인공지능 학습용 데이터
- 개인정보 영향평가
 - 개인정보파일의 운용으로 인하여 정보 주체의 개인정보 침해가 우려되는 경우에 그 위험요인의 분석과 개선사항 도출을 위한 평가
- 개인정보자기결정권
 - 자신에 관한 정보가 언제 누구에게 어느 범위까지 알려지고 이용되도록 할 것인지를 그 정보주체가 스스로 결정할 수 있는 권리, 즉 정보주체가 개인정보의 공개와 이용에 관하여 스스로 통제·결정할 수 있는 권리를 말함
- 객체
 - 실체가 있는 대상(사물) 자체를 지칭
예) 자동차, 자전거, 사람, 표지판, 건물, 도로, 하늘, 바다
- 검증 데이터셋(셋) (Validation Data Set)
 - 전체 학습데이터셋 중에서 일정한 비율을 정하여 인공지능의 기계학습에 따른 성능을 보정하거나 향상하는 용도로 사용하는 데이터셋
- 고유식별정보
 - ① 여권번호, ② 운전면허번호, ③ 외국인등록번호 등 단, 주민등록번호는 법률·대통령령 또는 헌법기관 규칙에 구체적 근거가 없는 경우 처리가 불가능함
 - * 출처: 개인정보보호법 제24조

- 과제조정위원회
 - 사업관리기관에서 사업수행기관이 제안한 사업수행계획서의 적정성을 검토하는 과정으로, 공모 절차에 의해 선정된 과제 대상으로 협약 이전에 진행
- 기계학습 Machine Learning
 - 인간이 자연적으로 수행하는 학습 능력과 같은 기능을 컴퓨터에서 실현하려는 기술이나 방법

ㄷ

- 데이터 획득/수집 (Data Acquisition/Collect)
 - 인공지능의 기계학습에 필요한 데이터를 현실 세계에서 직접 수집 또는 생성하거나, 이미 보유하고 있는 조직이나 시스템 등으로부터 법률적 제약이 없도록 ‘원시데이터’를 확보하는 활동
- 데이터 정제 (Data Refinement)
 - 획득한 원시데이터를 기계학습에 필요한 형식으로 맞추거나 불필요한 중복을 제거하며, 개인 정보를 비식별화하여 처리하는 등 일련의 전처리 과정을 통해 ‘원천데이터’를 확보하는 활동
- 데이터 가공 (Data Labeling)
 - 인공지능이 기계학습에 활용할 수 있도록 기능이나 목적에 부합하는 정보를 원천데이터에 부착하는 활동
- 데이터 학습 (Data Machine Learning)
 - 학습데이터셋의 훈련데이터셋, 검증데이터셋을 이용하여 선정된 인공지능 알고리즘을 학습시키고, 학습된 인공지능 모델의 성능을 향상시키거나 보정하는 활동

ㄹ

- 라벨링데이터 (Labeled Data)
 - 원천데이터에 부여한 ‘참값’, 파일형식이나 해상도 등의 속성, 그리고 설명이나 주석 등이 포함된 ‘어노테이션’의 집합

ㄴ

- 민감정보
 - ① 사상·신념, ② 노동조합·정당의 가입·탈퇴, ③ 정치적 견해, ④ 건강, ⑤ 성생활 등에 관한 정보, ⑥ 유전정보, ⑦ 범죄경력자료에 해당하는 정보, ⑧ 바이오정보(특징정보), ⑨ 인종이나

민족에 관한 정보 등 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보

* 출처: 개인정보보호법 제23조

ㅂ

- 범주
 - 개념적인 대상의 분류체계 예) 주제, 분야, 인종
- 비식별 조치
 - 정보 집합물(데이터셋)에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용, 개인을 알아볼 수 없도록 하는 조치
- 비정형데이터 (Unstructured Data)
 - 일정한 규격이나 형태를 지닌 숫자데이터와 달리 이미지, 영상, 음성, 보고서나 문서처럼 구조화 되지 않은 데이터

ㅅ

- 사업관리기관
 - 소관 부처가 사업에 대한 관리 등의 업무를 대행하기 위하여 설립하거나 지정한 전문기관
- 사업수행기관
 - 전담기관의 장과 별도의 협약 등을 통하여 사업을 수행하는 주관기관(기업), 참여기관(기업) 등을 통칭하는 것, 약칭 수행기관
- 시험 데이터(셋) (Test Data Set)
 - 전체 학습데이터셋 중에서 일정한 비율을 정하여 인공지능의 기계학습 성능 시험을 목적으로 사용하는 데이터셋

ㅇ

- 어노테이션 (Annotation)
 - 라벨링 공정에서 인간이 부여한 식별 기준을 기계가 인식할 수 있도록 선정된 데이터에 추가적인 정보를 기입하여 알고리즘이 이해할 수 있도록 만드는 과정

- 원시데이터 (Raw Data)
 - 기계학습을 목적으로 획득 단계에서 수집 또는 생성한 음성, 이미지, 영상, 텍스트 등의 데이터
- 원천데이터 (Source Data, Unlabeled Data)
 - 원시데이터를 라벨링 공정에 투입하기 위해 필요한 전처리 등 정제 작업을 수행한 데이터로 라벨링데이터가 부여되지 않은 상태의 데이터
- 인공지능 (Artificial Intelligence)
 - 자연 언어의 이해, 음성 번역, 문제 해결, 학습과 지식 획득, 인지 과학 등에 응용하기 위해 인간의 지능이 갖는 학습, 추리, 적응, 논증 등의 기능을 갖춘 컴퓨터 시스템
- 인공지능 학습용 공개 데이터셋 (AI Open Dataset)
 - 인공지능 통합플랫폼(AI Hub)에 등록하여 사용자가 활용할 수 있도록 공개된 학습데이터셋
- 인공지능 학습용 데이터 구축
 - 임무정의, 데이터 획득/수집, 데이터 정제, 데이터 가공(라벨링), 데이터 학습 등 인공지능 학습용 데이터를 구축하는 일련의 활동
- 인공지능 학습용 데이터 품질 (AI Data Quality)
 - 기계학습에 필요한 데이터의 적합성, 정확성, 유효성 등을 확보하여 사용자에게 유용한 가치를 줄 수 있는 수준
- 인공지능 학습용 데이터 품질관리 (AI Data Quality Management)
 - 인공지능 학습용 데이터 품질 확보에 필요한 조직, 절차, 기준, 제반 활동 등을 정의하여 점검하고 조치하는 일련의 활동
- 인스턴스 (Instance)
 - 클래스에 해당하는 개별 개체를 의미
 - 범주/클래스에 속하는 실제 데이터(example)로 어노테이션 대상을 의미
 - 예) ‘인공지능 구축과제’를 추진하는 경우 개 품종의 클래스인 ‘무들’, ‘비글’, ‘퍼그’ 등의 각 클래스에 해당하는 데이터 건수
- 임무 정의 (Task Define)
 - 인공지능 기계학습을 통해 해결하고자 하는 문제를 명확하게 정의하고, 문제 해결에 필요한 학습용 데이터의 요구사항을 구체적으로 정의하고 설계하는 활동

ㄱ

● 작업자

- 본 가이드라인에서 작업자란 인공지능 학습용 데이터 구축 과정에서 생성되는 원시데이터, 원천 데이터, 라벨링데이터 등의 구축 활동을 수행하는 인력을 의미하며, 크라우드워커(Crowd worker)라고 표현하기도 함
- 작업자는 역할에 따라 '수집 작업자', '정제 작업자', '가공 작업자'로 구분 가능하며, 각 구축 과정 내 '검수자' 또한 작업자의 범위에 속함
- 다만, 전문적인 지식을 기반으로 데이터를 구축하는 과업의 경우, 크라우드소싱 기반으로 작업자를 무작위 선정하지 않고 별도의 채용 과정을 통해 선정할 수 있음

● 제3자 품질검증

- 인공지능 학습용 데이터 구축 사업에 직접 참여하는 당사자인 수행기관 및 참여기관 등 이외의 자를 제3자라고 하며, 제3자가 데이터의 다양성, 정확성, 유효성 등을 검토하는 과정

● 주관 사업수행기관

- 사업수행기관이 다수인 경우, 사업비 신청·관리 및 실적보고와 정산 등 해당 협약사업에 대한 책임이 있는 사업수행기관 또는 사업비 지분이 가장 많은 사업수행기관, 약칭 주관기관

ㄴ

● 참값 (Ground Truth)

- 인공지능의 기계학습 목적에 따라 원천데이터에 라벨링된 정확한 값이나 사실의 의미적 표현

● 참여 사업수행기관

- 수행기관과 공동으로 해당 과제에 참여하여 과제를 수행하는 기관, 약칭 참여기관

ㄷ

● 크라우드소싱 (Crowdsourcing)

- 대중(crowd)과 아웃소싱(outsourcing)의 합성어로 기업 활동의 일부 과정에서 일반대중(크라우드워커)을 참여시키는 것을 의미

- 크라우드워커 (Crowd worker)

- 일반인이 기업의 업무 용역을 대행 수행하고, 일정 대가를 받는 경우를 의미하며, 집이나 재택근무 등의 형태로도 업무 수행이 가능하고, 자유롭게 할당된 과제물을 수행하는 일자리
- 예) A 참여기관이 '음식데이터 AI구축과제'를 추진하는 경우 '크라우드워커'를 고용하여 일반인으로부터의 음식 사진 데이터를 대량 수집

- 클래스 (Class)

- 분류/탐지하고자 하는 대상을 카테고리화 한 것으로 분류체계를 의미
- 실체가 있는 대상(사물)의 분류체계 명칭
- 예) 자동차, 자전거, 사람, 표지판, 건물, 도로, 하늘, 바다
- 예) A 참여기관이 '스팸 감지 인공지능 구축과제'를 추진하는 경우 스팸에 대한 클래스는 '스팸', '스팸 아님'이며, B 참여기관이 '개 품종을 판별하는 인공지능 구축과제'를 추진하는 경우 개 품종의 클래스는 '푸들', '비글', '퍼그' 등등

II

- 품질관리 지표

- 데이터의 품질 수준을 측정하기 위한 관점을 정의한 것으로 무엇을 측정할 것인가에 대한 기준

- 품질 자가점검

- 데이터 품질 이슈 조기발견 및 조치를 위한 사업수행기관 자체적인 전 구축과정(획득/수집 → 정제 → 가공 → 학습)에 대한 품질점검 활동

III

- 학습데이터셋 (AI Data Set)

- 인공지능의 기계학습에 사용하는 원천데이터와 라벨링데이터의 묶음을 말하며, 사용하는 목적에 따라 '훈련데이터셋', '검증데이터셋', '시험데이터셋'으로 구분

- 훈련데이터셋 (Training Data Set)

- 전체 학습데이터셋 중에서 일정한 비율을 정하여 인공지능의 기계학습에 직접 사용하는 데이터셋

- 학습모델 임무 (Learning Model TASK)

- 학습모델 알고리즘의 목적을 구분하기 위한 단위로서, 지능별 기술*에 따라 '시각지능', '언어·청각지능', '음성지능', '융합지능' 등 지능별 기술에 따라 다양하게 분류할 수 있음

* 출처: NIA, AI INSIGHT REPORT('19.12.)

- 다만, 본 『제1권. 품질관리 가이드라인 v3.0』에서는 NIA에서 공모하는 ‘인공지능 학습용 데이터 구축사업’의 구축사례들을 기반으로 널리 활용되는 Task 분야를 정리하여 ‘언어(이해/생성)지능’ 및 ‘시각(인지/생성)지능’으로 분류함

G

- General Data Protection Regulation (EU GDPR)
 - 유럽연합의 개인정보보호법

P

- PbD 원칙 (Privacy by Design)
 - 제품·서비스 개발 시 기획 단계부터 개인정보 처리의 전체 생애주기에 걸쳐 이용자의 프라이버시를 고려한 기술·정책을 설계에 반영하는 것을 의미하며, 국제적으로 광범위하게 통용되는 개인정보 보호 원칙임

1

- 1-Cycle 자가점검
 - 인공지능 학습용 구축 초기의 데이터를 이용하여 「획득/수집 → 정제 → 가공 → 학습」까지 전체 공정에 대한 1-Cycle 데이터(적합성, 정확성, 유효성)를 검사하여, 품질 이슈 조기발견 및 신속한 개선조치를 위한 활동

제2장 | 참고자료

1. Data quality for analytics and machine learning — Part 4: Data quality process framework(ISO/IEC WD 5259-4:2020(E))
2. “Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation”. ResearchGate. <https://www.researchgate.net/publication/311531910_Optimizing_Intersection-Over-Union_in_Deep_Neural_Networks_for_Image_Segmentation>
3. “Augmented Lagrangian-based Algorithm for Projective Reconstruction from Multiple Views with Minimization of 2D Reprojection Error”. SEMANTIC SCHOLAR. <<https://www.semanticscholar.org/paper/Augmented-Lagrangian-based-Algorithm-for-Projective-Mai-Hung/fb3b33345a25d36cdcecee5966abf8086d2d0f36>>
4. Moumita Mukherjee et al. Electronics and Telecommunication Engineering, National Institute of Technology, “Kinship verification using Compound Local Binary Pattern and Local Feature Discriminant Analysis” (2019)
5. Gabriela Csurka et al. Xerox Research Centre Europe, “What is a good evaluation measure for semantic segmentation?” (2013)
6. Keni Bernardin et al, “Multiple Object Tracking Performance Metrics and Evaluation in a Smart Room Environment” (2006)
7. Sergio Escalera Guerrero, ChaLearn, Computer Vision Center (UAB) and University of Barcelona, “Tutorial on Datasets, Annotations and Metrics in Computer Vision” (2015)
8. Ergys Ristani et al. “Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking” (2016)
9. Jinbao Wang et al. “Deep 3D human pose estimation: A review” (2020)
10. Shenqi Lai et al, “Enhanced Normalized Mean Error loss for Robust Facial Landmark detection” (2019)
11. [TTA] 2020.09 인공지능 데이터 품질관리 요구사항
12. [NIA] 2011 정보화사업 PMO 운영관리 매뉴얼

13. [NIA] 2014 데이터베이스 구축 방법론 V4.0
14. [NIA] 2018.01 공공데이터 품질관리 매뉴얼 v2.0
15. [NIA] 2018.12 우리나라 인공지능(A.I.) 분야 수준 조사 연구
16. [NIA] 2019 빅데이터 플랫폼 및 센터 데이터 품질관리 가이드
17. [NIA] 2022 인공지능 학습용 데이터 구축 사업관리 매뉴얼
18. [NIA] 2022 인공지능 학습용 데이터 가이드라인 v2.0
19. [NIA] 2022.12.31. [IT & Future Strategy 제12호] ‘가짜’ 데이터가 만드는 ‘진짜’ 인공지능 시대
20. [K-DATA] 2006.09 데이터 품질관리 지침
21. [관계부처 합동] 2016.06.30 개인정보 비식별 조치 가이드라인
22. [개인정보보호위원회] 2021.05.31 인공지능(AI) 개인정보보호 자율점검표
23. [행정안전부] 행정안전부훈령 제 173호(공간정보 보안관리규정, 개정 2020. 12. 29)
24. [국토교통부] 공간정보 품질기준(국가공간정보센터 운영세부규정, 시행 2018. 1. 1)
25. [NIPA 싱가포르지원센터] 2020.01.03 2019 글로벌 ICT 이슈리포트-싱가포르 정부의 인공지능(AI) 정책
26. [NIA] 2022 인공지능 학습용 데이터 라벨링 전문교육 - 데이터 라벨러 과정 - 품질관리 심화과정 교재
27. [NIA] 2022 인공지능 학습용 데이터 라벨링 전문 교육 - 데이터 라벨러 과정 - 입문, 이미지/영상
28. [NIA] 2022 인공지능 학습용 데이터 라벨링 전문 교육 - 데이터 라벨러 과정 - 입문, 음성/텍스트
29. Google Machine Learning
[〈https://developers.google.com/machine-learning/crash-course?hl=ko〉](https://developers.google.com/machine-learning/crash-course?hl=ko)
309. Paperwithcode [〈https://paperswithcode.com/〉](https://paperswithcode.com/)
31. “Metric: F1”, Hugging Face [〈https://huggingface.co/metrics/f1〉](https://huggingface.co/metrics/f1)
32. DeepAI [〈https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate〉](https://deepai.org/machine-learning-glossary-and-terms/accuracy-error-rate)
33. “Metric: BLEU”, Hugging Face [〈https://huggingface.co/metrics/bleu〉](https://huggingface.co/metrics/bleu)
34. “Exact Matching”, MatchIt
[〈https://kosukeimai.github.io/MatchIt/reference/method_exact.html〉](https://kosukeimai.github.io/MatchIt/reference/method_exact.html)
35. “ROUGE”, ACL Anthology. [〈https://aclanthology.org/W04-1013/〉](https://aclanthology.org/W04-1013/)
36. Daniel J Bell, “Dice similarity coefficient”, Radiopaedia,
[〈https://radiopaedia.org/articles/dice-similarity-coefficient〉](https://radiopaedia.org/articles/dice-similarity-coefficient)
37. “Recall”. C3.ai. [〈https://c3.ai/glossary/data-science/recall/〉](https://c3.ai/glossary/data-science/recall/)
38. “Precision”. C3.ai [〈https://c3.ai/glossary/machine-learning/precision〉](https://c3.ai/glossary/machine-learning/precision)

39. “Accuracy Metrics”. Humboldt State University – GSP216 Introduction to remote sensing
<http://gsp.humboldt.edu/olm_2019/courses/GSP_216_Online/lesson6-2/metrics.html>
40. “Root Mean Square Error”. C3.ai.
<<https://c3.ai/glossary/data-science/root-mean-square-error-rmse/>>
41. “Evaluate: Keypoints”, COCO Common Objects in Context
<<https://cocodataset.org/#keypoints-eval>>
42. “Metric: cer”, Hugging Face <<https://huggingface.co/metrics/cer>>
43. “Metric: wer”, Hugging Face <<https://huggingface.co/metrics/wer>>
44. “Perplexity in Language Models”, towards data science
<<https://towardsdatascience.com/perplexity-in-language-models-87a196019a94>>
45. “M.O.S. MEAN OPINION SCORE”, Arimas
<<https://arimas.com/2017/08/01/mos-mean-opinion-score/>>
46. “fbeta_score”.
<https://scikit-learn.org/stable/modules/generated/sklearn.metrics.fbeta_score.html>
47. F-beta score. <<https://hasty.ai/docs/mp-wiki/metrics/f-beta-score>>
48. 국가법령정보센터<<https://www.law.go.kr/>>
49. AI Hub, 전문분야 학습용 한영 번역 말뭉치 <https://aihub.or.kr/sites/default/files/Sample_data/구축활용가이드북_2020-01/006.전문분야_한영번역말뭉치_데이터_구축_가이드라인.pdf>

품질관리 가이드라인 및 구축 안내서 v3.0

제1권 품질관리 가이드라인 v3.0

2023년 2월 발행

발행처: 한국지능정보사회진흥원

〈작성 참여진〉

한국지능정보사회진흥원 이용진 본부장	한국지능정보사회진흥원 손기문 팀장
한국지능정보사회진흥원 오현목 팀장	한국지능정보사회진흥원 유호진 팀장
한국지능정보사회진흥원 신다울 팀장	한국지능정보사회진흥원 박현우 수석
한국지능정보사회진흥원 전진우 선임	한국지능정보사회진흥원 윤주미 연구원
(주)티지 오윤환 이사	(주)비트레스 류동주 대표
(주)아이티지엔 양원 대표	

〈자문 위원〉

한국정보통신기술협회 신준호 단장	가톨릭대학교 이상국 교수
세종대학교 김진철 교수	한국전자기술연구원 김동철 책임
한국표준과학연구원 이재두 초빙연구위원	(주)아이삭 김상복 기술사

- 본 『제1권 품질관리 가이드라인 v3.0』 내용의 무단전재 및 재배포를 금하며, 가공·인용 시에는 반드시 과학기술정보통신부, 한국지능정보사회진흥원의 「인공지능 학습용 데이터 품질관리 가이드라인 및 구축 안내서 v3.0 - 제1권 품질관리 가이드라인 v3.0」임을 밝혀주시기 바랍니다.
- 본 「제1권 품질관리 가이드라인 v3.0」은 지능정보산업 인프라 조성을 위한 인공지능 학습용 데이터 구축사업 중 ‘인공지능 학습용 데이터 품질관리 컨설팅 및 체계 고도화’ 영역 사업의 결과 산출물입니다.
- 한국지능정보사회진흥원 지능데이터본부 AI데이터활용팀 품질유닛 : aidq@nia.or.kr, 053-230-4220