

Microsoft AI-900 용어 정리 (Terminology Guide)

❖ 사용 가이드

- 시험 전 최종 암기용
- 한글 ↔ 영문 모두 숙지
- ★ 표시: 시험 빈출 용어
- [예시] 포함된 용어는 개념 이해 필수

A

Accountability (책임성) ★

- **한글**: 책임성, 책무성
- **정의**: AI 시스템에 대한 책임 소재가 명확해야 함
- **책임 있는 AI 6원칙 중 하나**
- [예시] AI가 오류 발생 시 책임자가 명확해야 함

Accuracy (정확도)

- **한글**: 정확도
- **정의**: 전체 예측 중 올바른 예측의 비율
- **공식**: (올바른 예측 수) / (전체 예측 수)
- [예시] 100개 중 85개 맞춤 → 85% 정확도

AI (Artificial Intelligence) ★

- **한글**: 인공지능
- **정의**: 인간의 지능을 모방하는 컴퓨터 시스템
- **포함**: ML, Deep Learning, NLP, Computer Vision

Anomaly Detection (이상 탐지)

- **한글**: 이상 감지, 이상치 탐지
- **정의**: 정상 패턴에서 벗어난 데이터 찾기
- [예시] 신용카드 사기 거래 탐지

API (Application Programming Interface)

- **한글**: API, 응용 프로그래밍 인터페이스
- **정의**: 서비스 간 통신을 위한 인터페이스
- [예시] Azure AI Vision을 호출하는 REST API

Automated ML (AutoML) ★

- **한글**: 자동화된 머신러닝
- **정의**: 최적의 ML 모델을 자동으로 찾아주는 기능

- **포함:** 알고리즘 선택, 하이퍼파라미터 튜닝, 전처리
- [예시] Azure ML의 Automated ML 기능

Azure AI Foundry ★ (New!)

- **한글:** Azure AI 파운드리
- **구명칭:** Azure AI Studio
- **정의:** AI 앱 구축 통합 플랫폼
- **기능:** Playground, Prompt Flow, 모델 배포

Azure AI Services ★

- **한글:** Azure AI 서비스
- **구명칭:** Azure Cognitive Services
- **정의:** 사전 학습된 AI 기능을 API로 제공
- **포함:** Vision, Language, Speech, OpenAI

Azure Machine Learning ★

- **한글:** Azure 머신러닝
- **약어:** Azure ML
- **정의:** 엔터프라이즈급 ML 플랫폼
- **기능:** AutoML, Designer, Notebooks, MLOps

Azure OpenAI Service ★

- **한글:** Azure OpenAI 서비스
- **정의:** OpenAI 모델을 Azure에서 제공
- **모델:** GPT-4, GPT-3.5, DALL-E, Whisper
- **장점:** 엔터프라이즈 보안, 데이터 프라이버시

B

Batch Endpoint

- **한글:** 배치 엔드포인트
- **정의:** 대량 데이터를 일괄 처리하는 배포 방식
- **용도:** 정기적인 대량 예측 작업

BERT (Bidirectional Encoder Representations from Transformers)

- **한글:** 버트
- **정의:** Google의 양방향 Transformer 모델
- **용도:** 텍스트 이해, 질의응답

Bias (편향) ★

- **한글:** 편향, 편견
- **정의:** 학습 데이터 또는 모델의 불공정한 성향

- [예시] 특정 성별/인종에 불리한 AI 결정

Bounding Box (경계 상자) ★

- **한글**: 경계 상자, 바운딩 박스
- **정의**: 객체의 위치를 표시하는 사각형
- **포함**: x, y 좌표, 너비, 높이
- [예시] Object Detection에서 차량 위치 표시

Bot Service

- **한글**: 봇 서비스
 - **정의**: 챗봇을 구축하고 배포하는 Azure 서비스
 - **통합**: CLU, QnA Maker와 연동
-

C

Classification (분류) ★★★

- **한글**: 분류
- **정의**: 데이터를 미리 정의된 카테고리로 분류
- **출력**: 범주형 (클래스)
- [예시] 이메일 스팸 필터 (스팸/정상)
- **관련**: Regression, Clustering과 구분 필수

Clustering (클러스터링) ★★★

- **한글**: 군집화, 클러스터링
- **정의**: 레이블 없는 데이터를 유사한 그룹으로 묶기
- **학습 유형**: 비지도 학습
- [예시] 고객 세분화
- **알고리즘**: K-Means, DBSCAN

CLU (Conversational Language Understanding) ★

- **한글**: 대화형 언어 이해
- **구명칭**: LUIS (Language Understanding)
- **정의**: 사용자 의도와 개체를 파악
- **구성**: Intent + Entity

Cognitive Services → Azure AI Services

- **한글**: 코그너티브 서비스
- **변경**: 현재는 "Azure AI Services"로 명칭 변경
- **시험 주의**: 둘 다 같은 서비스!

Computer Vision (컴퓨터 비전) ★

- **한글**: 컴퓨터 비전

- **정의:** 이미지와 비디오를 이해하는 AI
- **포함:** 분류, 객체 감지, OCR, 얼굴 인식

Content Filtering (콘텐츠 필터링) ★

- **한글:** 콘텐츠 필터링
- **정의:** 유해 콘텐츠 자동 차단
- **카테고리:** Hate, Sexual, Violence, Self-harm
- **심각도:** Safe, Low, Medium, High

Custom Vision ★

- **한글:** 커스텀 비전
- **정의:** 사용자 정의 이미지 모델 학습 서비스
- **용도:** 특정 도메인의 분류/객체 감지
- [예시] 제조업 불량품 감지

Compute Instance (컴퓨팅 인스턴스)

- **한글:** 컴퓨팅 인스턴스
- **정의:** Azure ML의 개발용 가상 머신
- **용도:** Jupyter Notebook, 모델 개발

Compute Cluster (컴퓨팅 클러스터)

- **한글:** 컴퓨팅 클러스터
- **정의:** 학습 작업용 확장 가능한 클러스터
- **특징:** 자동 확장, CPU/GPU 선택

D

DALL-E ★

- **한글:** 달리, 달-E
- **정의:** 텍스트로 이미지를 생성하는 모델
- **버전:** DALL-E 3 (최신)
- [예시] "해변의 석양" → 이미지 생성

Dataset (데이터셋)

- **한글:** 데이터셋, 데이터 집합
- **정의:** ML 학습에 사용하는 데이터 모음
- **분할:** Training / Validation / Test

Datastore (데이터스토어)

- **한글:** 데이터스토어
- **정의:** Azure Storage 연결 정보 저장
- **예시:** Blob Storage, Data Lake 연결

Deep Learning (딥러닝) ★

- **한글**: 딥러닝, 심층학습
- **정의**: 다층 신경망을 사용하는 ML
- **특징**: 대량 데이터 필요, 자동 특징 추출
- **유형**: CNN, RNN, Transformer

Designer (디자이너)

- **한글**: 디자이너
- **정의**: Azure ML의 드래그 앤 드롭 파이프라인 도구
- **용도**: 코드 없이 ML 워크플로우 구축

Document Intelligence ★★

- **한글**: 문서 인텔리전스
- **구명칭**: Form Recognizer
- **정의**: 구조화된 문서에서 데이터 자동 추출
- **모델**: 영수증, 송장, 신분증, 명함
- **차이점**: OCR은 텍스트만, Document Intelligence는 구조+의미

E

Embedding (임베딩) ★

- **한글**: 임베딩
- **정의**: 텍스트를 숫자 벡터로 변환
- **차원**: 일반적으로 1536차원
- **용도**: 의미 검색, 유사도 비교, RAG

Endpoint (엔드포인트)

- **한글**: 엔드포인트, 종단점
- **정의**: 배포된 모델에 접근하는 URL
- **유형**: Real-time, Batch

Entity (개체) ★

- **한글**: 개체, 엔티티
- **정의**: 텍스트에서 식별된 특정 정보
- **유형**: Person, Location, Organization, DateTime
- [예시] "서울"은 Location 개체

Entity Recognition (개체 인식) ★

- **한글**: 개체 인식, 엔티티 인식
- **약어**: NER (Named Entity Recognition)
- **정의**: 텍스트에서 사람, 장소, 조직 등을 식별
- [예시] "삼성전자는 서울에 있다" → 삼성전자(조직), 서울(위치)

F

Face Detection (얼굴 감지) ★

- **한글**: 얼굴 감지, 얼굴 탐지
- **정의**: 이미지에서 얼굴 위치 찾기
- **출력**: 얼굴 경계 상자, 랜드마크

Face Analysis (얼굴 분석) ★

- **한글**: 얼굴 분석
- **정의**: 얼굴의 속성 분석
- **속성**: 나이, 감정, 안경 착용 여부
- **주의**: 제한된 액세스 (승인 필요)

Fairness (공정성) ★★

- **한글**: 공정성
- **정의**: AI가 모든 사람을 공평하게 대우
- **책임 있는 AI 6원칙 중 하나**
- [예시] 채용 AI가 성별/인종 차별 없어야 함

Feature (특징) ★★

- **한글**: 특징, 피처, 변수
- **정의**: 모델의 입력 데이터
- [예시] 주택 가격 예측 시 면적, 방 개수, 위치

Few-shot Learning

- **한글**: 퓨샷 러닝, 소수 예시 학습
- **정의**: 적은 예시로 학습하는 기법
- **용도**: Prompt Engineering에서 예시 제공

Fine-tuning (미세 조정)

- **한글**: 미세 조정, 파인튜닝
 - **정의**: 사전 학습된 모델을 특정 작업에 맞게 조정
 - **비교**: RAG보다 비용 높음
-

G

Generative AI (생성형 AI) ★★★

- **한글**: 생성형 AI, 생성 AI
- **정의**: 새로운 콘텐츠를 생성하는 AI
- **유형**: 텍스트, 코드, 이미지, 음악
- **모델**: GPT, DALL-E, Stable Diffusion

- **비중**: 시험의 20-25% (가장 큰 비중!)

GPT (Generative Pre-trained Transformer) ★★

- **한글**: GPT
 - **정의**: OpenAI의 대규모 언어 모델
 - **버전**: GPT-3.5, GPT-4, GPT-4o
 - **특징**: Transformer 기반
-

H

Hallucination (환각) ★

- **한글**: 환각, 허위 정보 생성
- **정의**: AI가 사실이 아닌 내용을 그럴듯하게 생성
- **위험**: 잘못된 정보 전파
- **대응**: 사실 확인, RAG 사용

Hyperparameter (하이퍼파라미터)

- **한글**: 하이퍼파라미터, 초매개변수
 - **정의**: 학습 과정을 제어하는 설정값
 - [예시] Learning rate, Batch size
-

I

Image Classification (이미지 분류) ★★

- **한글**: 이미지 분류
- **정의**: 이미지 전체를 하나의 카테고리로 분류
- **출력**: 1개 레이블 + 신뢰도
- [예시] 고양이/개 분류
- **차이점**: Object Detection은 여러 객체 + 위치

Inclusiveness (포용성) ★

- **한글**: 포용성
- **정의**: 모든 사람이 AI 혜택을 누려야 함
- **책임 있는 AI 6원칙 중 하나**
- [예시] 접근성 기능, 다국어 지원

Inference (추론)

- **한글**: 추론, 예측
- **정의**: 학습된 모델로 새 데이터 예측
- **단계**: Training → Inference

Intent (의도) ★

- **한글**: 의도, 인텐트
 - **정의**: 사용자가 원하는 것
 - **예시**: "피자 주문" 의도
 - **구성**: Intent + Entity
-

K

Key Phrase Extraction (핵심 구문 추출) ★

- **한글**: 핵심 구문 추출, 키워드 추출
 - **정의**: 텍스트에서 주요 개념 자동 추출
 - [예시] "Azure AI는 훌륭합니다" → "Azure AI"
-

L

Label (레이블) ★★

- **한글**: 레이블, 정답, 타겟
- **정의**: 모델이 예측하려는 값 (지도학습)
- [예시] 스팸 필터의 "스팸/정상"

Language Modeling (언어 모델링) ★

- **한글**: 언어 모델링
- **정의**: 언어 패턴을 학습하여 텍스트 이해/생성
- **모델**: GPT, BERT

LLM (Large Language Model) ★★

- **한글**: 대규모 언어 모델
 - **정의**: 방대한 텍스트로 학습된 거대 모델
 - [예시] GPT-4, Claude, LLaMA
-

M

Machine Learning (머신러닝) ★★★

- **한글**: 머신러닝, 기계학습
- **약어**: ML
- **정의**: 데이터로부터 패턴을 학습하는 AI
- **유형**: 지도학습, 비지도학습, 강화학습
- **관계**: AI > ML > Deep Learning

Model (모델)

- **한글**: 모델
- **정의**: 학습된 ML 알고리즘

- **생명주기:** 학습 → 평가 → 배포 → 모니터링

Model Catalog ★ (New!)

- **한글:** 모델 카탈로그
 - **정의:** Azure에서 사용 가능한 모든 AI 모델의 저장소
 - **포함:** OpenAI, Microsoft, Llama, Mistral 등
-

N

Neural Network (신경망)

- **한글:** 신경망, 인공신경망
- **정의:** 뇌의 뉴런 구조를 모방한 ML 모델
- **종류:** CNN, RNN, Transformer

NLP (Natural Language Processing) ★★

- **한글:** 자연어 처리
 - **정의:** 인간 언어를 이해하고 생성하는 AI
 - **포함:** 감정 분석, 번역, 음성 인식, 텍스트 생성
-

O

Object Detection (객체 감지) ★★★

- **한글:** 객체 감지, 객체 탐지
- **정의:** 이미지 내 여러 객체의 위치와 종류 식별
- **출력:** 여러 Bounding Box + 레이블
- [예시] 자율주행차의 보행자/차량 감지
- **차이점:** Image Classification은 1개 레이블만

OCR (Optical Character Recognition) ★★

- **한글:** 광학 문자 인식
- **정의:** 이미지에서 텍스트 추출
- [예시] 명함 스캔, 문서 디지털화
- **차이점:** Document Intelligence는 구조도 이해

Overfitting (과적합) ★

- **한글:** 과적합, 오버피팅
 - **정의:** 학습 데이터에만 너무 잘 맞아서 새 데이터에서 성능 저하
 - **해결:** 더 많은 데이터, 정규화
-

P

Playground ★

- **한글**: 플레이그라운드
- **정의**: Azure AI Foundry에서 모델을 테스트하는 환경
- **종류**: Chat, Completions, DALL-E

Precision (정밀도)

- **한글**: 정밀도
- **정의**: 양성 예측 중 실제 양성 비율
- **공식**: $TP / (TP + FP)$

Prompt (프롬프트) ★★

- **한글**: 프롬프트, 지시문
- **정의**: AI에게 주는 입력 질문/지시사항
- [예시] "Azure AI를 설명해줘"

Prompt Engineering (프롬프트 엔지니어링) ★★

- **한글**: 프롬프트 엔지니어링
- **정의**: 효과적인 프롬프트를 작성하는 기술
- **기법**: 명확성, 구체성, 역할 부여, Few-shot

Prompt Flow

- **한글**: 프롬프트 플로우
 - **정의**: Azure AI Foundry의 AI 워크플로우 구축 도구
 - **용도**: 여러 AI 모델을 체인으로 연결
-

R

RAG (Retrieval-Augmented Generation) ★★

- **한글**: 검색 증강 생성
- **정의**: 외부 문서를 검색하여 LLM 응답 개선
- **구성**: 문서 검색 + LLM 생성
- [예시] 회사 내부 문서 기반 Q&A

Real-time Endpoint

- **한글**: 실시간 엔드포인트
- **정의**: 즉시 응답하는 배포 방식
- **용도**: 낮은 지연시간 필요 시

Recall (재현율)

- **한글**: 재현율
- **정의**: 실제 양성 중 찾아낸 비율
- **공식**: $TP / (TP + FN)$

Regression (회귀) ★★★

- **한글**: 회귀
- **정의**: 연속적인 숫자 값을 예측
- **출력**: 숫자 (가격, 온도, 수량 등)
- [예시] 주택 가격 예측, 매출 예측
- **관련**: Classification, Clustering과 구분 필수

Reliability (신뢰성) ★

- **한글**: 신뢰성
- **정의**: AI가 일관되게 작동
- **책임 있는 AI 6원칙 중 하나** (Safety와 함께)

Responsible AI (책임 있는 AI) ★★★

- **한글**: 책임 있는 AI
- **6원칙** (암기 필수!):
 1. Fairness (공정성)
 2. Reliability & Safety (신뢰성/안전성)
 3. Privacy & Security (개인정보/보안)
 4. Inclusiveness (포용성)
 5. Transparency (투명성)
 6. Accountability (책임성)

S

Safety (안전성) ★

- **한글**: 안전성
- **정의**: AI가 안전하게 작동
- **책임 있는 AI 6원칙 중 하나** (Reliability와 함께)

Security (보안) ★

- **한글**: 보안
- **정의**: 데이터와 시스템을 보호
- **책임 있는 AI 6원칙 중 하나** (Privacy와 함께)

Self-Attention

- **한글**: 셀프 어텐션, 자기 주의
- **정의**: Transformer의 핵심 메커니즘
- **역할**: 문장 내 단어 간 관계 파악

Sentiment Analysis (감정 분석) ★★

- **한글**: 감정 분석, 감성 분석
- **정의**: 텍스트의 감정적 톤 분석

- **출력**: 긍정/부정/중립 + 점수
- [예시] 고객 리뷰 "최고예요!" → 긍정 (0.95)

Speech-to-Text (음성 인식) ★★

- **한글**: 음성 인식, STT
- **영어**: STT
- **정의**: 음성을 텍스트로 변환
- [예시] 회의록 자동 작성

Speech Translation (음성 번역)

- **한글**: 음성 번역
- **정의**: 음성을 다른 언어의 텍스트/음성으로 변환
- [예시] 한국어 음성 → 영어 텍스트

Supervised Learning (지도 학습) ★

- **한글**: 지도 학습
- **정의**: 레이블이 있는 데이터로 학습
- **포함**: Regression, Classification
- **특징**: Features + Labels 필요

System Message

- **한글**: 시스템 메시지
- **정의**: AI의 역할/행동을 정의하는 초기 지시
- [예시] "당신은 친절한 고객 서비스 담당자입니다"

T

Temperature (온도) ★★

- **한글**: 온도, 템퍼러처
- **정의**: 생성 다양성을 조절하는 파라미터
- **범위**: 0.0 (일관적) ~ 1.0 (창의적)
- [예시] Temperature 0.1 → 예측 가능한 응답

Test Set (테스트 셋) ★

- **한글**: 테스트 데이터, 테스트 셋
- **정의**: 최종 모델 성능 평가 데이터
- **비율**: 10-15%
- **주의**: 학습에 절대 사용 안 됨

Text-to-Speech (음성 합성) ★★

- **한글**: 음성 합성, TTS
- **영어**: TTS

- 정의: 텍스트를 자연스러운 음성으로 변환
- [예시] 내비게이션 안내, 오디오북

Training Set (학습 셋) ★★

- 한글: 학습 데이터, 훈련 셋
- 정의: 모델을 학습시키는 데이터
- 비율: 70-80%

Transformer ★★

- 한글: 트랜스포머
- 정의: Self-Attention 기반 신경망 구조
- 특징: 병렬 처리, 장거리 의존성
- 모델: GPT, BERT, Claude 등의 기반

Translation (번역) ★

- 한글: 번역
- 정의: 언어 간 텍스트/음성 변환
- 서비스: Azure AI Translator
- [예시] 영어 → 한국어

Transparency (투명성) ★★

- 한글: 투명성
- 정의: AI 작동 방식을 이해할 수 있어야 함
- 책임 있는 AI 6원칙 중 하나
- [예시] AI 의사결정 과정 설명 가능

U

Underfitting (과소적합)

- 한글: 과소적합, 언더피팅
- 정의: 모델이 너무 단순해서 패턴 학습 실패
- 해결: 더 복잡한 모델, 더 많은 특징

Unsupervised Learning (비지도 학습) ★

- 한글: 비지도 학습
- 정의: 레이블 없는 데이터로 학습
- 포함: Clustering
- 특징: 패턴 발견, 그룹화

V

Validation Set (검증 셋) ★★

- **한글**: 검증 데이터, 검증 셋
- **정의**: 학습 중 모델 성능 평가 데이터
- **비율**: 10-15%
- **역할**: 과적합 방지, 하이퍼파라미터 조정

Vision Studio

- **한글**: 비전 스튜디오
- **정의**: Azure AI Vision을 테스트하는 웹 환경
- **기능**: 이미지 분석, OCR, 객체 감지 테스트

W

Whisper

- **한글**: 위스퍼
- **정의**: OpenAI의 음성 인식 모델
- **용도**: Speech-to-Text
- **특징**: 다국어 지원, 높은 정확도

시험 필수 암기 용어 Top 20 ★★★

1. 책임 있는 AI 6원칙

1. Fairness (공정성)
2. Reliability & Safety (신뢰성/안전성)
3. Privacy & Security (개인정보/보안)
4. Inclusiveness (포용성)
5. Transparency (투명성)
6. Accountability (책임성)

2. ML 기법 3가지

1. Regression (회귀) - 숫자 예측
2. Classification (분류) - 카테고리
3. Clustering (클러스터링) - 그룹화 (레이블 없음)

3. Computer Vision 워크로드

1. Image Classification - 1개 레이블
2. Object Detection - 여러 객체 + 위치
3. OCR - 텍스트 추출
4. Face Detection/Analysis - 얼굴
5. Document Intelligence - 문서 구조 이해

4. NLP 작업

1. Key Phrase Extraction - 핵심 구문
2. Entity Recognition - 개체 식별
3. Sentiment Analysis - 감정 분석
4. Speech-to-Text - 음성 인식
5. Text-to-Speech - 음성 합성
6. Translation - 번역

5. 생성형 AI 관련

1. GPT - 텍스트 생성 모델
2. DALL-E - 이미지 생성
3. LLM - 대규모 언어 모델
4. Prompt Engineering - 프롬프트 작성 기술
5. Temperature - 다양성 조절
6. RAG - 검색 증강 생성
7. Hallucination - 거짓 정보 생성

6. Azure 서비스

1. Azure AI Services - 통합 AI 서비스
2. Azure Machine Learning - ML 플랫폼
3. Azure OpenAI Service - GPT, DALL-E
4. Azure AI Foundry - 통합 개발 환경
5. Model Catalog - 모델 저장소
6. Custom Vision - 사용자 정의 분류/객체 감지
7. Azure AI Document Intelligence - 문서 처리

7. ML 데이터 분할

1. Training Set (70-80%) - 학습용
2. Validation Set (10-15%) - 검증용
3. Test Set (10-15%) - 최종 평가용

8. 핵심 개념

1. Features - 입력 변수
2. Labels - 출력 값 (정답)
3. Overfitting - 과적합

- 4. Transformer - 현대 AI의 기반 구조
- 5. Automated ML - 자동 모델 찾기