

글로벌 ICT 표준 컨퍼런스 2022

Global ICT Standards Conference 2022

2022. 11.9.(수)~11.(금)
서울 양재 엘타워 오르체홀(5F)



글로벌 표준화 세미나

인공지능과 데이터 품질 표준화

하 수 욱 책임연구원 ETRI

글로벌 ICT 표준 컨퍼런스 2022

Global ICT Standards Conference 2022

INDEX

01 Data-driven age

02 Standardization activities

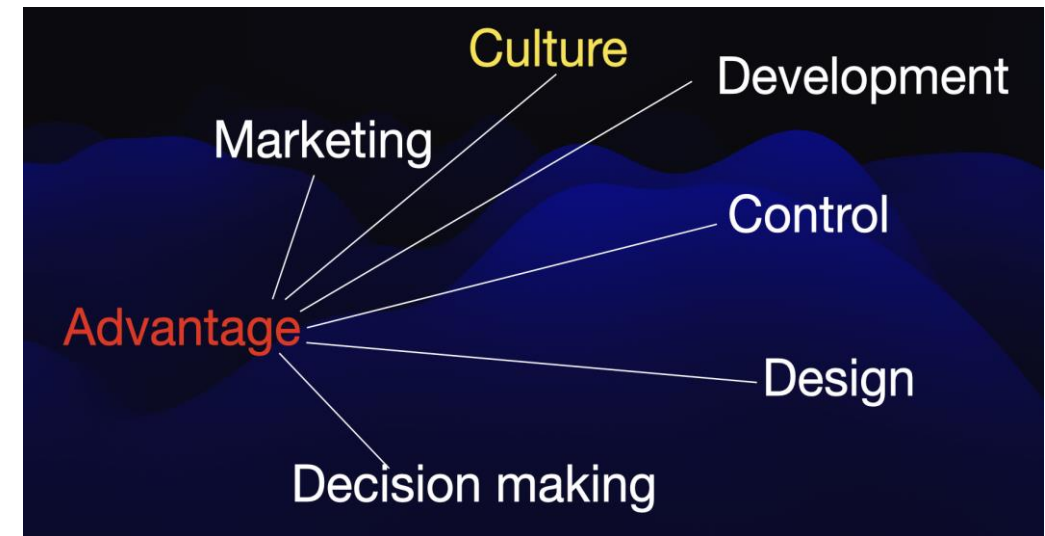
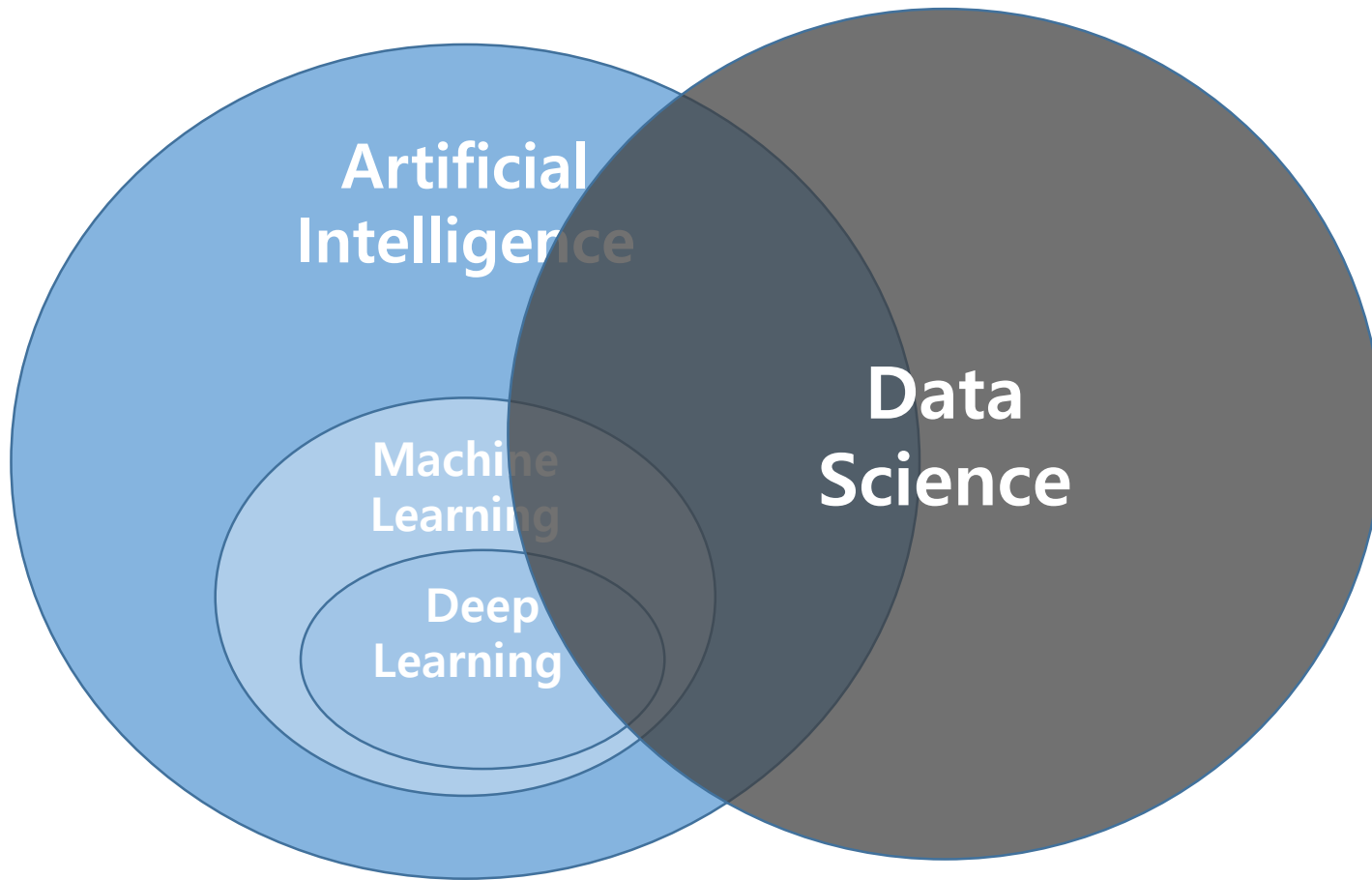
03 Conclusion

ABOUT PRESENTATION

AI and Standardization activities of Data

인공지능 기술의 확산에 따른 데이터 생태계의 변화와
공적(De jure) 국제 표준화 기구(ISO/IEC JTC 1, ITU-T)의 활동 현황 및 전망

01. Data-driven Age



Google trends search(2022.09.13)

Data driven approaches also raise ...

Privacy

Bias

Inequality

Safety and security

Robustness

Trustworthiness

Societal concern

공공데이터포털

이 누리집은 대한민국 공식 전자정부 누리집입니다.

DATA . GO . KR

데이터찾기

AI Hub

AI 데이터찾기 AI 개발지원

AI 나눔

인공지능 응용서비스에 활용 가능한 데이터 자원과
하스데이터 그 밖에 필요한 특용 사용자에게 제공하고

ETRI AI 나눔

데이터 공유 ETRI 기술체험 개발자 커뮤니티 나눔 공모전 더보기

AI 나눔

인공지능 응용서비스에 활용 가능한 데이터 자원과
하스데이터 그 밖에 필요한 특용 사용자에게 제공하고

Open Datasets

Tasks

- 3D Object Detection
- 3D Object Pose
- 3D Reconstruction / Photogrammetry
- 3D Semantic Segmentation
- Autonomous Driving
- Computational Photography

Only Browsable on V7

Annotation Types

- 3D Point Cloud
- Bounding Boxes
- Classification Tags
- Image Pairs
- Instance Segmentation
- Keypoint

Use Cases

- AI Coaching
- AI Referee

COVID-19 Chest X-ray Dataset

Lung segmentation of cases of COVID-19 collected in 2020

Pens Segmentation

Manually and semi-automatically segmented pens for instance segmentation experiments

Kaggle Datasets

Explore, analyze, and share quality data. Learn more about data

Sign In Register

Dataset Search

데이터세트 검색

예시 coronavirus covid-19 또는 water quality site:canada.ca.

데이터 세트 검색에 관해 자세히 알아보세요.

modelzoo.co

github.com/AI

AI | MODEL

Product Solutions

AI

AI Model Share

AIModelShare

Follow

3 followers · 0 following

Achievements

Beta

Send feedback

Block or Report

ModelZoo Frameworks

Filter models...

OpenPose

★ 24000

OpenPose represents real-time multi-person pose estimation (e.g. to jointly detect human pose, hand, and facial keypoint) on images.

Caffe

tensorflow.org/resources/models-datasets

설치 리소스 더보기

검색

Language

GitHub 로그인

모델 및 데이터 세트

도구 라이브러리 및 확장 프로그램 TensorFlow 인증 프로그램 ML 알아보기 책임감 있는 AI

모델 및 데이터 세트

저장소와 기타 리소스를 탐색하여 TensorFlow 커뮤니티에서 만든 모델, 모듈 및 데이터 세트를 찾아보세요.

TensorFlow Hub

즉시 미세 조정 후 어디에나 배포 가능한 학습 모델이 포함된 종합 저장소입니다.

tfhub.dev 살펴보기

모델 모음

TensorFlow의 상위 API를 사용해 개발된 머신러닝 모델 및 예시입니다.

GitHub 살펴보기

TensorFlow.js 모델

클라이언트 측 웹브라우저 또는 Node.js와 같이 자바스크립트를 실행할 수 있는 모든 곳에서 사용 가능한 선행 학습된 머신러닝 모델입니다.

GitHub 살펴보기

01. Data-driven Age

≡ 매일경제

뉴스 오피니언 프리미엄 연예 스포츠 증권 부동산

경제 기업 사회 국제 부동산 증권 정치 IT·과학 문화 기획·연재 Special Edition 인기뉴스 암호화폐 오늘의 매경

디지털시대..."데이터 목록 쌓였지만, 실제 쓸만한 데이터 부족"

입법조사처 "정부 주도 데이터 공급정책 장기적으로 재검토해야"

입력 : 2021.05.12 07:00:05 0



출처: <https://www.mk.co.kr/news/it/view/2021/05/45484>



출처: <https://www.electimes.com/news/articleView.html?idxno=307685> 8

Data standard, standardization?

Target & objectives

02. Standardization activities

← ISO/IEC JTC 1

ISO/IEC JTC 1/SC 42

Artificial intelligence

About

Secretariat: **ANSI**

Committee Manager: **Ms Heather Benko**

Chairperson (until end 2024): Mr Wael William Diab

ISO Technical Programme Manager [TPM]: **Mr Andrew Dryden**

ISO Editorial Manager [EM]: **Ms Christelle Gansonre**

Creation date: 2017

15

published ISO standards *
under the direct responsibility
of ISO/IEC JTC 1/SC 42

24

ISO standards under
development *
under the direct responsibility
of ISO/IEC JTC 1/SC 42

35

Participating members

15

Observing members

02. Activities in JTC 1/SC42

Structure

Liaisons

Meetings

Reference ↑	Title	Type
ISO/IEC JTC 1/SC 42/AG 3 ⓘ	AI standardization roadmapping	Working group
ISO/IEC JTC 1/SC 42/AHG 4 ⓘ	Liaison with SC 27	Working group
ISO/IEC JTC 1/SC 42/JWG 2 ⓘ	Joint Working Group ISO/IEC JTC1/SC 42 - ISO/IEC JTC1/SC 7 : Testing of AI-based systems	Working group
ISO/IEC JTC 1/SC 42/WG 1 ⓘ	Foundational standards	Working group
ISO/IEC JTC 1/SC 42/WG 2 ⓘ	Data	Working group
ISO/IEC JTC 1/SC 42/WG 3 ⓘ	Trustworthiness	Working group
ISO/IEC JTC 1/SC 42/WG 4 ⓘ	Use cases and applications	Working group
ISO/IEC JTC 1/SC 42/WG 5 ⓘ	Computational approaches and computational characteristics of AI systems	Working group

02. Activities in JTC 1/SC42 WG 3

ISO/IEC DTR 5469 Information technology — AI — Functional safety and AI systems

ISO/IEC PWI 17866 Information technology — AI — Guidance for mitigating ethical and societal concerns

ISO/IEC FDIS 23894 Information technology — AI — Guidance on risk management

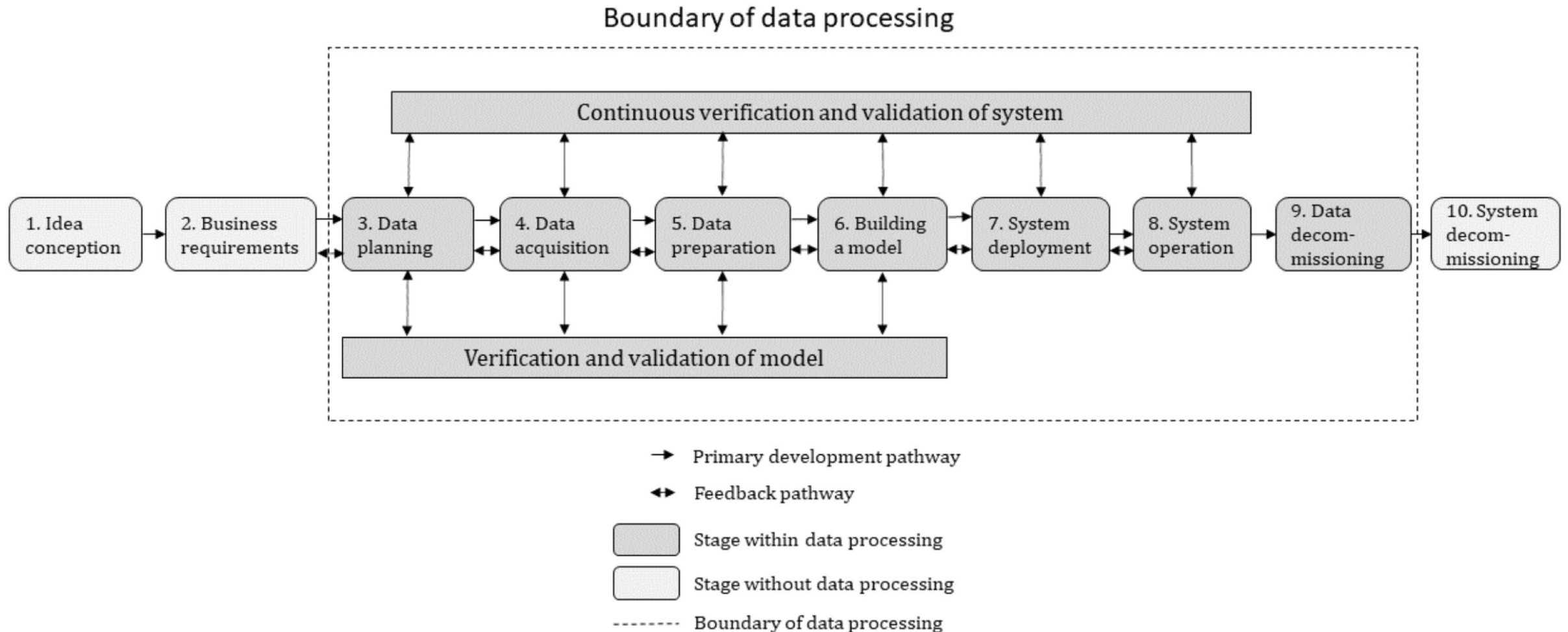
ISO/IEC TR 24027:2021 Information technology — AI — Bias in AI systems and AI aided decision making

ISO/IEC TR 24028:2020 Information technology — AI — Overview of trustworthiness in artificial intelligence

ISO/IEC TR 24368:2022 Information technology — AI — Overview of ethical and societal concerns

02. Activities in JTC 1/SC42 WG 2

ISO/IEC DIS 8183 Information technology - Artificial intelligence - **Data life cycle framework**



02. Activities in JTC 1/SC42 WG 2

ISO/IEC 5259 Artificial intelligence - Data quality for analytics and machine learning

Part 1: Overview, terminology, and examples

Part 2: Data quality measures

Part 3: Data quality manage requirements and guidelines

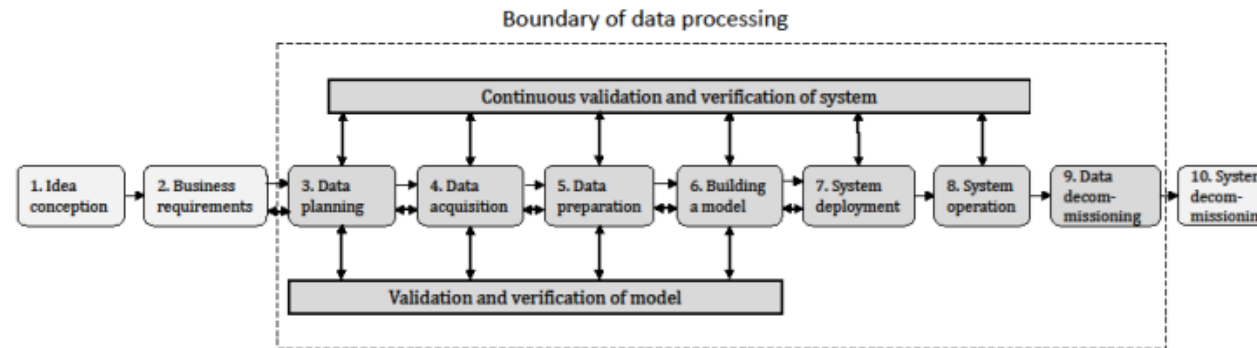
Part 4: Data quality process framework

Part 5: Data quality governance framework

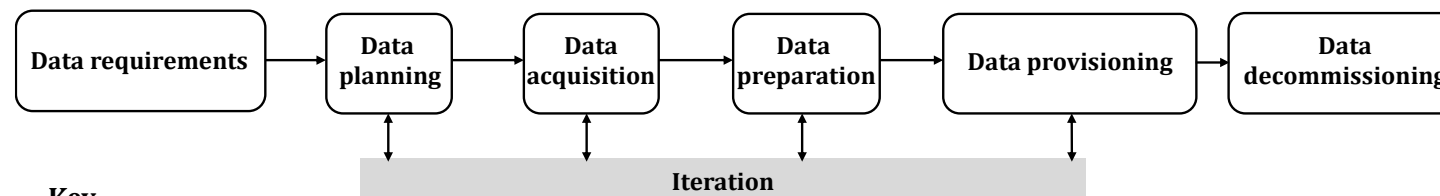
02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-1 Artificial intelligence - Data quality for analytics and machine learning - **Part 1: Overview, terminology, and examples**

Data life cycle framework (from ISO/IEC 1183)



Instantiations of DLC model for 5259 series

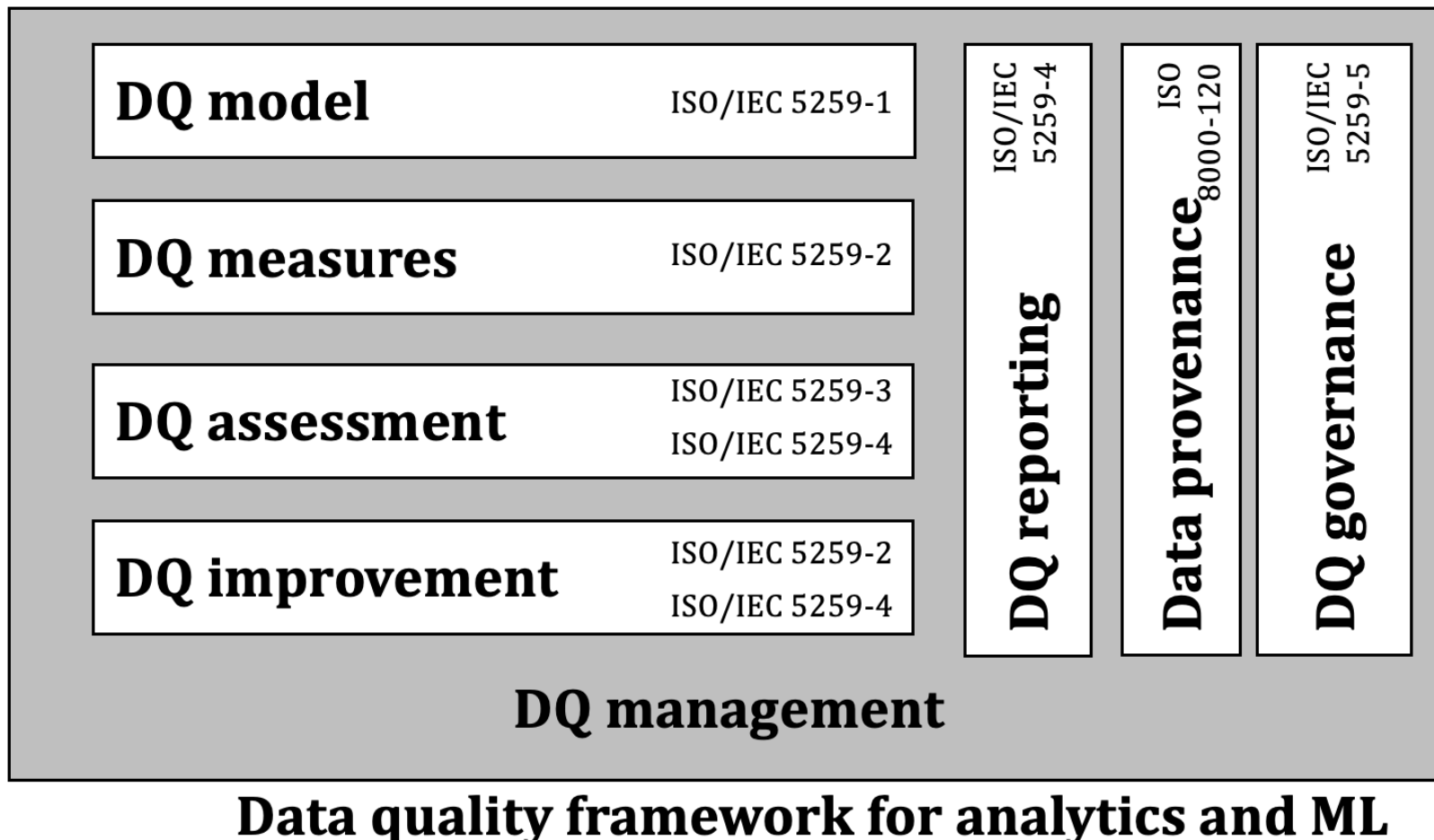


Key

- Stage where data are processed
- Iteration
- Primary development pathway
- Feedback pathway

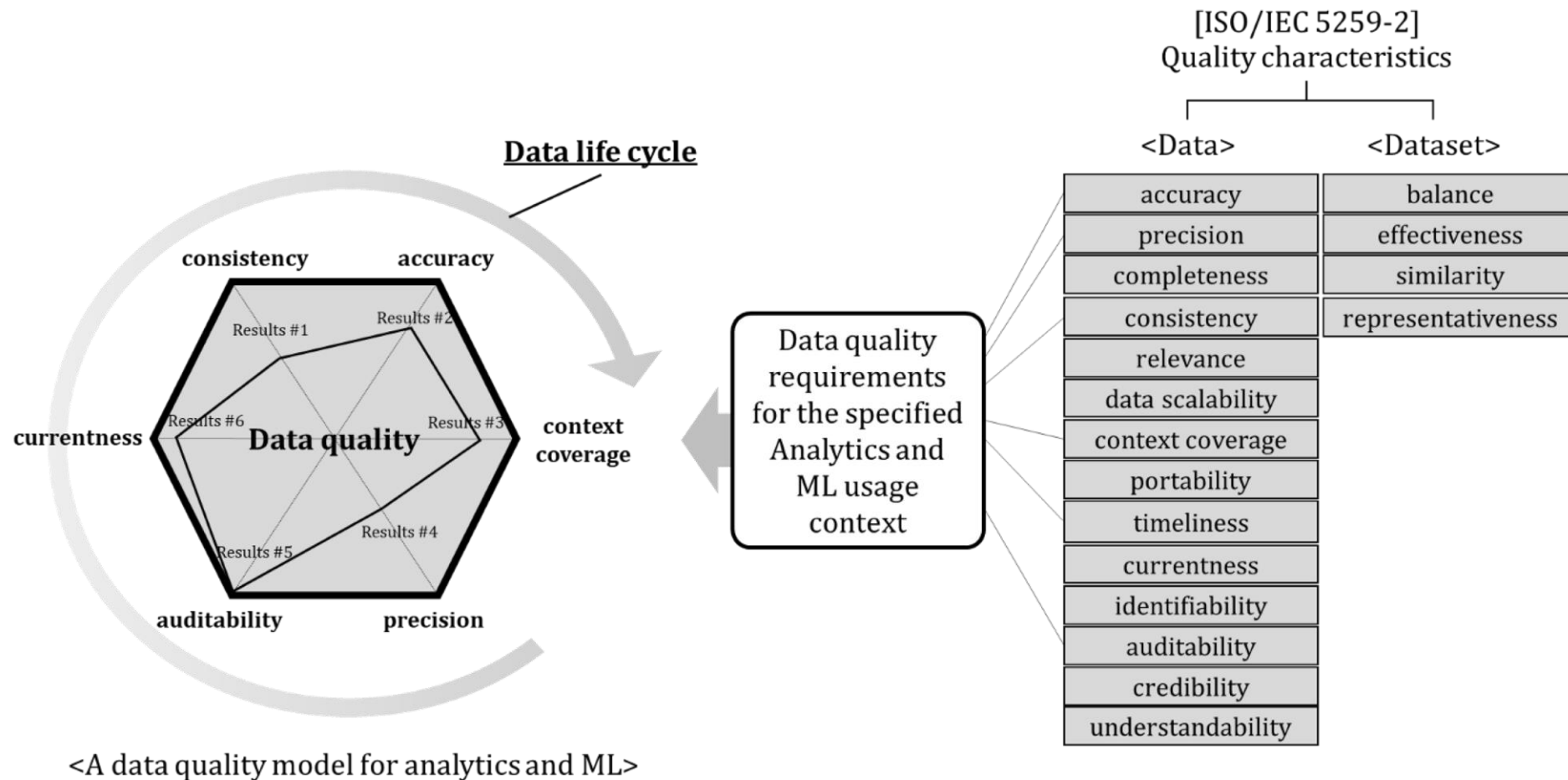
02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-1 Artificial intelligence - Data quality for analytics and machine learning - **Part 1: Overview, terminology, and examples**



02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-1 Artificial intelligence - Data quality for analytics and machine learning - **Part 1: Overview, terminology, and examples**



02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-2 Artificial intelligence - Data quality for analytics and machine learning - Part 2: Data quality measure

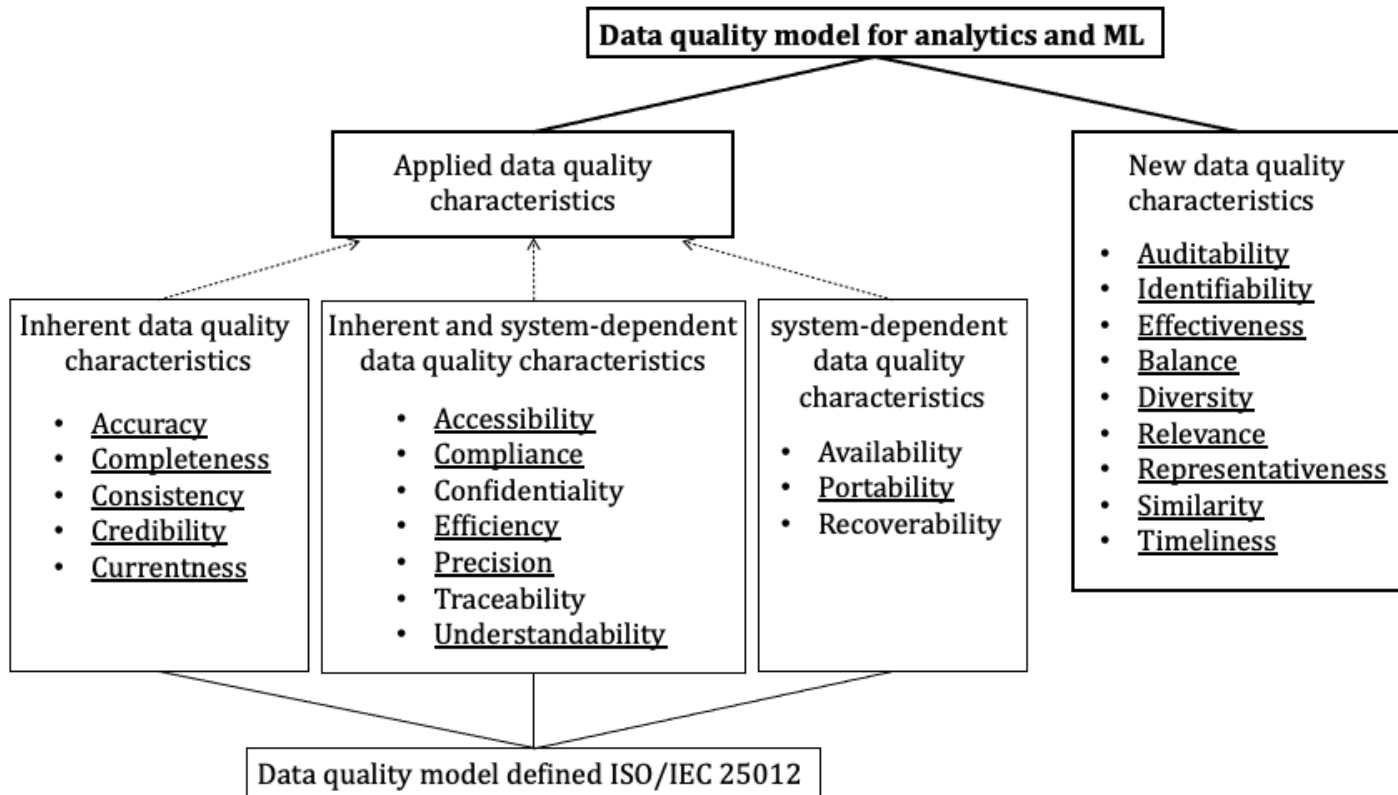


Table 7 — Effectiveness measures

ID	Name	Description	Measurement function	Target data category
Eff-ML-1	Brightness effectiveness	Ratio of samples with unacceptable brightness in a dataset	$\frac{A}{B}$ where A is the number of the samples with unacceptable brightness; B is the number of all samples in the dataset.	Training data Validation data Test data
Eff-ML-2	Resolution effectiveness	Ratio of samples with unacceptable resolution in a dataset	$\frac{A}{B}$ where A is the number of the samples with unacceptable resolution; B is the number of all samples in the dataset.	Training data Validation data Test data
Eff-ML-3	Category size effectiveness	Ratio of <u>categories</u> where the number of categorized samples is lower than a threshold	$\frac{A}{B}$ where A is the number of categories where the number of categorized samples is lower than a threshold; B is the number of categories in total.	Training data Validation data Test data
Eff-ML-4	Bounding box area effectiveness	Ratio of bounding boxes which areas are smaller than an acceptance threshold over all boxes	$\frac{A}{B}$ where A is the number of bounding boxes which areas are smaller than an acceptance threshold;	Training data Validation data Test data

02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-3 Artificial intelligence - Data quality for analytics and machine learning - **Part 3: Data quality manage requirements and guidelines**

▪ 1 Scope↵

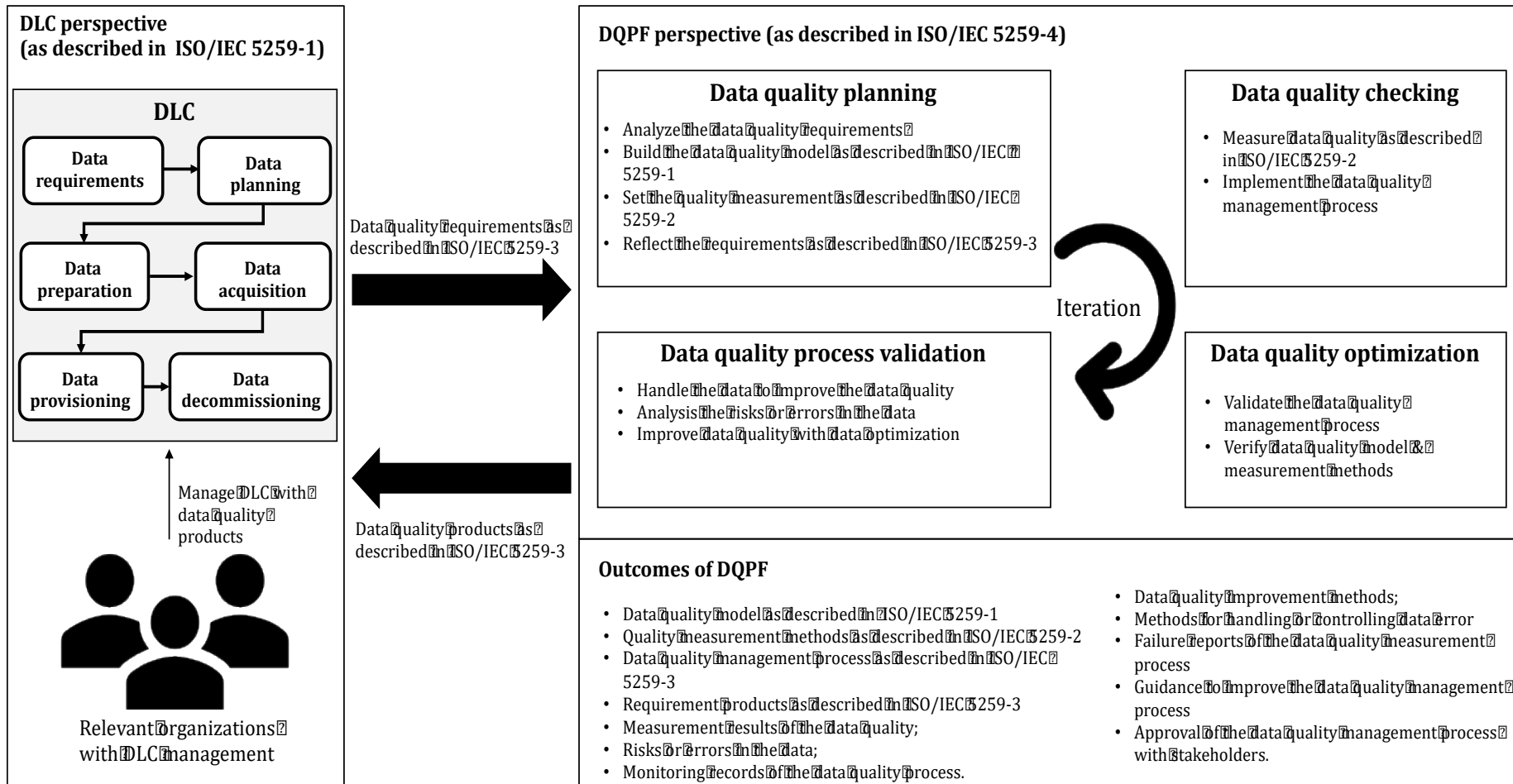
This document specifies requirements and provides guidance for establishing, implementing, maintaining and continually improving the quality for data used in the areas of analytics and machine learning.↵

This document does not define a detailed process, methods or metrics. Rather it defines the requirements and guidance for a quality management process along with a reference process and methods that can be tailored to meet the requirements in this document. ↵

The requirements and recommendations set out in this document are generic and are intended to be applicable to all organizations, regardless of type, size or nature.↵

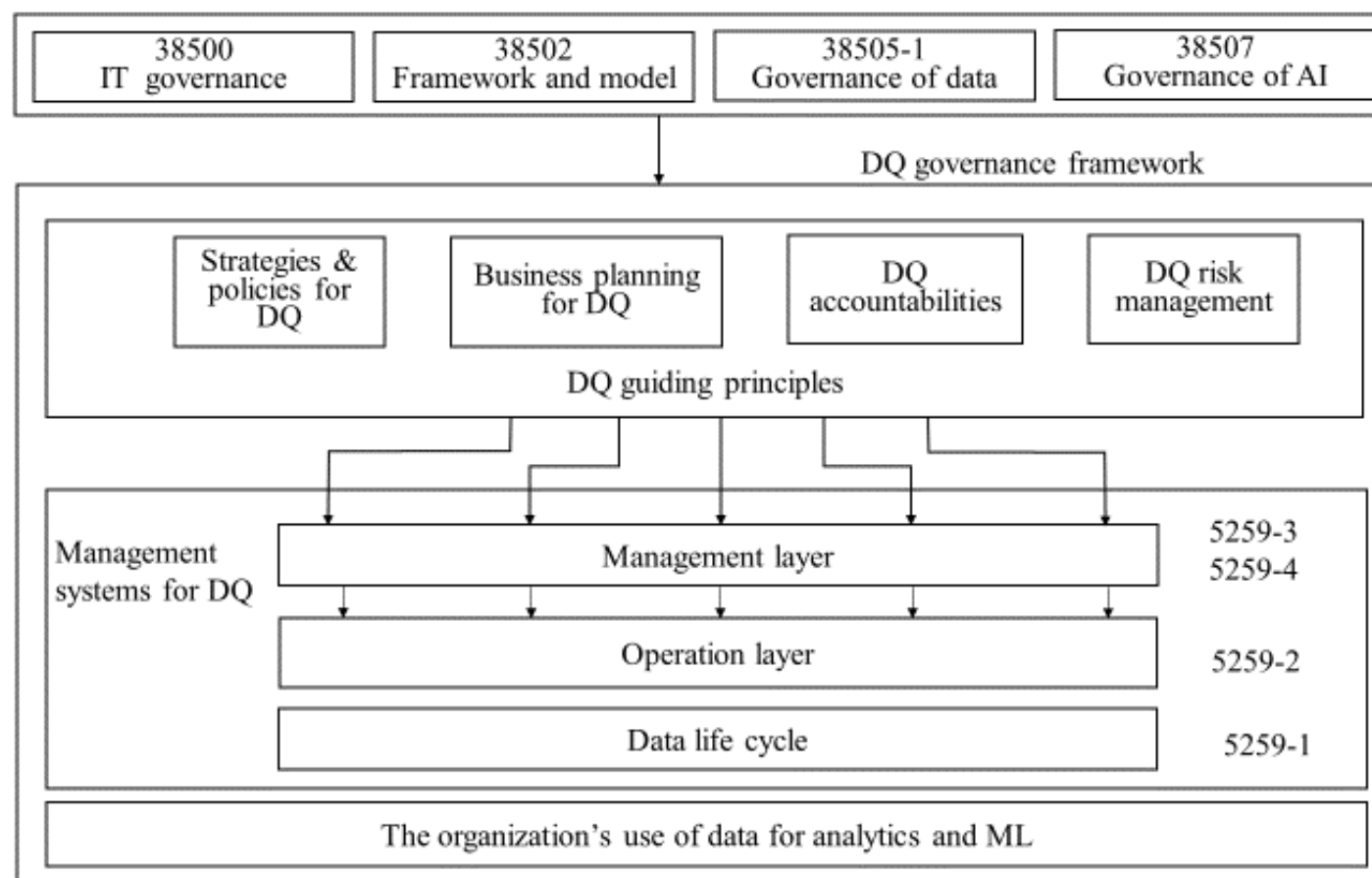
02. Activities in JTC 1/SC42 WG 2

ISO/IEC CD 5259-4 Artificial intelligence - Data quality for analytics and machine learning - Part 4: Data quality process framework



02. Activities in JTC 1/SC42 WG 2

ISO/IEC WD 5259-5 Artificial intelligence - Data quality for analytics and machine learning - **Part 5: Data quality governance framework**



02. Activities in ITU-T SG13

Future networks and emerging network technologies

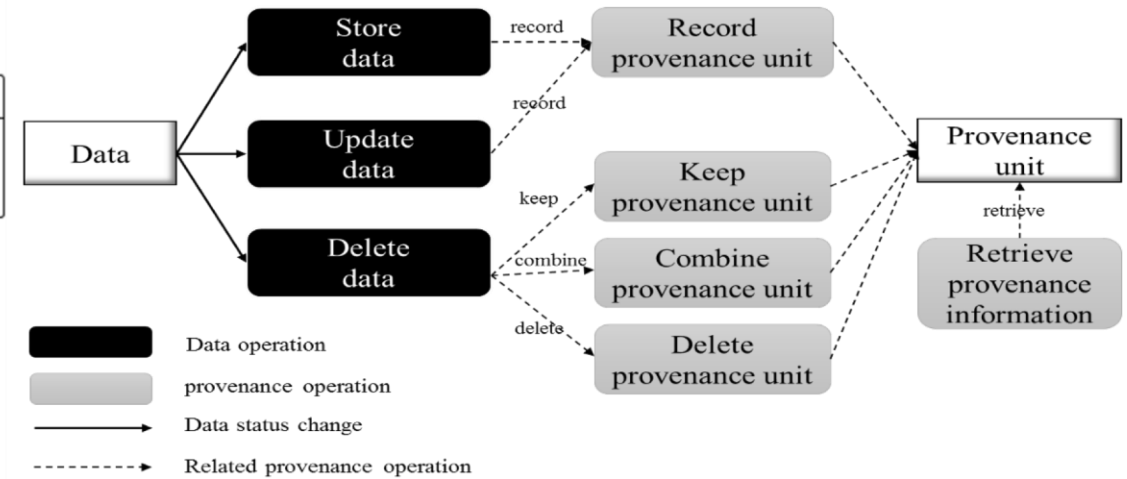
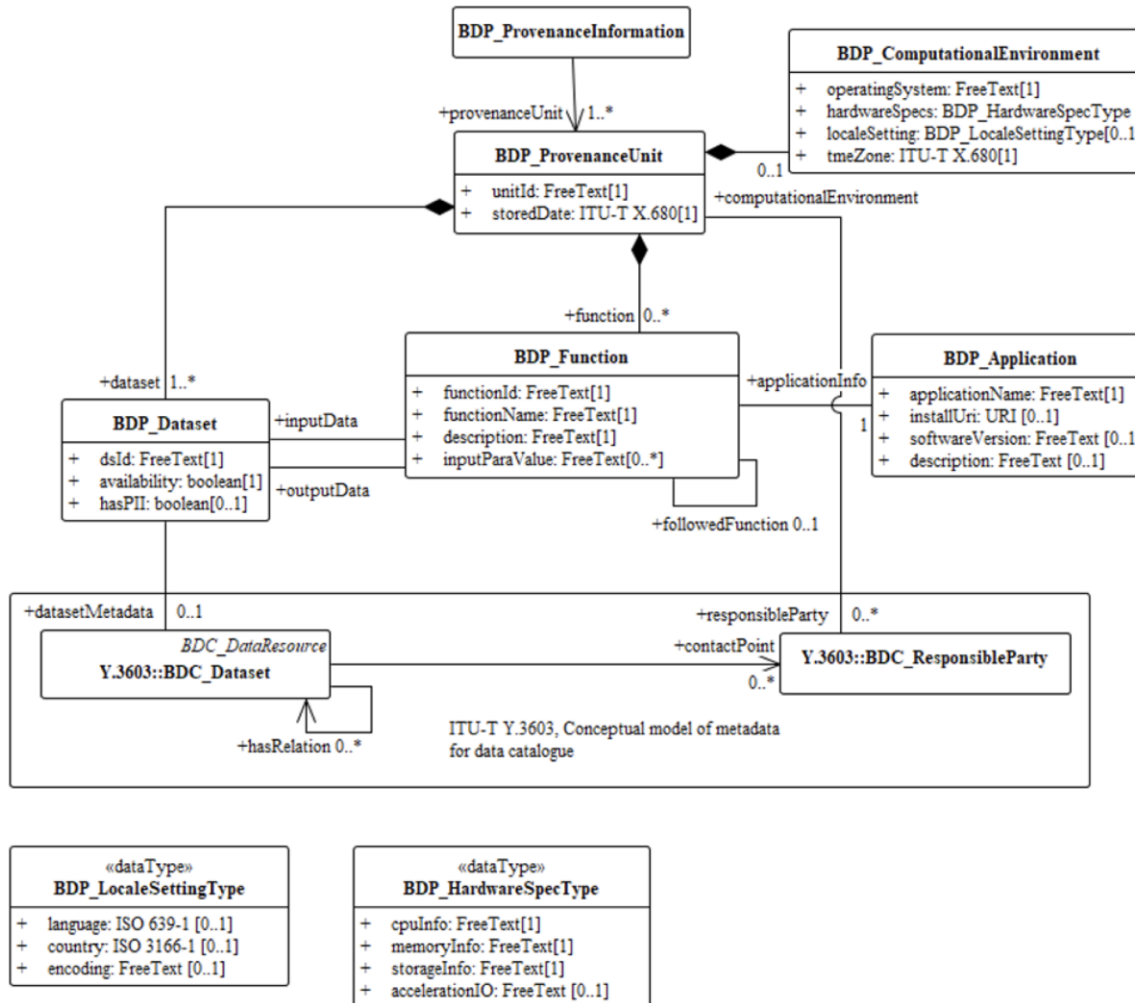
Q17/13 is to provide the necessary overall **frameworks, definitions, and ecosystems** including **requirements, capabilities** related to the integration or support of future computing including cloud computing and **data handling** in telecommunication ecosystem.

Q18/13 deals with the **architectures**, infrastructures and networking views related to the integration and support of the future computing (including cloud computing and data handling) in telecommunication ecosystems.

Q19/13 deals with end-to-end **management, governance**, and security for future computing including cloud computing and data handling from the perspective of telecommunication. The novel methods based on artificial intelligence and machine learning are essential to handle complexity of future computing management and optimally orchestrate its operation and lifecycle management.

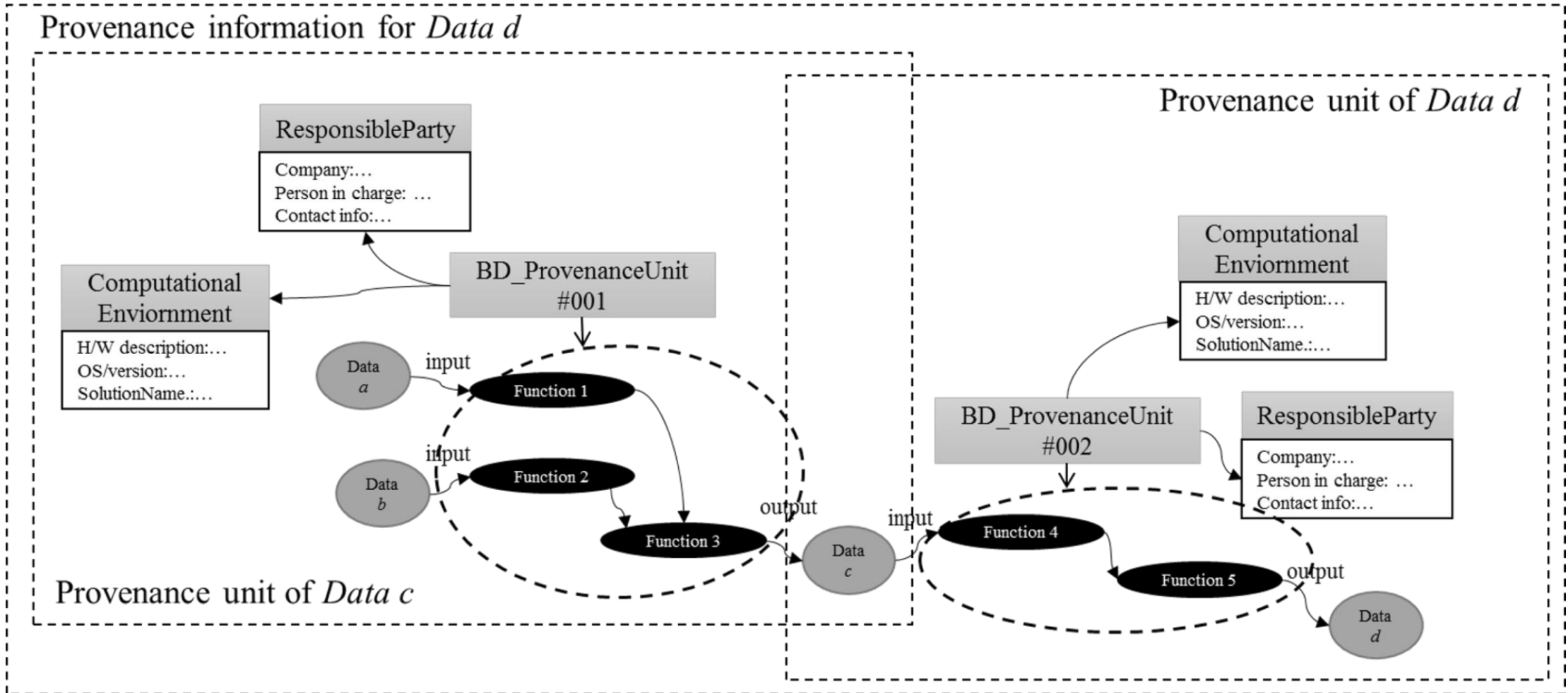
02. Activities in ITU-T SG13

ITU-T Y.3602 Big data - Functional requirements for data provenance (2022)



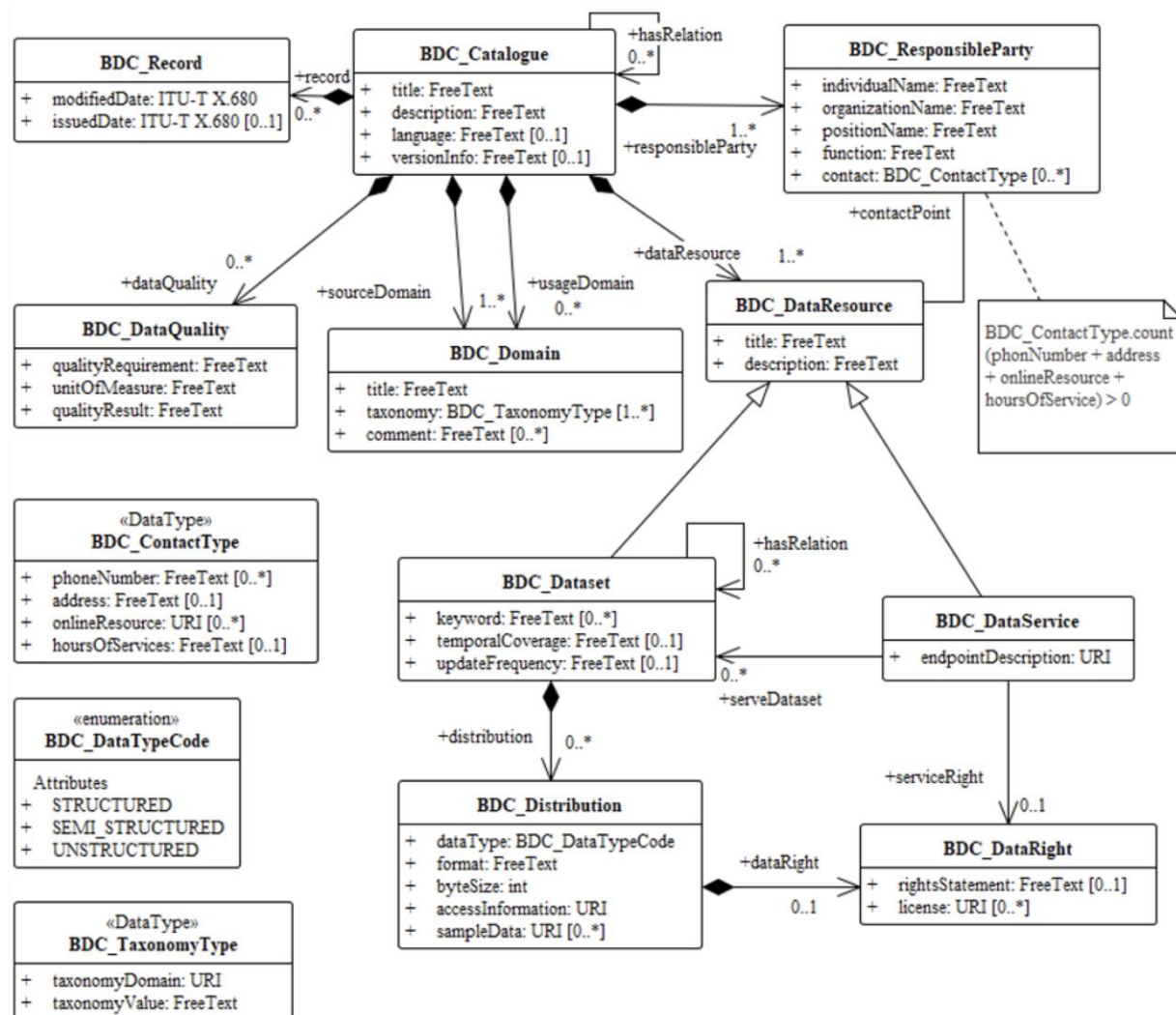
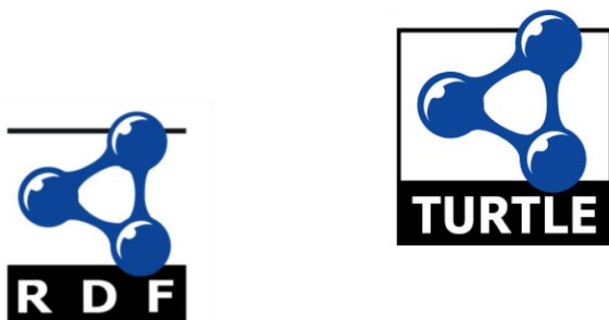
02. Activities in ITU-T SG13

ITU-T Y.3602 Big data - Functional requirements for data provenance (2022)



02. Activities in ITU-T SG13

ITU-T Y.3603 Big data - Requirements and conceptual model of metadata for data catalogue



03. Conclusion

Data life cycle and Data ecosystem are changed.

Data model, format
Interface
Management within organization
Data quality for single application



Support to {Reuse, Share, Find} data fit for user's requirements
Realtime measurement of the context based data quality
Trustworthiness of data
Personal data vs. Privacy

글로벌 ICT 표준 컨퍼런스 2022

Global ICT Standards Conference 2022

Thank you

하수욱 책임연구원, ETRI
E-mail@email.kr