ARTIFICIAL INTELLIGENCE AND DATA INTEROPERABILITY

Standardisation in ISO/IEC and its relation to interoperability

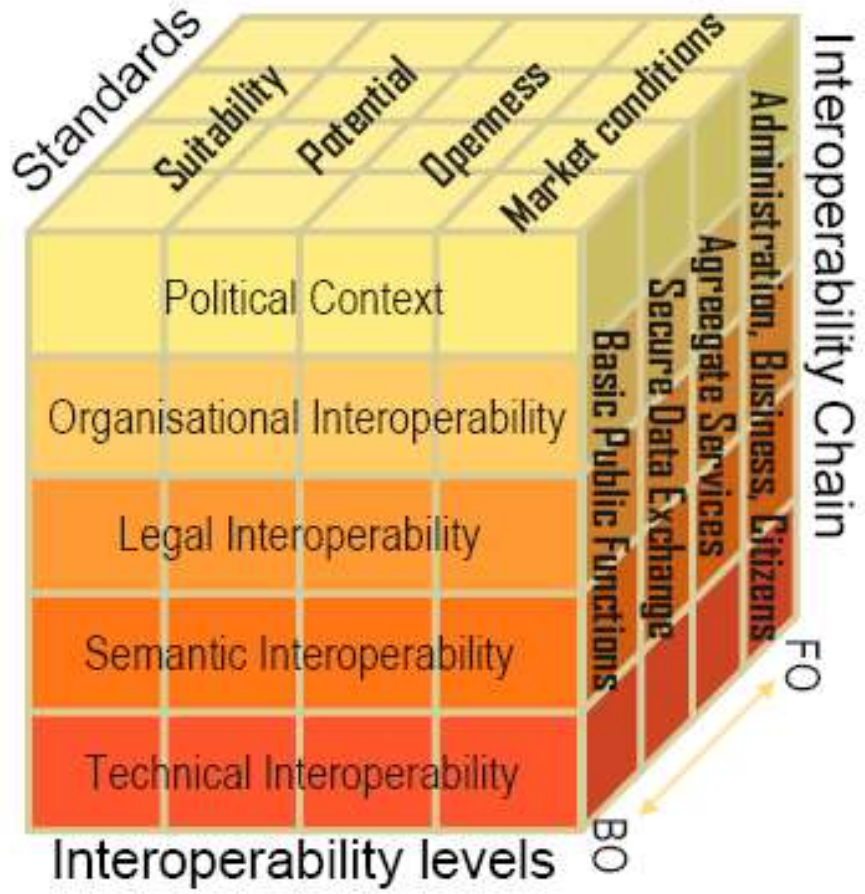
Arne.J.Berre@sintef.no

Leader of BDVA TF6, Norwegian representative in ISO SC42 AI (and Big Data)

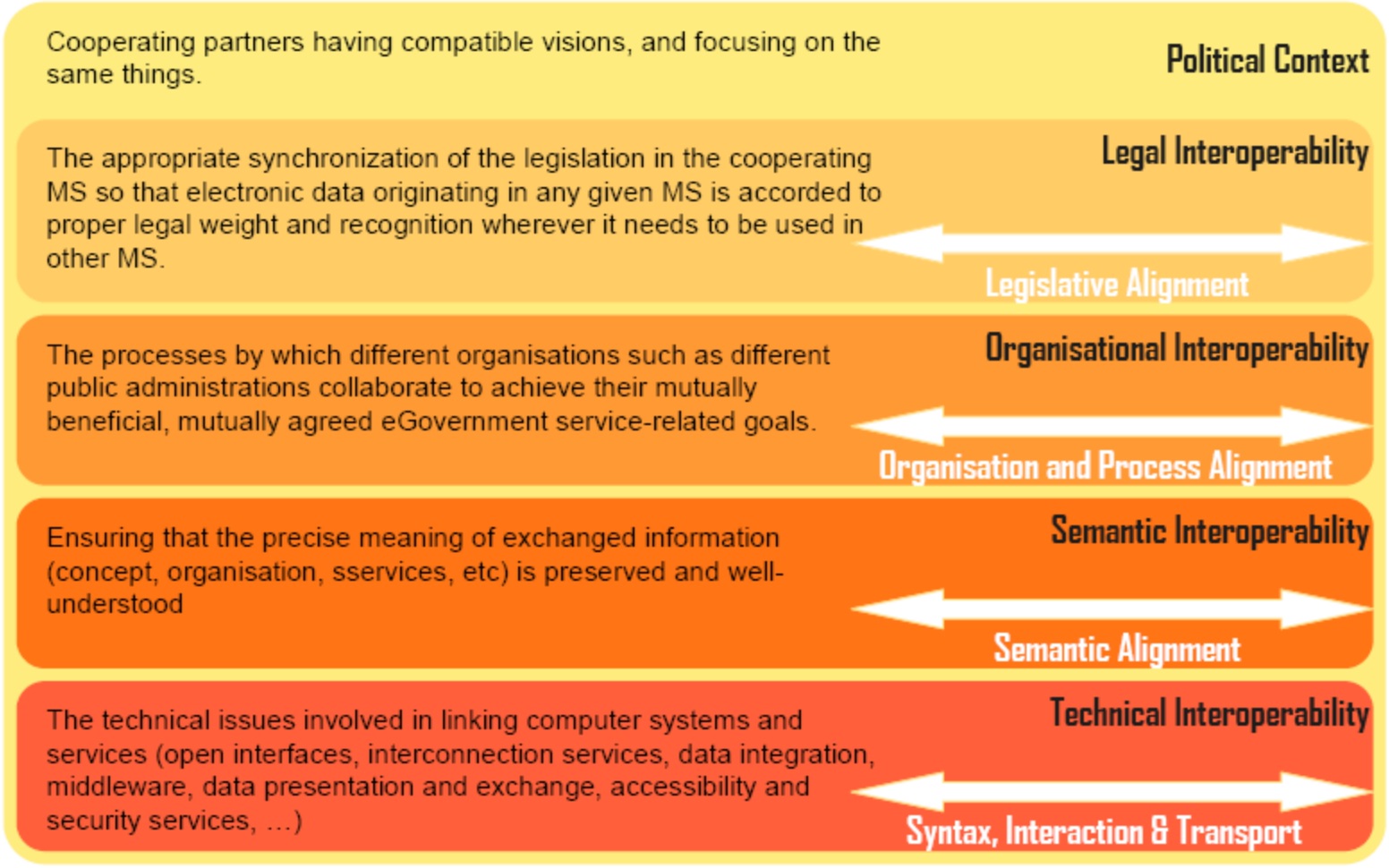
Agenda

- Interoperability – European Interoperability Framework (EIF)
- BDVA – Big Data Value Association – Reference Model - Future AI PPP model - and AIOTI HLA
- Standardisation in ISO/IEC and its relation to interoperability
- ISO SC42 Artificial Intelligence (and Big Data)
- AI Use Cases – DEMETER Agriculture example – The Digital Cow/Farmers Dashboard
- ONNX – Open Neural Network Exchange
- Trustworthiness - incl. European Guidelines for Ethical AI
- AI Data Ecosystem

European Interoperability Framework (EIF) - Dimensions of Interoperability



Interoperability levels

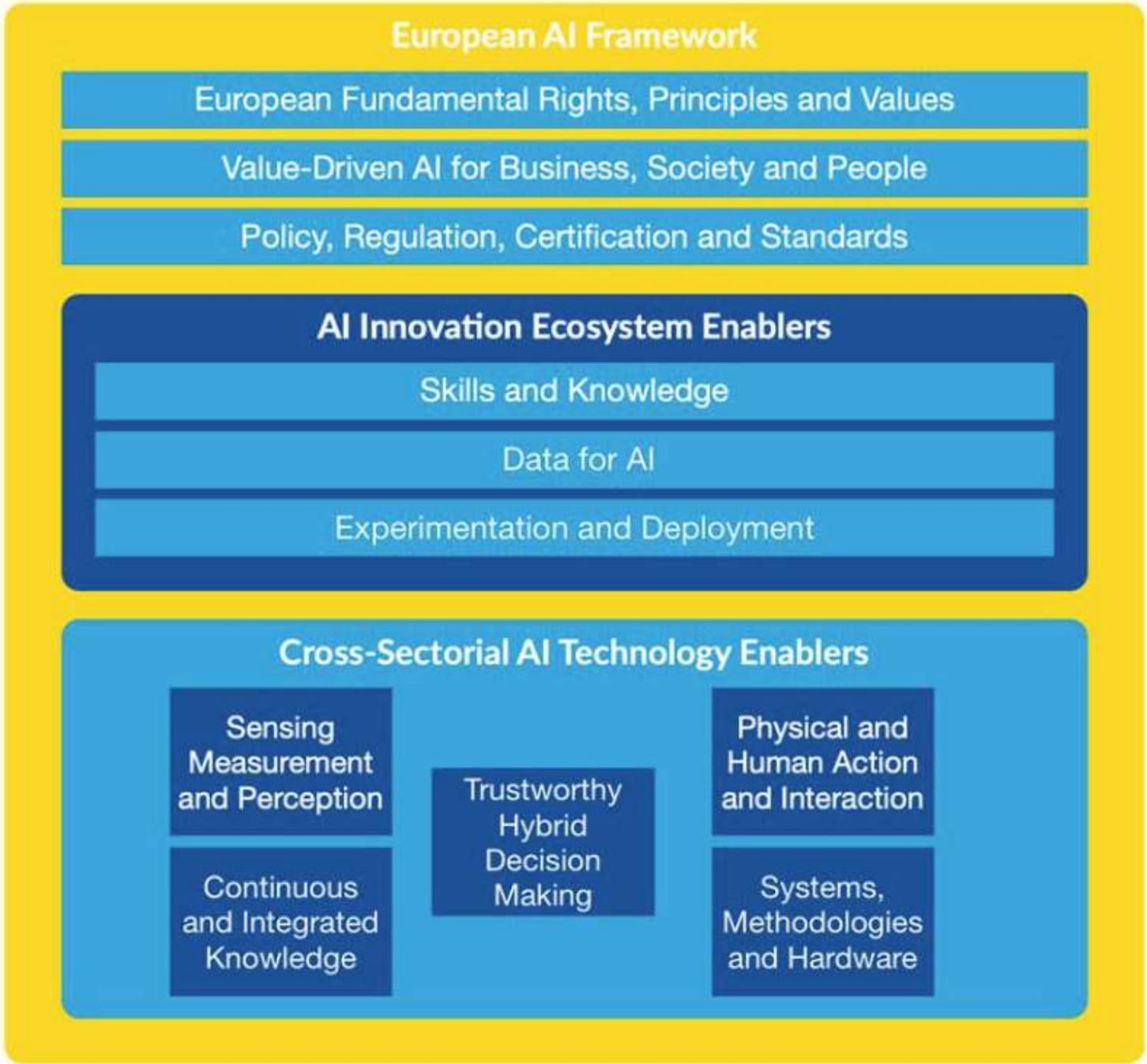


SINTEF

Available at
www.bdva.eu

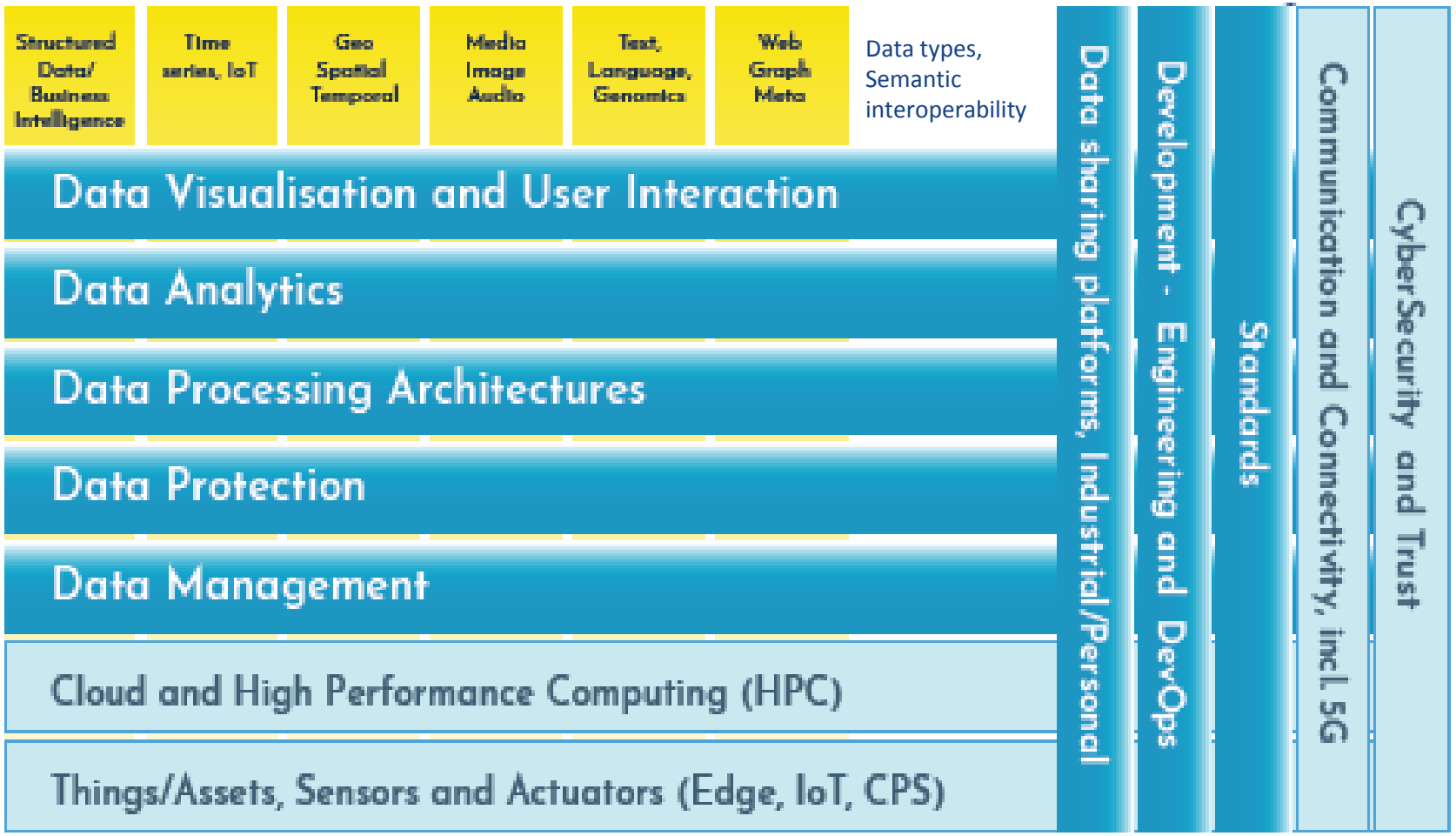
The Vision of the AI Public Private Partnership is to boost European industrial competitiveness and lead the world in developing and deploying value-driven trustworthy AI based on European fundamental rights, principles and values.

SINTEF

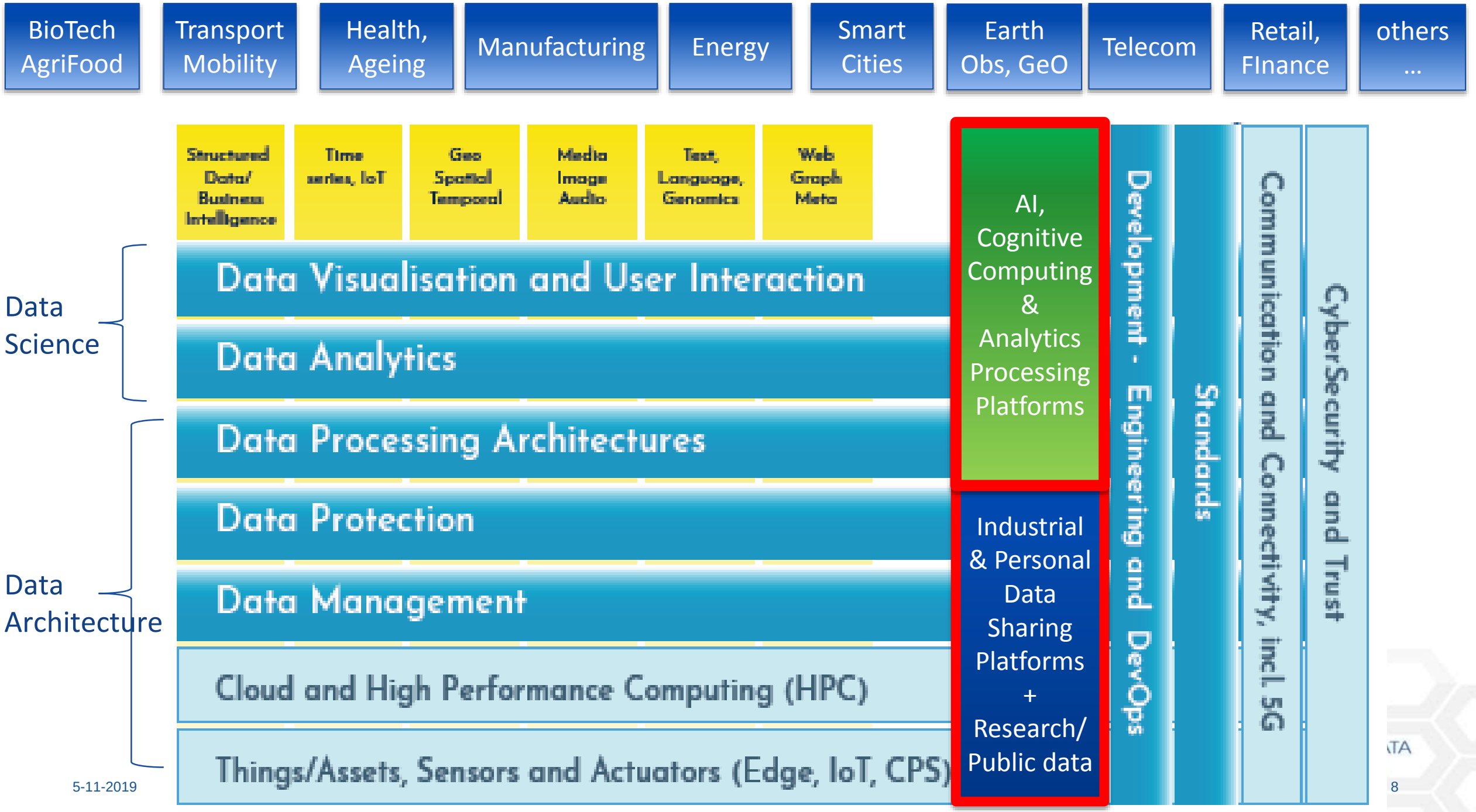

European AI Framework
European Fundamental Rights, Principles and Values
Value-Driven AI for Business, Society and People
Policy, Regulation, Certification and Standards
AI Innovation Ecosystem Enablers
Skills and Knowledge
Data for AI
Experimentation and Deployment
Cross-Sectorial AI Technology Enablers
Sensing Measurement and Perception
Continuous and Integrated Knowledge
Trustworthy Hybrid Decision Making
Physical and Human Action and Interaction
Systems, Methodologies and Hardware

SINTEF

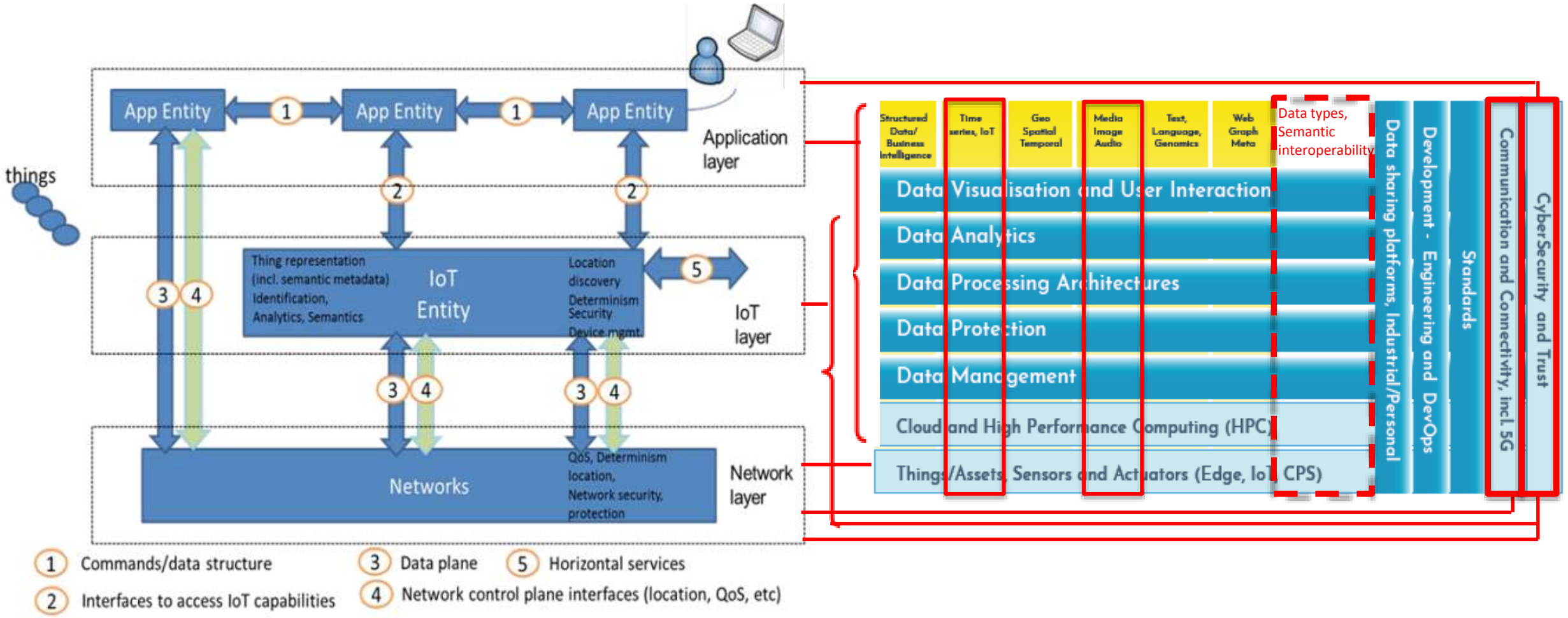
BDV – Big Data and Analytics/Machine Learning Reference Model



SINTEF


BioTech AgriFood
Transport Mobility
Health, Ageing
Manufacturing
Energy
Smart Cities
Earth Obs, GeO
Telecom
Retail, FInance
others …
Structured Data/ Business Intelligence
Time series, IoT
Geo Spatial Temporal
Media Image Audio
Text, Language, Genomics
Web Graph Meta
Data Visualisation and User Interaction
Data Analytics
Data Processing Architectures
Data Protection
Data Management
Cloud and High Performance Computing (HPC)
Things/Assets, Sensors and Actuators (Edge, IoT, CPS)
AI, Cognitive Computing & Analytics Processing Platforms
Industrial & Personal Data Sharing Platforms + Research/ Public data
Development - Engineering and DevOps
Standards
Communication and Connectivity, incl. 5G
CyberSecurity and Trust
Data Science
Data Architecture

BDV Reference Model mapped into the AIOTI High Level Architecture (HLA)



ISO/IEC JTC 1

ISO/IEC JTC 1/SC 42

Artificial intelligence

ABOUT

SECRETARIAT: ANSI

Committee Manager: Ms Heather Benko

Chairperson (until end 2021): Mr Wael William Diab

ISO Technical Programme Manager [TPM]: Mr José Alcorta
ISO Editorial Programme Manager [EPM]: Mr Bastien Gavoille

Creation date: 2017

SCOPE

Standardization in the area of Artificial Intelligence

- Serve as the focus and proponent for JTC 1's standardization program on Artificial Intelligence
- Provide guidance to JTC 1, IEC, and ISO committees developing Artificial Intelligence applications

QUICK LINKS

WORK PROGRAMME
Drafts and new work items

WORKING AREA
on ISOTC and Public information folder

ISO ELECTRONIC APPLICATIONS
IT Tools that help support the standards development process

Use cases and applications (ISO SC42 WG4)

WG 4 - Use cases and applications

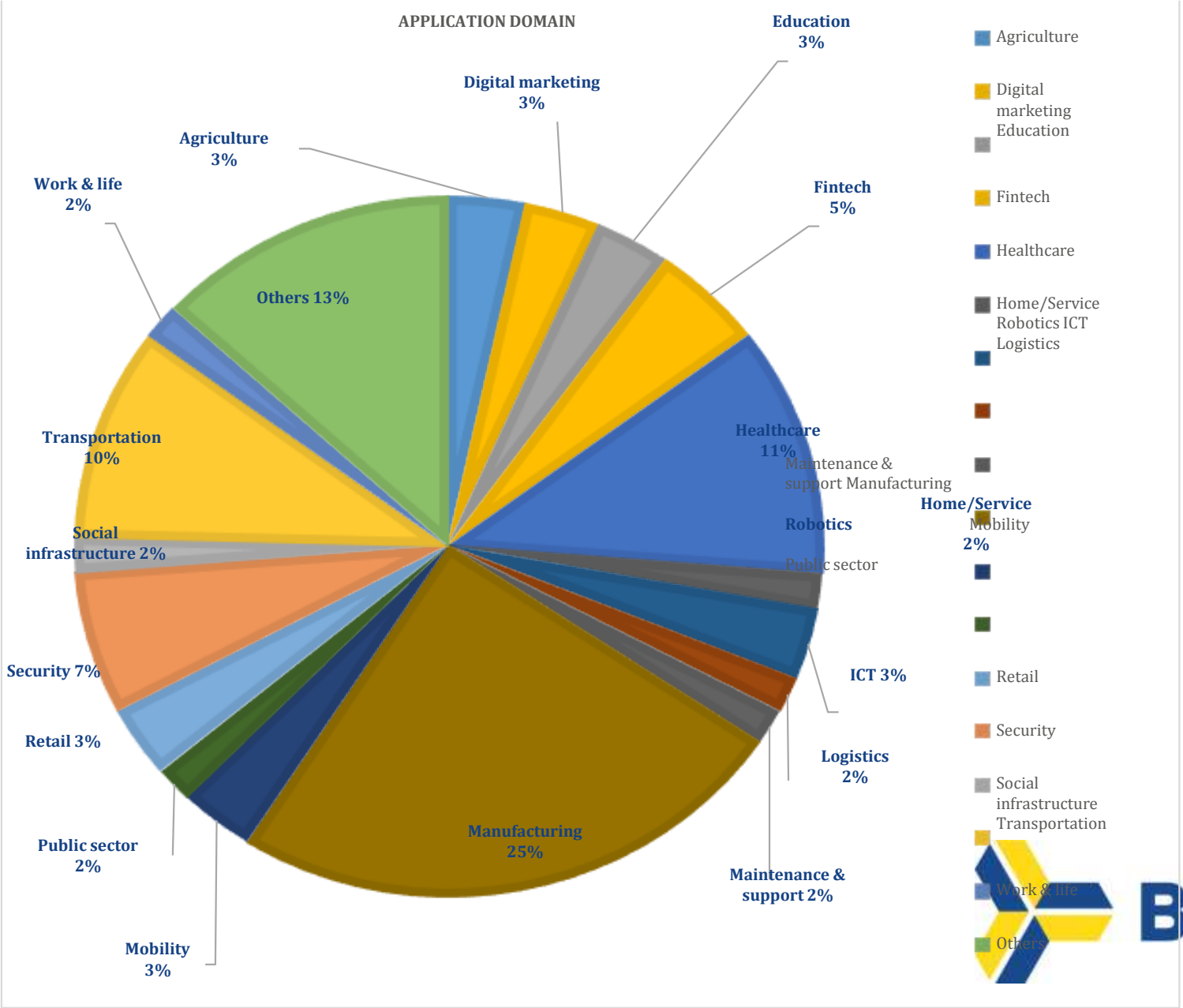
- ToR: Use cases and applications for AI standardization
- ISO/IEC TR 24030: AI Use cases [New]
  - Identify different AI applications domains and the different context and their use
  - 50 use cases submitted → **Request for more use cases submission to BDVA according to the SC 42 template**

JTC 1/SC 42 is developing a Technical Report collecting use cases relative to Artificial Intelligence. With this liaison statement, we request that your organization provide us with your relevant AI use cases, using the attached Word template (contained in SC 42 N 315), by 31 August 2019 for consideration by the committee.

For your information, please find attached the first WD of ISO/IEC TR 24030, which includes a number of use cases.

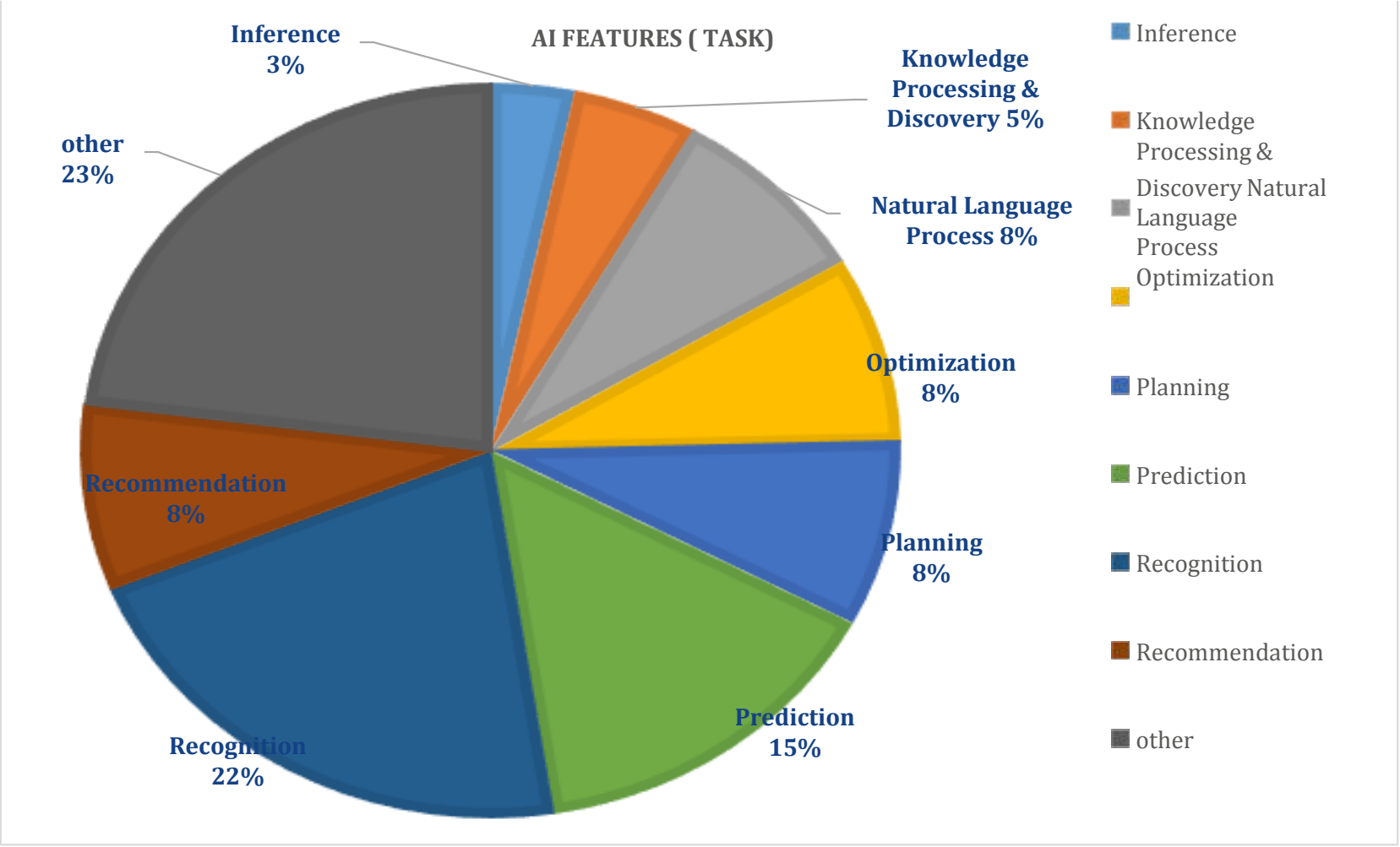
Thank you for your assistance.

WG4



- Media
- Energy
- Maritime
- Telecom
- ....

BDV BIG DATA VALUE ASSOCIATION


AI FEATURES (TASK)
Inference 3%
Knowledge Processing & Discovery 5%
Natural Language Process 8%
Optimization 8%
Planning 8%
Prediction 15%
Recognition 22%
Recommendation 8%
other 23%
Inference
Knowledge Processing & Discovery Natural Language Process Optimization
Planning
Prediction
Recognition
Recommendation
other

AI Use case Template

ID				
Use case name				
Application domain				
Deployment Model				
Status				
Scope				
Objective(s)				
Narrative	Short description (not more than 150 words)			
	Complete description			
Stakeholders				
Stakeholders' assets, values				
System's threats & vulnerabilities				
Key performance indicators (KPIs)	ID	Name	Description	Reference to mentioned use case objectives
AI features	Task(s)			
	Method(s)			
	Hardware			
	Topology			
	Terms and concepts used			
Standardization opportunities/ requirements				
Challenges and issues				
Societal	Description			

ID	Use case name	Application domain	Deployment model	Status
1	Explainable artificial intelligence for Genomic Medicine	Healthcare	Cloud services	Prototype
2	Revolutionizing clinical decision-making using artificial intelligence	Healthcare	On-premise systems	PoC
3	AI solution to calculate amount of contained material from mass spectrometry measurement data	Manufacturing	Embedded systems	PoC
4	AI solution to quickly identify defects during quality assurance process on wind turbine blades	Manufacturing	On-premise systems	In operation
5	Solution to detect signs of failures in wind power generation system	Manufacturing	On-premise systems	PoC
6	Computer-aided diagnosis in medical imaging based on machine learning	Healthcare	On-premise systems	PoC
7	AI Ideally Matches Children to Daycare Centers	Public sector	On-premise systems	In operation
8	Deep Learning Technology Combined with Topological Data Analysis Successfully Estimates Degree of Internal Damage to Bridge Infrastructure	Social infrastructure	Cloud services	PoC
9	AI Components for Vehicle Platooning on Public Roads	Transportation	Self-driving vehicles	Prototype
10	Self-Driving Aircraft Towing Vehicle	Transportation	Self-driving vehicles	Prototype
11	Unmanned Protective Vehicle for Road Works on Motorways	Transportation	Self-driving vehicles	Prototype
12	autonomous apron truck	Mobility	Embedded systems	PoC
13	AI solution to identify automatically false positives from a specific check for	Other (please specify) This will be relevant	Cloud services	PoC

14	Behavioural and sentiment analytics	Security	On-premise systems	PoC
15	Generative design of mechanical parts	Manufacturing	On-premise systems	In operation
16	Robotic prehension of objects	Other (please specify) Robotics	Embedded systems	PoC
17	Robotic vision - scene awareness	Work & life	Embedded systems	PoC
18	AI solution for Car Damage Classification	Other (Insurance)	Cloud services	PoC
19	AI to understand adulteration in commonly used food items	Agriculture	Cloud services	PoC
20	Detection of frauds based on collusions	Fintech	On-premise systems	In operation
21	Information Extraction from Hand-marked Industrial Inspection Sheets	Manufacturing	Cloud services	PoC
22	AI (Swarm Intelligence) solution for Attack Detection in IoT Environment	Security	Hybrid or other (Agent Based Hub-Spoke)	Prototype
23	VTrain recommendation engine	Education	On-premise systems	In operation
24	AI solution to predict Post-Operative Visual Acuity for LASIK Surgeries	Healthcare	Cloud services	In operation
25	Use of robotic solution for traffic policing and control	Security	On-premise systems	PoC
26	Robotic solution for replacing human labour in Hazardous condition	Security	On-premise systems	PoC
27	Credit scoring using KYC data	Fintech	On-premise systems	PoC
28	Recommendation algorithm for improving member experience and discoverability of resorts in the booking portal of a hotel chain	Healthcare	Cloud services	In operation
29	Enhancing traffic management efficiency and infraction detection accuracy with AI technologies	Transportation	Hybrid or other (please specify) Cloud services and on-premise systems	In operation
30	Autonomous network and automation level definition	ICT	Cyber-physical systems	PoC
31	Autonomous network scenarios	ICT	Cyber-physical systems	PoC
32	AI solution to help mobile phone to have better picture effect	Mobility	Hybrid or other (please specify)	In operation
33	Automated defect classification on product surfaces	Manufacturing	On premise system	PoC
34	Robotic task automation: Insertion	Manufacturing	Embedded systems – Cloud service	PoC
35	Causality-based Thermal Prediction for Data Center	Other (data center)	On-premise systems	Prototype
36	Powering Remote Drilling Command Centre	Manufacturing	Cloud services	In operation
37	Leveraging AI to enhance adhesive quality	Manufacturing	On-premise systems	In operation

38	Machine learning driven approach to identify the weak spots in the manufacturing of the circuit breakers.	Manufacturing	Prototype	On-premise system
39	Machine Learning Driven Analysis of Batch Process Operation Data to Identify Causes for Poor Batch Performance	Manufacturing	On-premise systems	Prototype
40	Empowering Autonomous Flow meter control- Reducing time taken to "proving of meters"	Manufacturing	Cloud services	In operation
41	Improving Productivity for Warehouse Operation	Logistics	On-premise systems	PoC
42	Emotion-sensitive AI Customer Service	Retail	On-premise systems	In operation
43	Deep Learning Based User Intent Recognition	Retail	On-premise systems	In operation
44	Chromosome Segmentation and Deep Classification	Healthcare	Hybrid or other (please specify)	PoC
45	Anomaly Detection in Sensor Data Using Deep Learning techniques	Maintenance & support	Hybrid or other (Cloud or on premise deployment)	PoC
46	Adaptable Factory	Manufacturing	Cyber-physical System, Embedded System	PoC
47	Order-Controlled Production	Manufacturing	Cloud Services	Prototype
48	Value-based Service	Manufacturing	Hybrid or other (Cloud or on premise deployment)	PoC
49	AI solution for traffic signal Optimization based on multi-source data fusion	Transportation	Cloud services	In operation
50	AI solution to quality control of Electronic Medical Record(EMR) in real time	Healthcare	Cloud services	In operation
51	Machine Learning Tools in Support of Transformer Diagnostics	Performance evaluation and diagnostics	Prototype	Prototype
52	Automated Travel Pattern Recognition using Mobile Network Data for Applications to Mobility as a Service	Transportation	Activity- based Modelling for New mobility Services	PoC
53	Improving conversion rates and RoI (Return on Investment) with AI technologies	Digital marketing	On-premise	In operation
54	bioBotGuard	Agriculture	Hybrid or other (Cloud or on premise deployment)	PoC
55	RAVE	Education	Hybrid or other (Cloud or on premise deployment)	PoC
56	Logo and Trademark Detection	Digital Marketing	Hybrid or other (Cloud or on premise deployment)	PoC
57	Virtual Bank Assistant	Fintech	Cloud services	In operation

58	Video on Demand Publishing Intelligence Platform	TMT Industry, Technology Department	On premise	In operation
59	Predictive Testing	TMT Industry – Application development	On premise	PoC
60	Predictive Data Quality	Other (please specify) Data Management	Hybrid or other (Cloud or on premise deployment)	PoC
61	Robot consciousness	Home/Service Robotics	Embedded systems	PoC

Some sectors/domains not currently covered:

- Media
- Energy
- Building and Construction
- Ocean/Maritime
- Aquaculture / Fishery
- …

... Example AI use cases ….

demeter

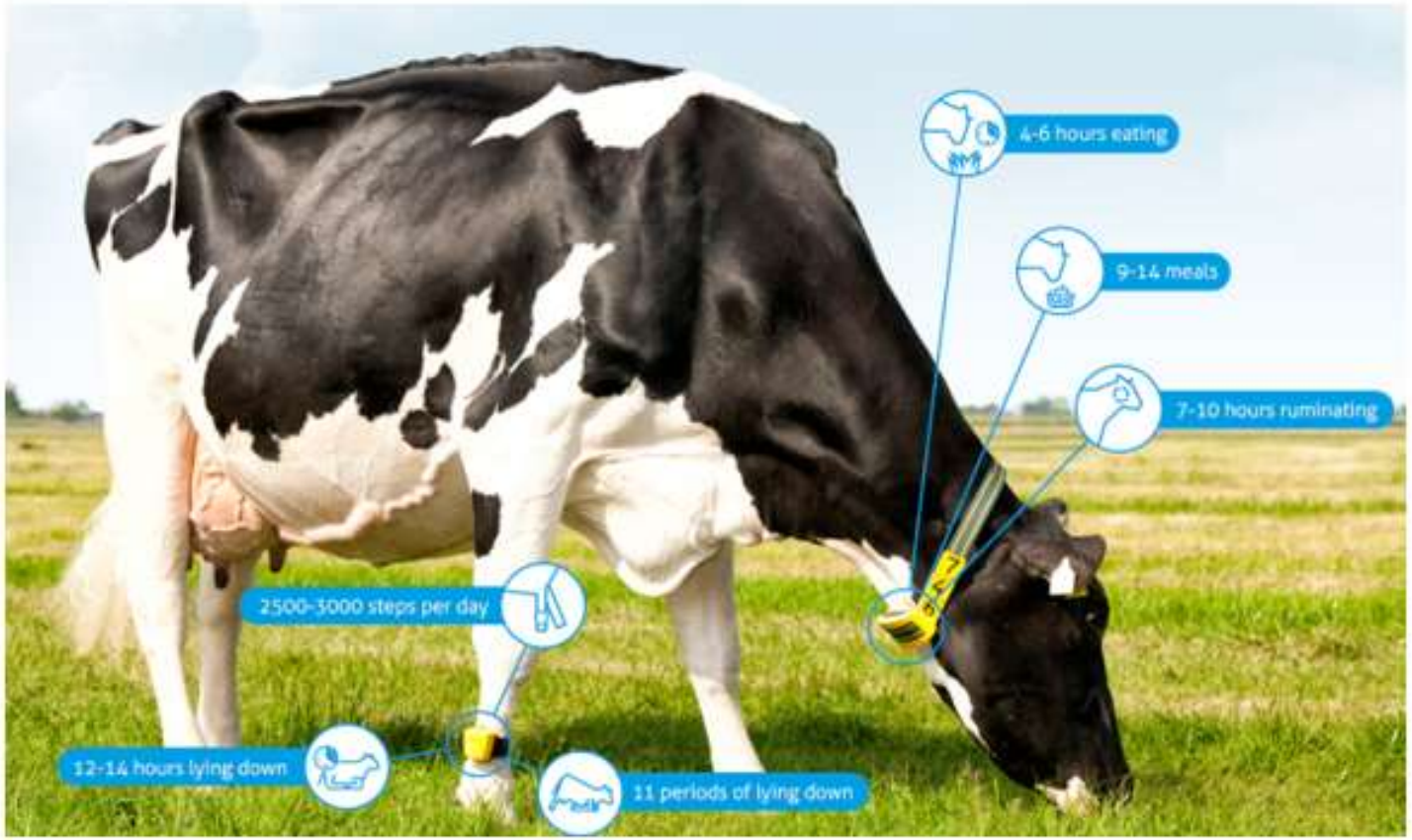
Building an Interoperable, Data-Driven, Innovative & Sustainable European Agri-Food Sector

DEMETER

September 2019 – September 2022

60 partners
15 Meuro funding

SINTEF



The Digital Cow ….

Feeding equipment sensors

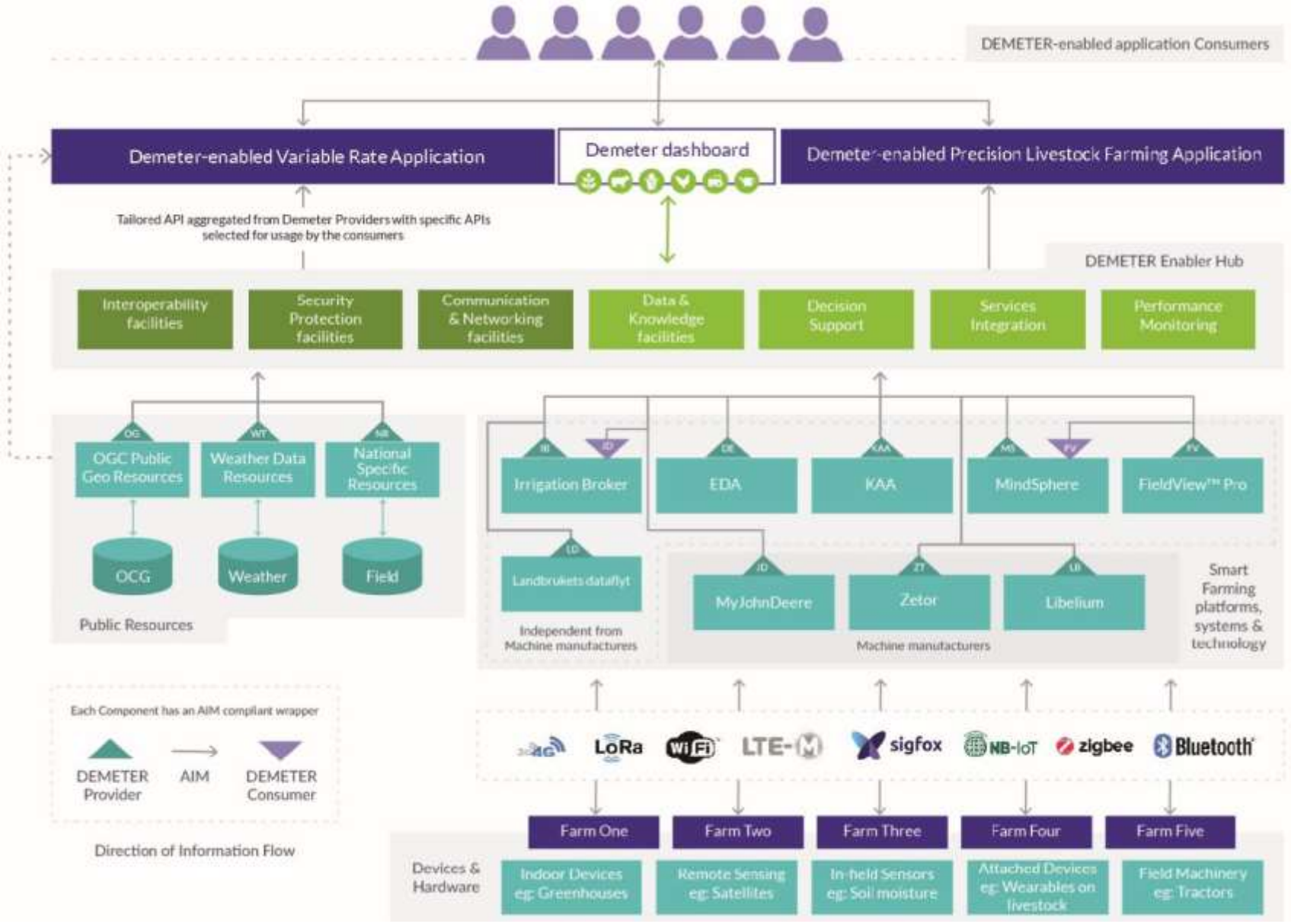
Cow health sensors

Milk fat sensor + other sensors/sources

Milking robots

SINTEF

DEMETER - Agriculture



SINTEF

Advisor

Farmer

Consignee
Supplier

Bank

Accountant

Farmers Dashboard

for the entire milk and meat production value chain

Benchmarking Services/apps

Farm / advisory management

Machine learning Based prediction applications

Authentication
Authorisation

Tailored API aggregated from Demeter Providers with specific APIs selected for usage by the Consumers

Demeter Dashboard

Demeter-enabled Precision Application

Demeter Enabler Hub

Interoperability facilities | Security Protection facilities | Communication & Networking facilities | Data & Knowledge facilities | Decision Support | Services Integration | Performance Monitoring

Open API's

Norway National Eco-system

National Registries

National Specific Resources

MyJohnDeere™ myjohndeere.deere.com

Public data

Field

Farmers registrations

Agricultural Dataflow system

Mimiro Farmers Data platform

Milk Production (Tine)

Meat Production (Nortura) ++

Invoicing & Accounting (Duett ++)

Insurance Bank (Gjensidige)

Livestock systems

Data quality control w/Machine learning

Farm Products (Felleskjøpet)

Communication/Networking infrastructure (Low Power Wide Area or Cellular Networks)

NB-IoT | LoRa | WiFi | sigfox | zigbee | LTE-M | 3G | 4G | Bluetooth

Weather data
Soil data

Objects
Farmer's
IoT

Feeding equipment sensors

Cow health sensors

Milk fat sensor

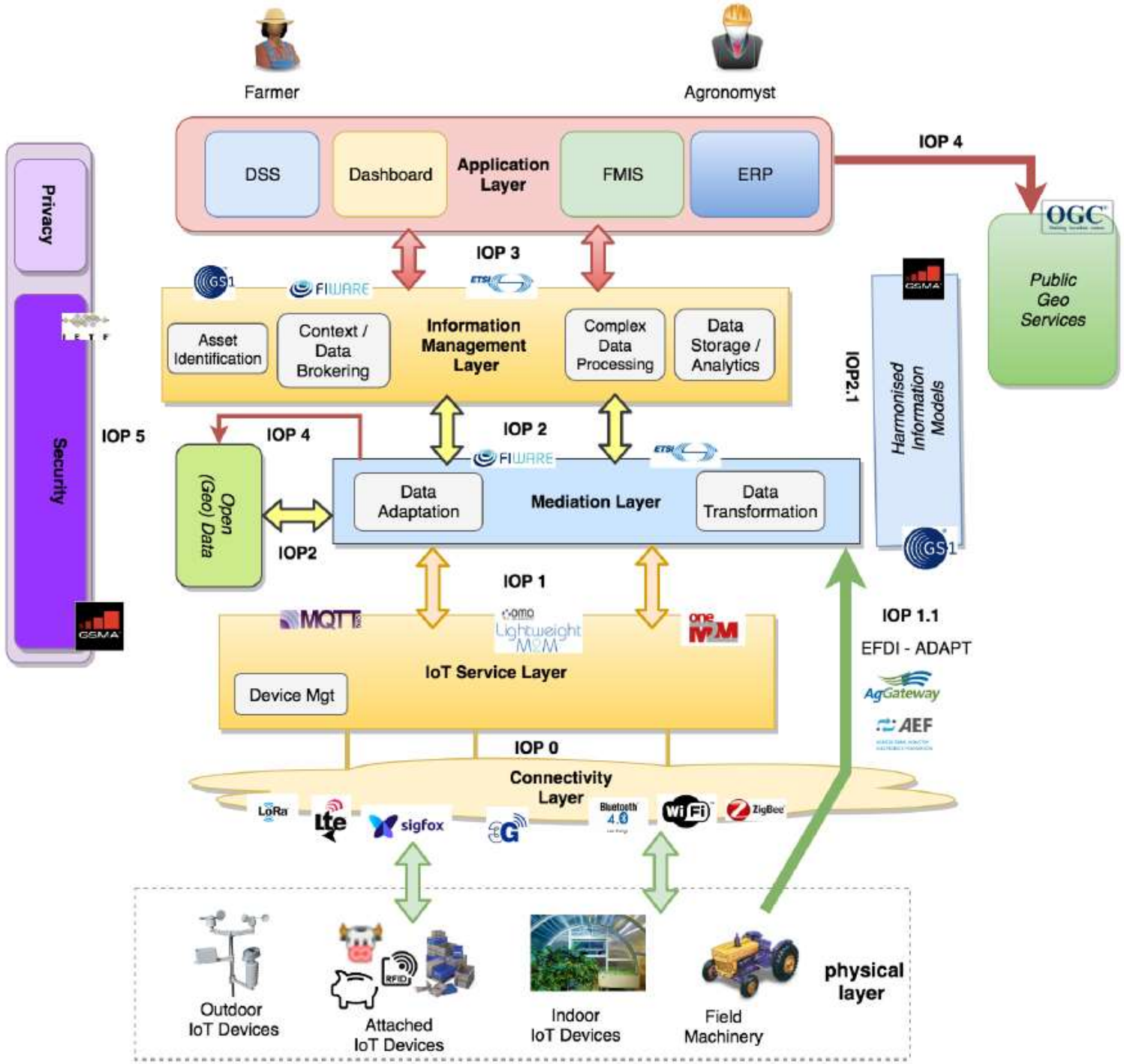
+ other sensors/sources

Milking robots

Automated data capture

SINTEF

IOF2020 Reference Architecture



SINTEF

Foundational AI standards (ISO SC42 WG1)

WG 1 - Foundational standards

- ToR: Development of foundational standards for AI
  - ISO/IEC 22989: AI Concepts and Terminology
    - settled on an initial definition for AI having looked at 5 different proposals. But still not final!
    - AI definition will be shared once this project is promoted to the CD stage after the 4th plenary meeting in October 2019
  - ISO/IEC 23053: Framework for AI Systems Using Machine Learning
  - New study item: AI Lifecycle Model

ISO 22989

ISO/IEC WD 22989:20190524 (E)

ISO/IEC JTC1 SC42/WG1

Secretariat: ANSI

Information Technology — Artificial Intelligence — Artificial Intelligence Concepts and Terminology

The current wave of progress and enthusiasm for artificial intelligence (AI) began around 2010, driven by three factors that built upon each other:

1. the availability of *big data* from many sources;
2. dramatically *improved machine learning approaches and algorithms*; and
3. the capabilities of *more powerful computers*.

Artificial intelligence is an area of information technology that has been researched for many decades, but only recently have advancements in computing capacity enabled the emergence of practical applications across many sectors. With these advancements, artificial intelligence is poised to become a new part of the foundation of information technology.

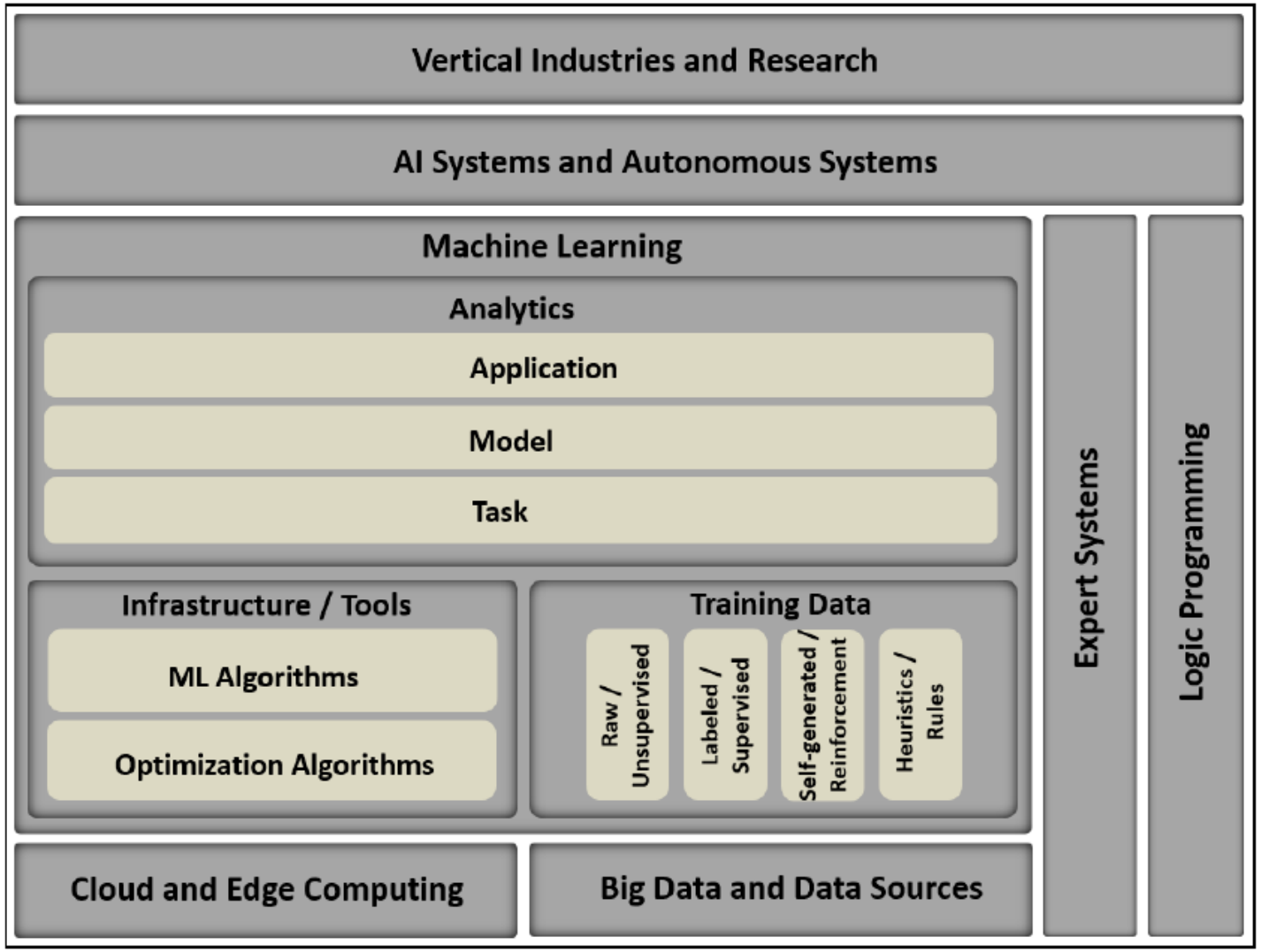
There is no single definition of AI that is universally accepted by practitioners. Some define AI loosely as a computerized system that exhibits behavior that is commonly thought of as requiring intelligence. Others define AI as a system capable of rationally solving complex problems or taking appropriate actions to achieve its goals in whatever real-world circumstances it encounters.

SINTEF

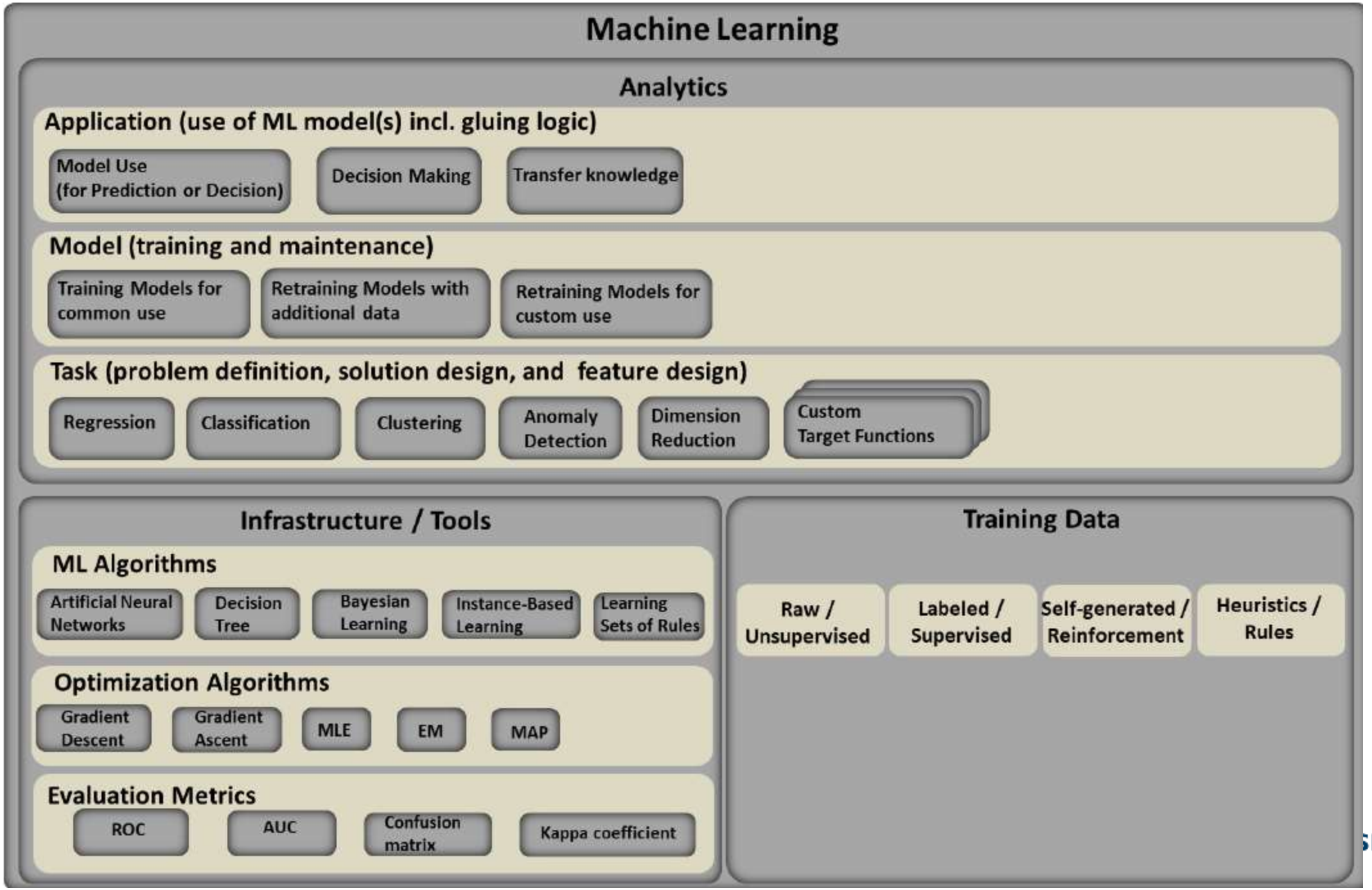
Framework for Artificial Intelligence (AI) using Machine Learning (ML) (ISO 23053)

Title of the proposed deliverable. **English title:** Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) **French title:** *(In the case of an amendment, revision or a new part of an existing document, show the reference number and current title)*
Scope of the proposed deliverable. This document defines an Artificial Intelligence (AI) and Machine Learning (ML) framework for describing a generic AI system using ML technology. The framework describes the system components and their functions in the AI ecosystem. This document is applicable to all types and sizes of organizations, including public and private companies, government entities, and not-for-profit organizations, that are implementing or using AI systems.

Framework for Artificial Intelligence (AI) using Machine Learning (ML)



SINTEF


Machine Learning
Analytics
Application (use of ML model(s) incl. gluing logic)
Model Use (for Prediction or Decision)
Decision Making
Transfer knowledge
Model (training and maintenance)
Training Models for common use
Retraining Models with additional data
Retraining Models for custom use
Task (problem definition, solution design, and feature design)
Regression
Classification
Clustering
Anomaly Detection
Dimension Reduction
Custom Target Functions
Infrastructure / Tools
ML Algorithms
Artificial Neural Networks
Decision Tree
Bayesian Learning
Instance-Based Learning
Learning Sets of Rules
Optimization Algorithms
Gradient Descent
Gradient Ascent
MLE
EM
MAP
Evaluation Metrics
ROC
AUC
Confusion matrix
Kappa coefficient
Training Data
Raw / Unsupervised
Labeled / Supervised
Self-generated / Reinforcement
Heuristics / Rules

Data for AI

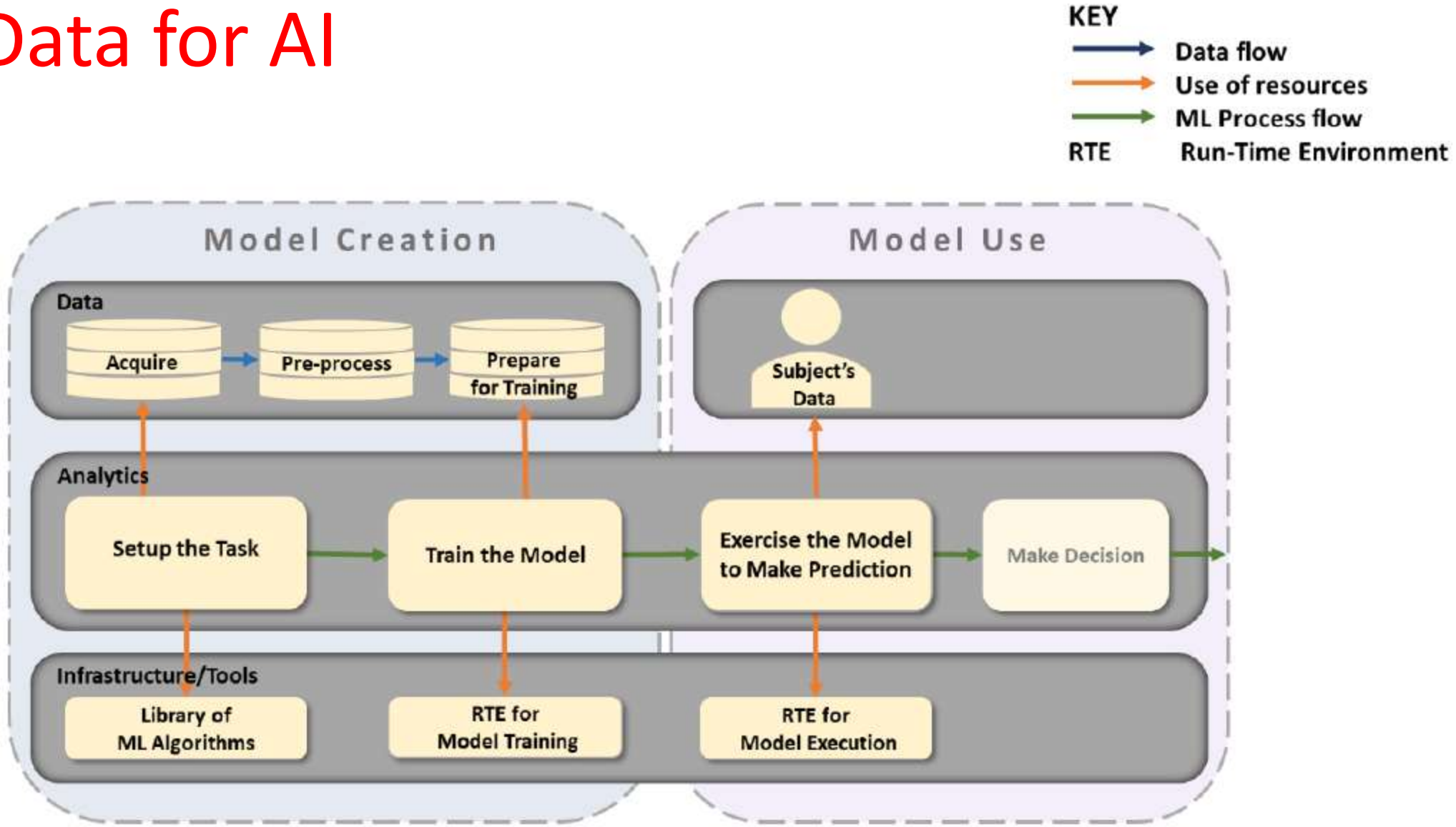


Figure 3 - Typical supervised machine learning process

SINTEF

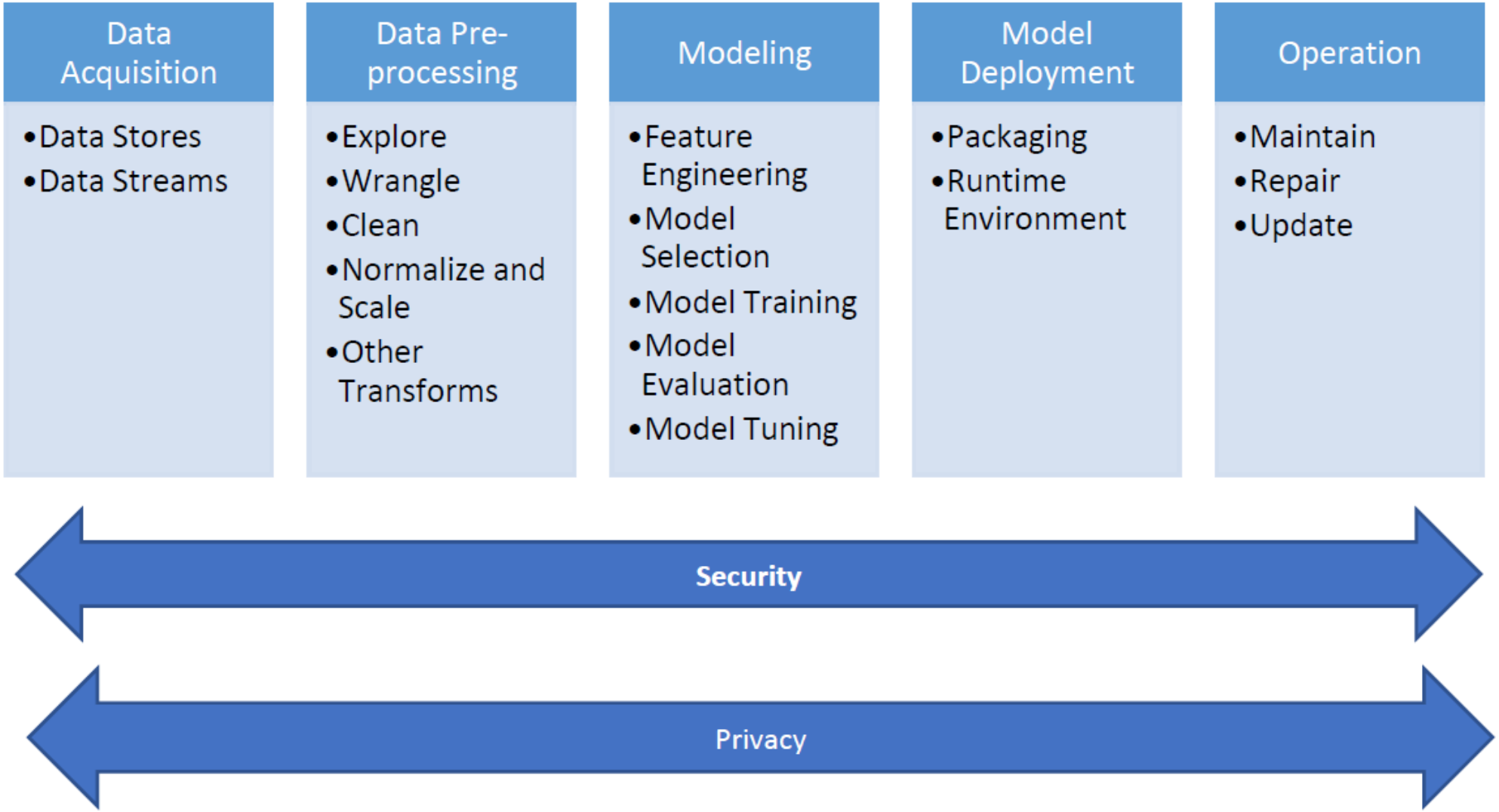


Figure 2. Machine learning pipeline

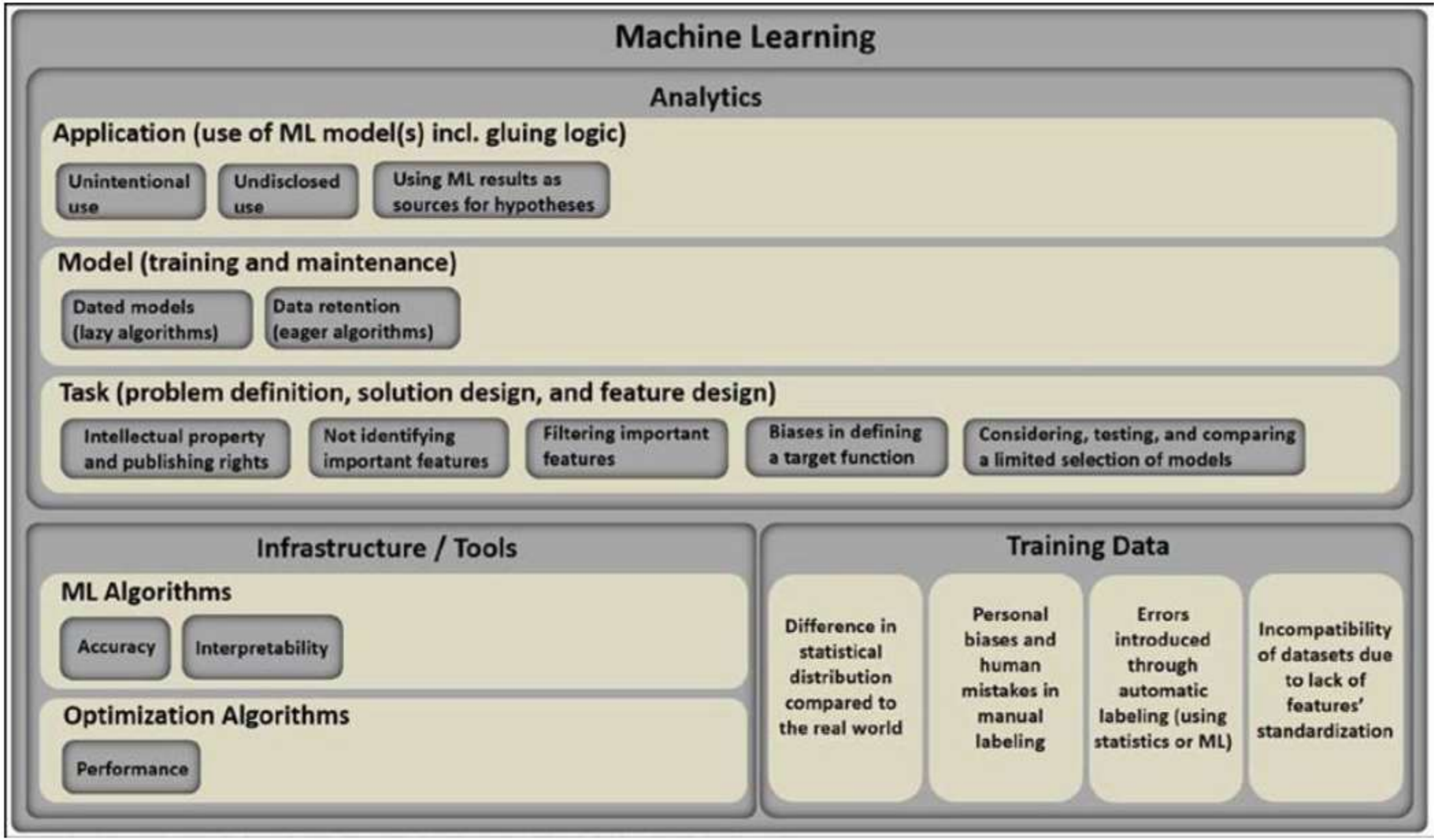
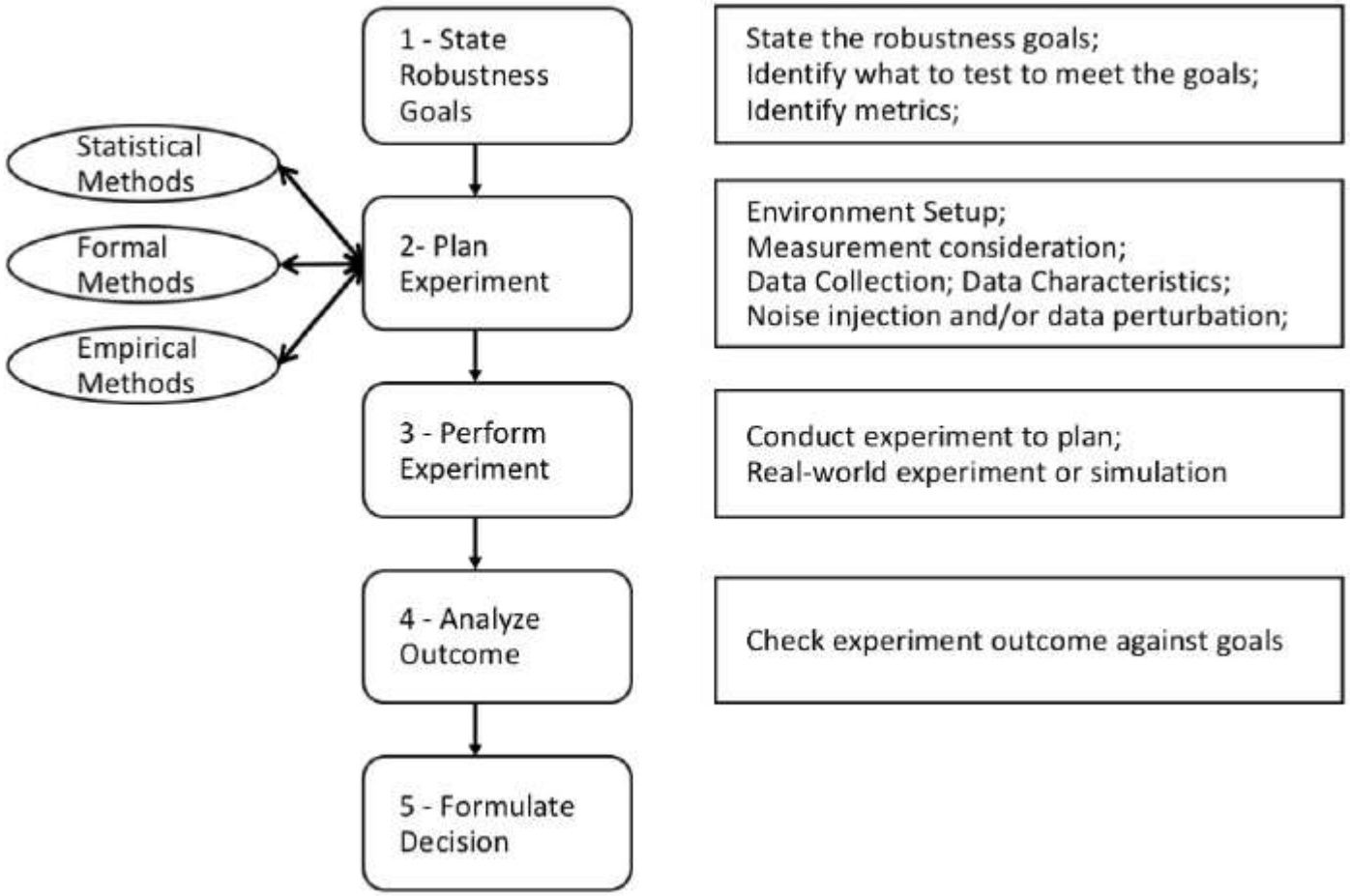


Figure 4 - Considerations in ML-based systems

Workflow to determine neural network robustness



SINTEF

Robustness of neural networks

4 Overview of the existing methods to assess the robustness of neural networks ....
4.1 General ....
4.2 Statistical methods ....
4.2.1 General ....
4.2.2 Measures of performance ....
4.2.3 Method for decision on robustness ....
4.3 Formal methods ....
4.3.1 General ....
4.3.2 Uncertainty of an interpolator ....
4.3.3 Maximum stable space for perturbation resistance ....
4.4 Empirical methods ....
4.4.1 Field trials ....
4.4.2 A posterior testing ....

A.2 Example image perturbations ....
A.2.1 Homogeneous noising ....
A.2.2 Brightening ....
A.2.3 Vibration and rotation ....
A.2.4 Atmospheric turbulance ....
A.2.5 Blurring ....
A.2.6 Blooming ....
A.2.7 Smear ....
A.3 Example sound perturbations ....
A.4 Other types of data perturbations ....

SINTEF

B.2 Training/Testing/Validation data

As described in this document, training data is used to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm. Testing and validation data are used to select the best model among candidates and to determine whether the final selected model will generalize well on input data.

For the purposes of this document, a category called "observed data" is included to help map the 23053 data type to the 19944 data taxonomy. For the purposes of this document, using terminology from ISO/IEC 19944, observed data is derived data [1] that is captured or collected. Customer data is defined in ISO/IEC 19944

23053 data type	19944 data category
Training	Customer Data, Observed Data
Testing	Customer Data, Observed Data
Validation	Customer Data, Observed Data

23053 data type	19944 data category
Input	Customer Data, Observed Data

Table B.2 Mapping of 23053 input data to 19944 data categories

23053 data type	19944 Data category Top Level	Data category 2nd Level
Output	Derived Data	Inferred Data

Table B.3 Mapping of 23053 output data to 19944 data categories

AI Neural Network Model Interoperability

https://onnx.ai/

SINTEF

ONNX

- There is one commonality between ML frameworks that enables an open format such as ONNX, and that is that they all make use of dataflow graphs in one way or another. While there are differences between each framework, they all provide APIs enabling developers to construct computational graphs and runtimes to process these graphs. Even though these graphs are conceptually similar, each framework has been a siloed stack of API, graph and runtime. The goal of ONNX is to empower developers to select the framework that works best for their project, by providing an extensible computational graph model that works as a common intermediate representation at any stage of development or deployment.

Converting to ONNX format

Framework / Tool	Installation	Tutorial
Caffe	apple/coremltools and onnx/onnxmltools	Example
Caffe2	part of caffe2 package	Example
Chainer	chainer/onnx-chainer	Example
Cognitive Toolkit (CNTK)	built-in	Example
CoreML (Apple)	onnx/onnxmltools	Example
Keras	onnx/keras-onnx	Example
LibSVM	onnx/onnxmltools	Example
LightGBM	onnx/onnxmltools	Example
MATLAB	Deep Learning Toolbox	Example
ML.NET	built-in	Example
MXNet (Apache)	part of mxnet package docs github	Example
PyTorch	part of pytorch package	Example1, Example2, export for Windows ML, Extending support
SciKit-Learn	onnx/sklearn-onnx	Example
SINGA (Apache) - Github (experimental)	built-in	Example
TensorFlow	onnx/tensorflow-onnx	Examples

Big Data (ISO SC42 WG2)

WG 2 - Big Data

- ToR: Standardization in the area of Big Data
- 20547-4 Big data security and privacy (SC 27/WG 4/5) [CD stage]
- NWIP on Process Management Framework - Big Data Analytics [New]

Initital Big Data Standards - transferred from JTC WG9 - 2018

Reference number of document: **ISO/IEC DIS 20546**

Committee identification: ISO/JTC 1/WG 9

Secretariat: ANSI

Information technology — Big data — Overview and vocabulary

Document type: International standard

Document subtype:

Document stage: (30) Committee Draft

Document language: E

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

SINTEF

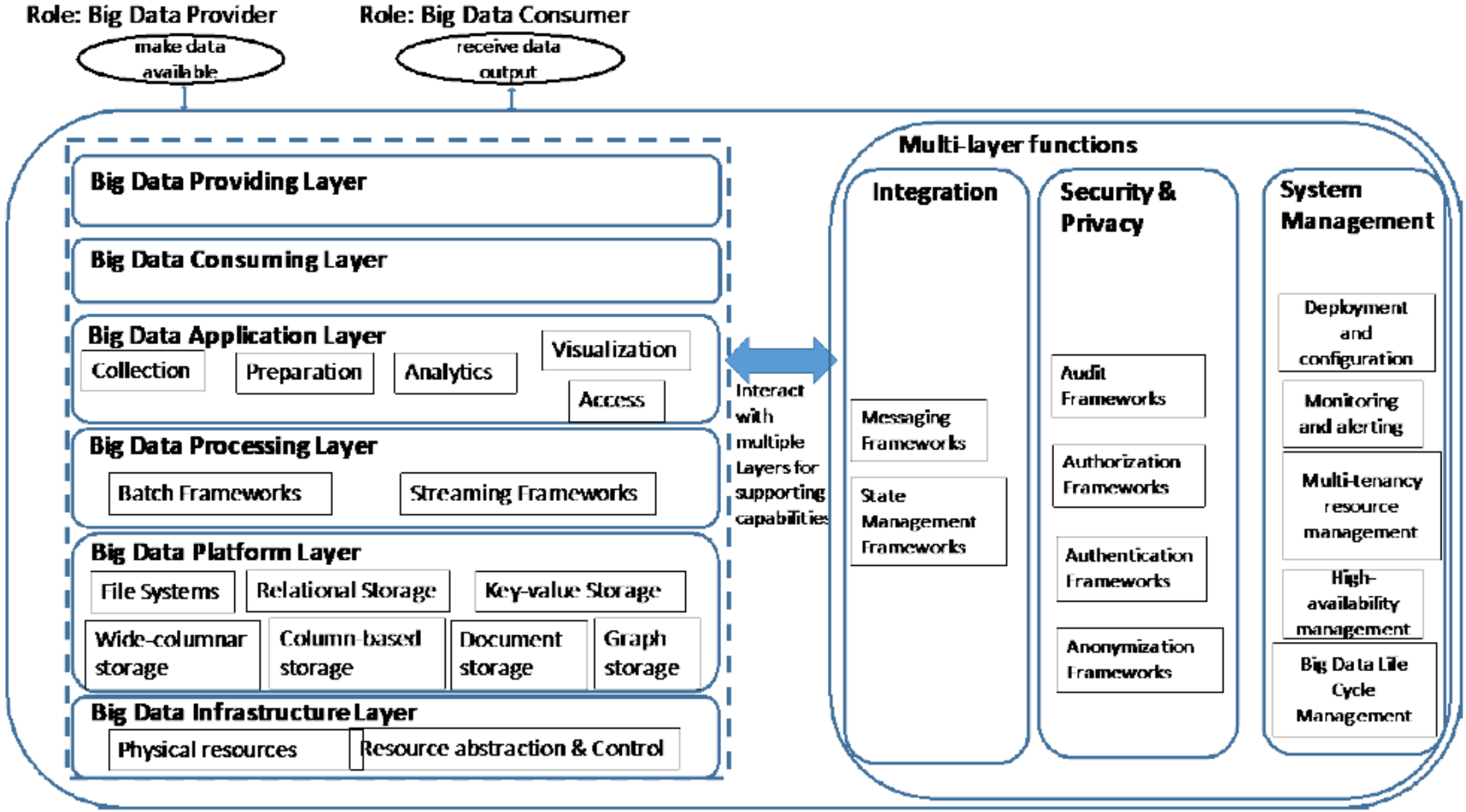
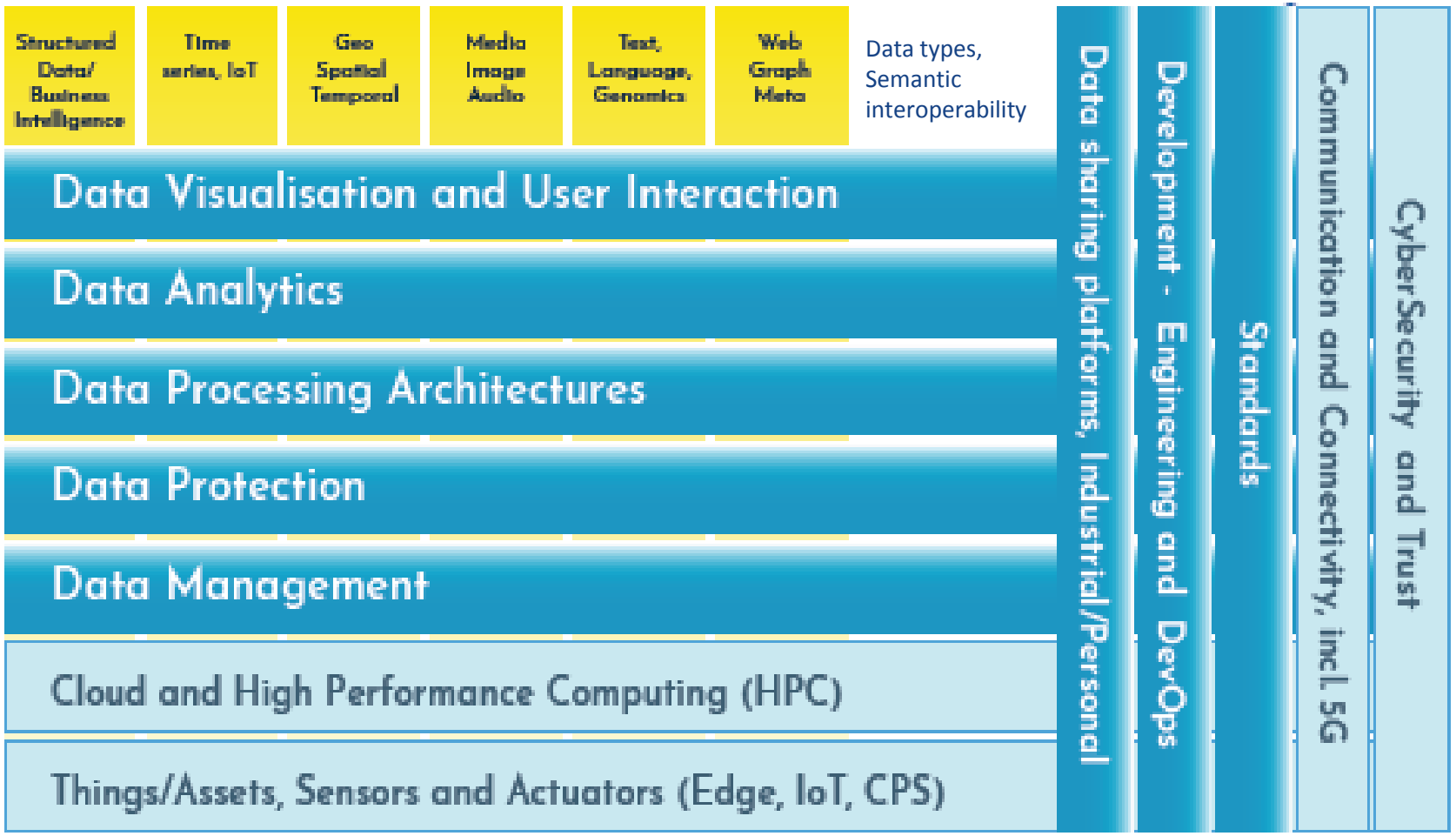


Figure 12: Functional Components of BDRA

BDV – Big Data and Analytics/Machine Learning Reference Model



SINTEF

New Query Language for Graph Databases to Become International Standard

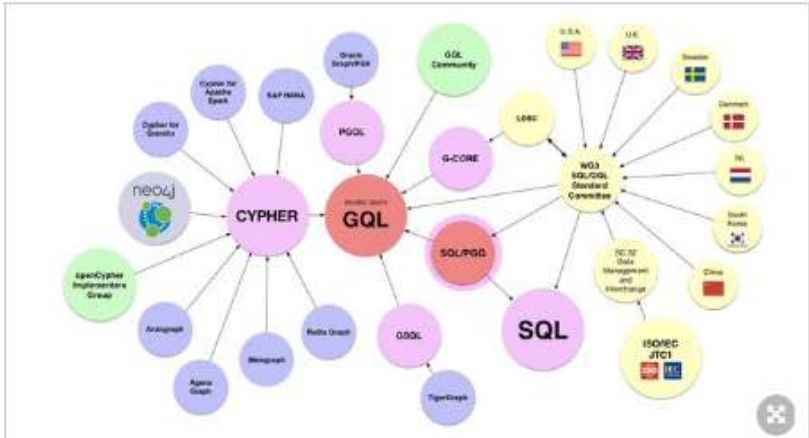
Neo4j Backs Launch of GQL Project: First New ISO Database Language Since SQL

NEWS PROVIDED BY
Neo4j →
Sep 17, 2019, 05:00 ET

SHARE THIS ARTICLE

SAN MATEO, Calif., Sept. 17, 2019 /PRNewswire/ -- Neo4j, the leader in graph databases, announced today that the international committees that develop the SQL standard have voted to initiate GQL (Graph Query Language) as a new database query language. Now to be codified as the international standard declarative query language for property graphs, GQL represents the culmination of years of effort by Neo4j and the broader database community.

The initiative for GQL was first advanced in the GQL Manifesto in May 2018. A year later, the project was considered at an international gathering in June. Ten countries including the United States, Germany, UK, Korea, and China have now voted in favor, with seven countries promising active participation by national experts.



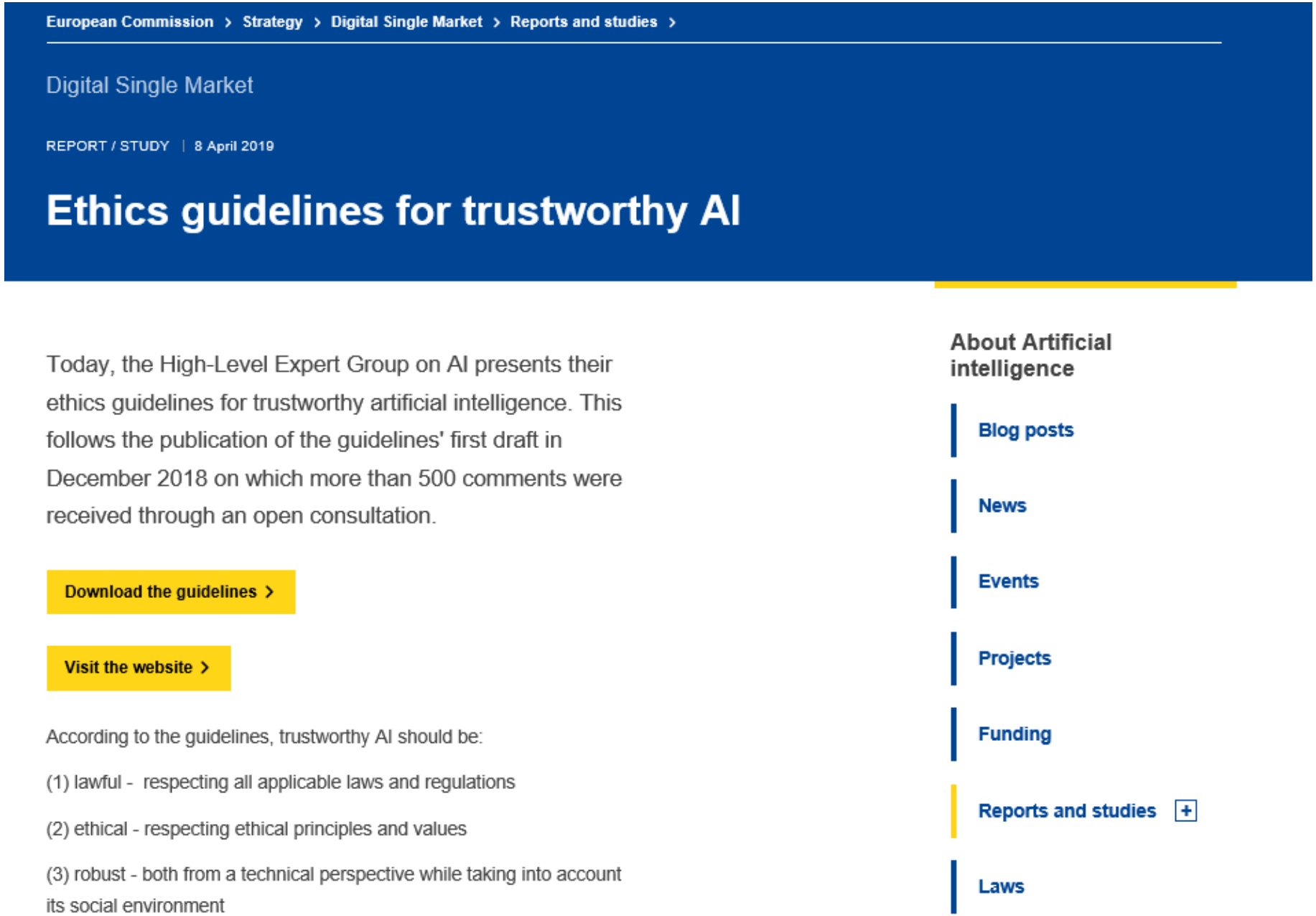
GQL to incorporate and consider several graph database languages.

It has been well over 30 years since ISO/IEC began the SQL project. SQL went on to become the dominant language for accessing relational data, achieving wide adoption by vendors and practitioners and dramatically accelerating the growth of the relational database market. The GQL project will initiate development of the next generation of technology standards for accessing data, optimized for today's world of connected data. Its charter mandates building on core foundations already established by SQL, as well as ongoing collaboration to ensure SQL and GQL compatibility and interoperability.

SINTEF

Trustworthiness (ISO SC42 WG3)

- **WG 3 - Trustworthiness**
  - ToR: Standardization in the area of AI Trustworthiness
  - ISO/IEC TR 24027: Bias in AI systems and AI aided decision making
    ISO/IEC TR 24028: Overview of Trustworthiness in AI [WD stage]
  - ISO/IEC TR 24029: Assessment of the Robustness of Neural Networks - Part 1: Overview [WD stage]
    - Possible future WI will be on Formal Methods to assess NNs Robustness
  - ISO/IEC 23894: AI Risk Management [New]
  - ISO/IEC NP TR: Overview of Ethical and Societal Concerns [New]
    - provides information in relation to principles, processes and methods in this area
    - intended for technologists, regulators, interest groups, and the society at large
    - not intended to advocate for any specific set of values (value systems)
  - JTC 1 SC 27 Study Period on AI Impact on Privacy



https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Computational approaches of AI (ISO SC42 WG5)

WG 5 - Computational approaches and computational characteristics of AI systems [New]

- ToR: Standardization in the area of computational approaches and computational characteristics of AI systems
- TR "Overview of computational approaches for AI systems" [New]

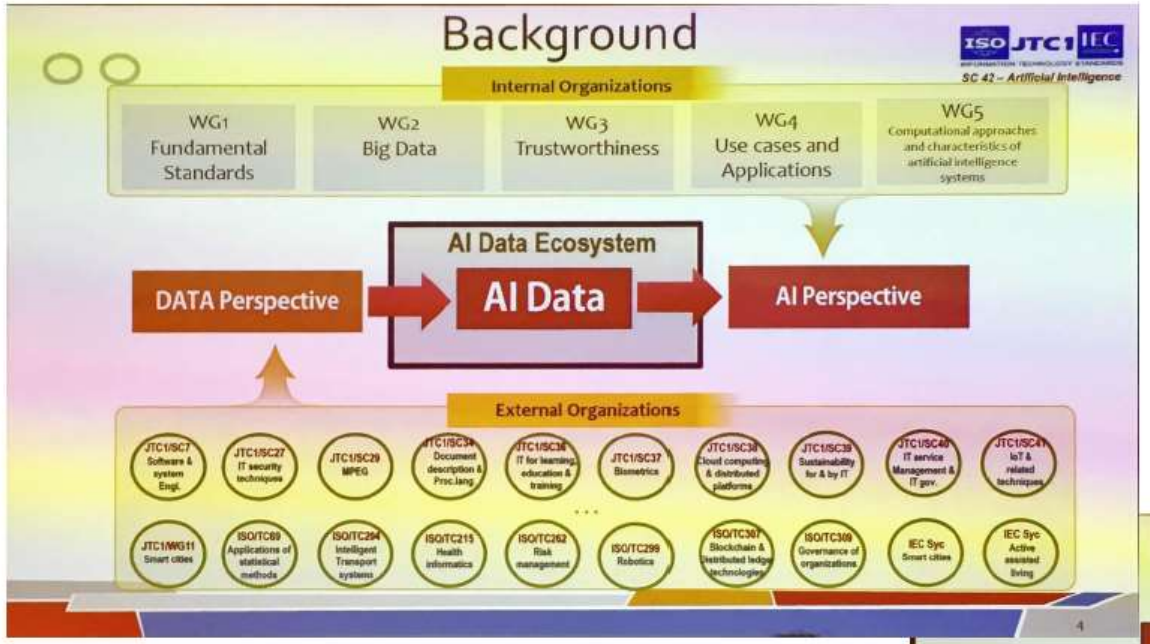
SC 42/SC 40 JWG 1 AI Governance

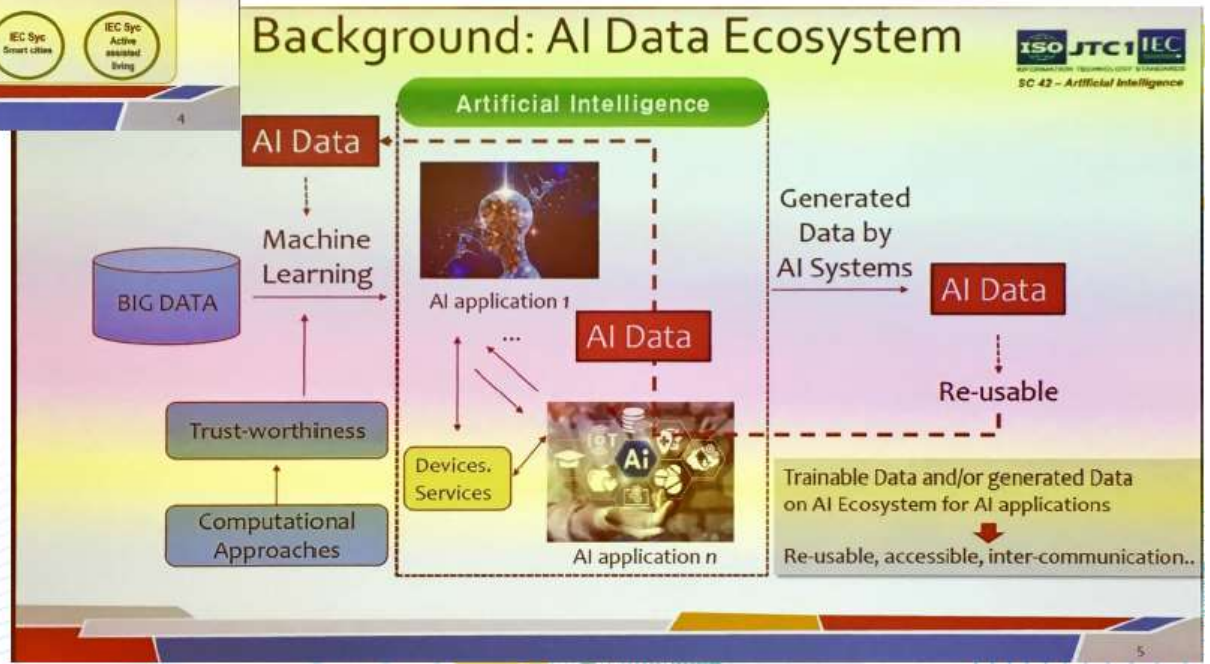
- ISO/IEC 38507: Governance implications of the use of artificial intelligence by organizations [New]

Ad-hoc group on Intelligent Systems Engineering

Future Proposal from the Korean NB on AI Data Ecosystem

ISO SC42 New work item proposal - AI Data Ecosystem





BDV BIG DATA VALUE ASSOCIATION

SINTEF

CEN CENELEC Focus Group on AI (European)

Scope:

- Explore the need for European standardization on AI within CEN-CENELEC. Taking into account the guidelines of the High-Level Expert Group on AI and the Commission Communication (237/2018)

Elected convenors

- Patrick Bezons (Chair of French standardization committee for AI, AFNOR/France)
- Sebastian Hallensleben (VDE/DKE Germany and chair IEC SEG-10 on "Ethics in Autonomous and AI Applications")

Perspective from the convenors

- Vision: Focus on a realistic and persuasive European approach for harnessing the huge potential of AI
- Priorities:
  1. Ethics/trustworthiness in AI
  2. Reliable jurisdiction and data protection
  3. Interoperability of AI algorithms, training data sets, etc
- For the AI ethics part, the objective will be to operationalize it and to make it tangible and concrete for technology
- Leverage the European research & industry ecosystem and have a close link with both ISO/IEC JTC 1 SC42 and IEC SEG-10