ITU Regional Forum on “Internet of Things, Telecommunication Networks and Big Data as basic infrastructure for Digital Economy”
Saint-Petersburg, Russian Federation, 4-6 June 2018

Data processing and management: opportunities, challenges and trends, including in ITU-T standardization activities related to IoT and SC&C

Marco Carugi
ITU-T Q2/20 Rapporteur and SG20 Mentor
Liaison rapporteur to ISO/IEC JTC1 SC41
marco.carugi@gmail.com

1956 2016
CCITT / ITU-T

Outline

- Data Processing and Management (DPM)
  - Opportunities, challenges and trends
- Internet of Things and Data - relevant studies with focus on ITU-T activities
  - Some standardization advances on Big Data related aspects
  - ITU-T SG20 (IoT and SC&C)
  - ITU-T Focus Group on “Data Processing and Management to support IoT and Smart Cities & Communities”
  - Enhancing DPM capabilities with AI/Machine Learning

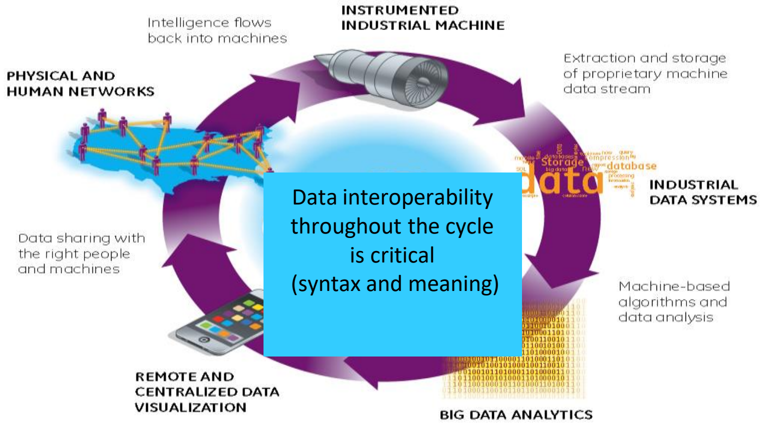
Data Processing and Management (DPM) – opportunities, challenges and trends

The Data deluge (in IoT) – from connecting devices to value creation

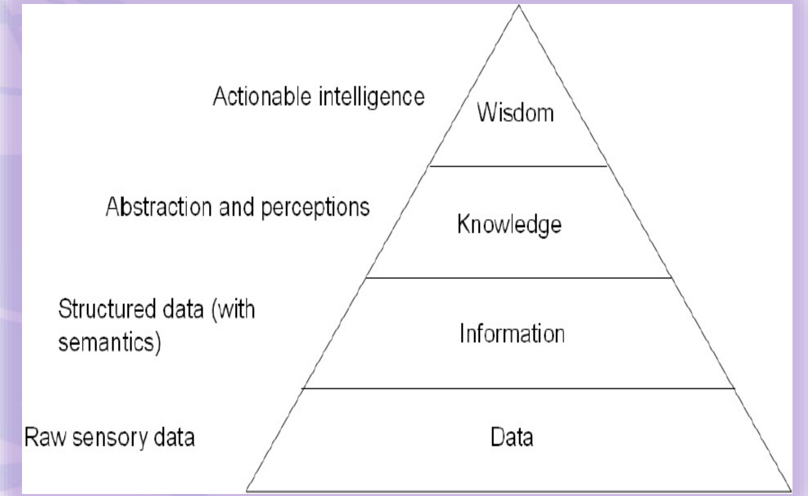
- Huge volume of data produced by IoT devices
- Multiple data sources (things, context, historical data, social data)
- Data with different velocity, formats, precision, confidence, quality
- Different operations on data for extraction of actionable intelligence

A key goal is to have ready for use the right data, at the right time, at the right location

The Industrial Internet Data loop *[source: GE whitepaper]*



From Data to Actionable Intelligence



Knowledge hierarchy applied in data processing
Source: Barnaghi and al., IJSWIS, 2012

Opportunities and challenges of Data (for networks and services)

Process optimization and data monetization via analytics - driving revenue by sharing, analysing and interpreting data, for multiple purposes

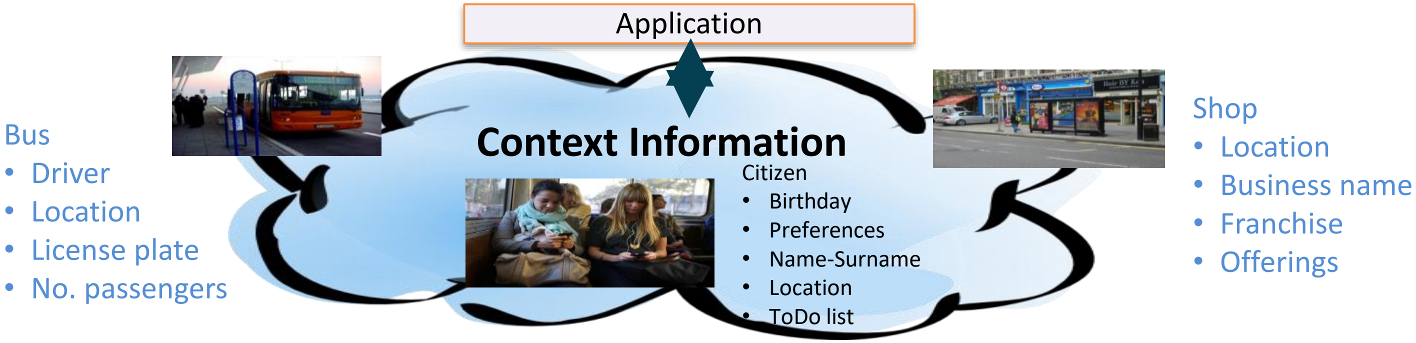
- Extraction of tangible business and technology value
- Response and action in real time, improving productivity/business processes, lowering costs
- Long-range forecasts enabling strategic actions - business differentiation, addressing society challenges
- New/improved business models and service offer, faster, more efficiently and agile

Technical and non-technical challenges

- Dealing with the “V”s of data : Volume, Variety, Velocity, Variability, Veracity
- Discovery of appropriate devices and data sources
- Integration of heterogeneous devices, networks and data
- Scalability for large device numbers, diverse and huge data, computational complexity of data interpretation
- Availability and (open) access to data, data query
- Trust, security and privacy of data
- Interpretation (extraction of actionable intelligence from data)
- Massive data mining, efficient processing, flexible learning
- **Other non-technical challenges also essential (incl. data ownership and data governance)**

Cross-domain Data Sharing – example of context information management

Cross-domain uses cases require access to information from different domains that is normally held in separate silos - e.g. sharing of context information



Different sources of context information

- Existing systems, users (through mobile apps), sensors/IoT devices

Standard specifications are needed for context information management

- To ensure vendor neutrality for users such as Cities
- To reduce technological barriers to development/deployment, to enable innovative services

Blockchain technology for Data Exchange and Sharing

Some benefits of Blockchain in IoT

- Efficient ensurance of integrity, authenticity, auditability and traceability of transactions (data) => trust based information transmission
- Decentralized approach (lower maintenance costs)
- Multi-party consensus (data security)
- Distributed approach (multi-party collaboration)
- Enabler of data monetization

Some limitations of Blockchain in IoT

- Not suitable for massive IoT data service transactions and frequent data transmissions
- High performance and capacity requirements not fitting constrained IoT environments
- Big storage capacity needs cannot cope with IoT devices' storage as the blockchain grows
- Data security and privacy is relative

Source: ITU-T FG-DPM D3.6 (ongoing)

*Blockchain technology fosters a new generation of transactional applications that **establish trust, accountability, transparency and efficiency.** It shows great promises across a wide range of business applications in many fields, including IoT and smart cities.*

DPM solutions' scalability requirement for stream processing, including video

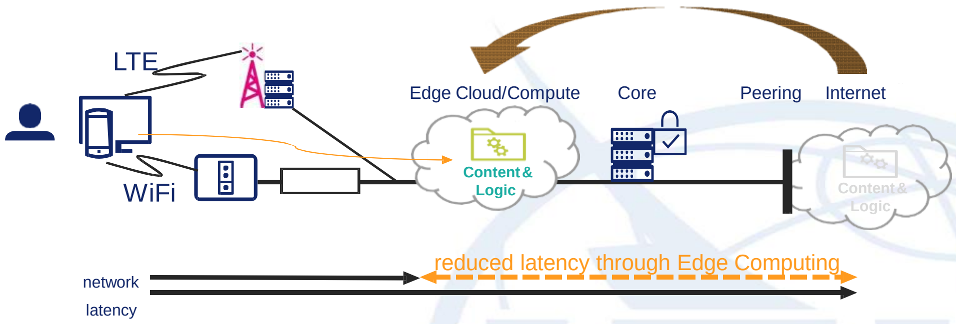
Pure centralized cloud solutions will not be able to scale for the continuous and timely processing of growing amounts of real-time streams

Solutions mixing edge and central cloud processing with high performance computing capabilities will be required

- Faster response to emergencies
- Improved decision-making
- Increased operational efficiency

1956 2016
CCITT / ITU-T

Edge Computing: computing and storage resources next to the user



Edge Computing benefits

- **(Ultra-)low latency**: disruptive improvement of customer experience
- **Reduction of backhaul/core network traffic**: cloud services near to user
- **In-network data processing**

Some issues not fully addressed

- Resource limitation, more complexity, service continuity/mobility, …

Low latency applications

Autonomous Devices

- Drones
- Self-Driving Cars
- Robotics

Immersive Experiences

- Interactive Environments
- Virtual Reality
- Augmented Reality

Natural Interfaces

- Voice Control
- Motion Control
- Eye-Tracking

[Ultra-low Latency < 20 ms]

ITU-T SG20 ongoing draft Recommendations

- "IoT requirements for edge computing" (Y.IoT-EC-reqts)
- "Capabilities and framework of edge computing-enabled gateway in the IoT" (new Y.IoT-EC-GW)

Edge Computing … and more: Fog/Device Computing

The hot topic of «Personal Data» processing and management

European Union's "General Data Protection Regulation (GDPR)" addresses usage and protection of personal data
GDPR has been adopted by EU Parliament in 2016, and is enforceable throughout EU since 25 May 2018

- Processed lawfully, fairly and in a transparent manner ("Lawfullness, fairness and transparency")
- Collected for specified, explicit and legitimate purposes ("Purpose limitation")
- Adequate, relevant and limited to what is necessary ("Data minimization")
- Accurate and, where necessary, kept up to date ("Accuracy")
- Identification no longer as necessary for the purposes ("Storage limitation")
- Processed in an appropriate manner to maintain security ("Integrity and confidentiality ")
- Accountability (documentation)

GDPR principles for personal data

Source: Austrian Data Protection Authority

*Among the current technical issues in the IoT space with respect to privacy protection and compliance to GDPR, **how to implement the GDPR principles in IoT technical architectures (mechanisms, standards, testing) requires specific attention and work in the standardization arena***

Integrating DPM with Machine Learning (ML) technologies

ML has potential for network design, operation and optimization

- coping with highly increased complexity
- enhancing efficiency and robustness of network operations (e.g. by reducing number of measurements and facilitating robust decisions)
- increasing network self-organization feasibility (cognitive network management)
- providing reliable predictions [pro-active strategies] (e.g. adaptive QoS in highly dynamic environments)

Also, ML has potential to enable new advanced services and applications

Various challenges, including trust *[specific slide further]*

Internet of Things and Data – relevant studies with focus on ITU-T activities

IoT and leading technologies

The IoT will benefit from the integration of a number of leading technologies, including those for

- Machine to Machine Communications
- Advanced sensing and actuation
- Big Data management
- Cloud Computing (and distributed computing)
- Semantics support
- Machine Learning (AI)
- Security, Privacy and Trust
- Softwarization (incl. Software Defined Networking, Network Functions Virtualization)
- Autonomic Networking
- Service Delivery Platforms
- Specific applications/industries

The key role of Standardization for IoT interoperability

Market research: “nearly 40% of IoT economic impact requires interoperability between systems”

IoT value will come solving interoperability issues within/across IoT domains (different dimensions)

Key issue with IoT interoperability is current diversity =>> key role of international SDOs in standards convergence/harmonization

Open innovation systems move fast =>> Standardization needs to cope - process, collaboration

Some standardization advances on Big Data related aspects

Big Data technical standardization

Some areas for standardization [SG13's Y.Supp 40 on Big Data standards roadmap and more]

- Big Data architecture, Data model and APIs (APIs with network infrastructure, users, auditors)
- Flexible Analytics (real-time, batch; remote, distributed and federated analytics; network-driven analytics)
- Data access, including Open Data frameworks and Data Governance within companies
- Framework for Data quality and trust (context dependent)
- Framework and standards for Data Exchange - data sharing, transaction, interconnection
- Security and Data protection, anonymization and de-identification of Personal Data (and reversibility)
- Integration of Big Data requirements in central/distributed cloud computing solutions for both infrastructure and services (interoperability, data/process security, traceability, personal data protection, SLAs, data storage)
- Standards and guidelines to address issues for legal implications of (Big) Data in telecom (e.g. Data ownership)
- Benchmarks for system performance evaluation
- Standardized visualization methods
- Domain-specific languages

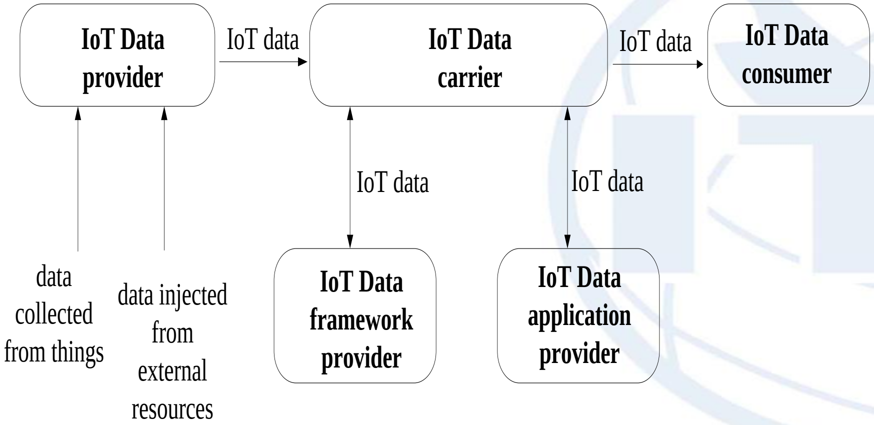
Different standards initiatives are addressing the different (technical) Big Data areas

- ***ITU-T*** *[SG13 (non-IoT), SG20 and FG-DPM (IoT)]*, ***ISO/IEC JTC1 WG9, BDVA and others***

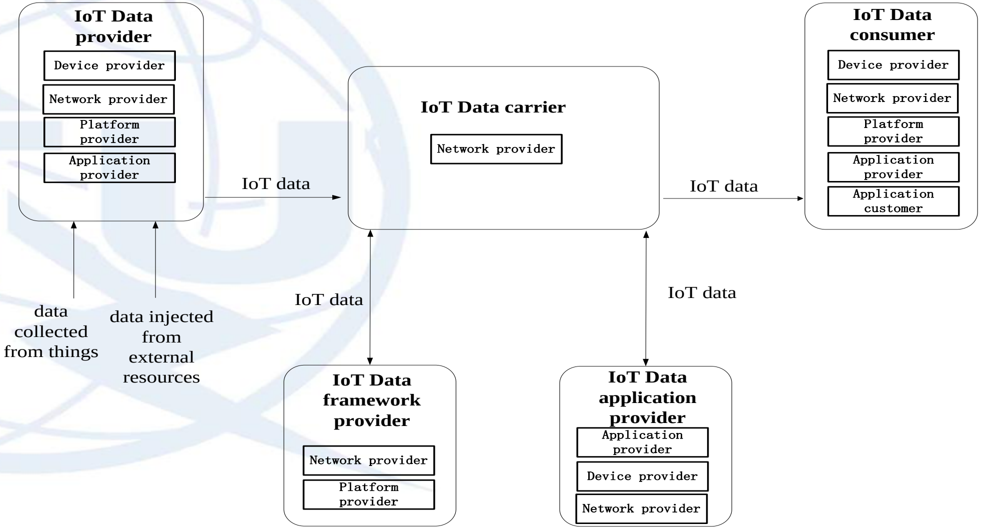
A foundational ITU-T Recommendation on Big Data in IoT:

ITU-T Y.4114 "Specific requirements and capabilities of the IoT for Big Data"

Specific requirements and capabilities the IoT is expected to support to address the challenges related to Big Data



The IoT data roles identified in Y.4114
[the key roles relevant in an IoT deployment from a data operation perspective]

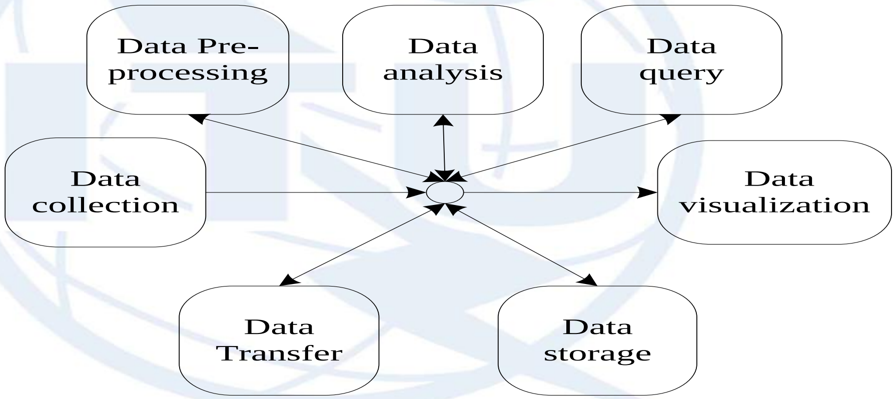


Mappings of IoT business roles (ITU-T Y.2060) to IoT data roles

Y.4114 – IoT data operations

Abstract representation of IoT data operations and related data flows
(diverse set of concrete IoT deployments does not imply unique logical sequencing of IoT data operations)

IoT data operations



The IoT data operations sequencing highly depends on service and deployment scenarios
(e.g. data analysis with respect to cloud computing vs edge computing scenarios)

Semantics based technologies to cope with essential requirements of data in IoT and networks

Semantics is the study of meaning

- A definition of semantics: “The rules and conventions governing the interpretation and assignment of meaning to constructions in a language” [ITU-T Z.341]
- Shared vocabularies (for network entities, services) and their relationships [ontologies]

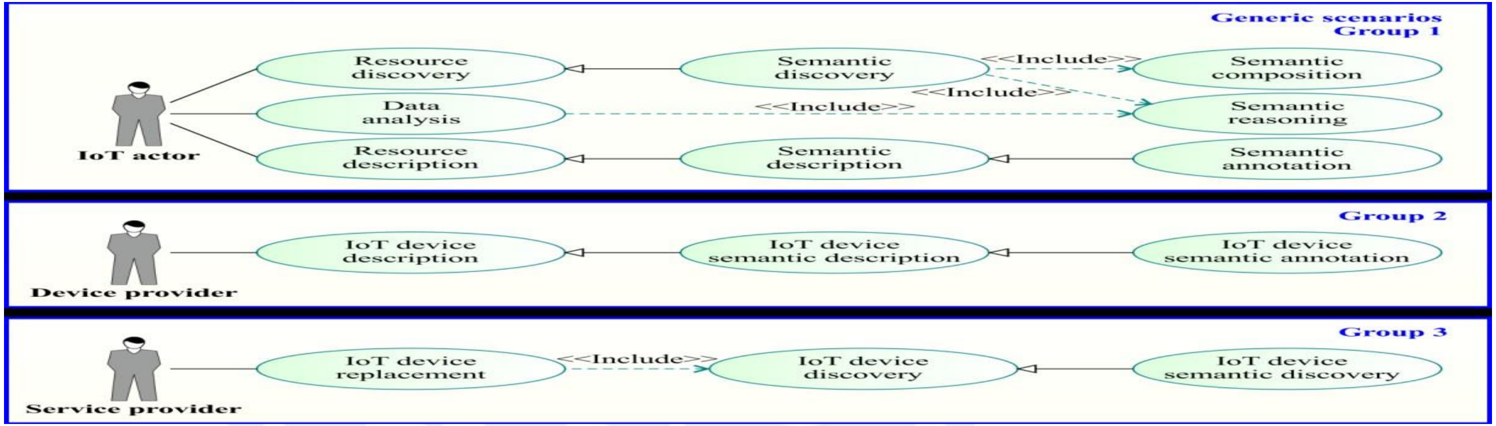
Requirements for data (and services) interoperability, scalability, consistency, discovery, reusability and composability, analytics and reasoning for actionable intelligence, are essential in IoT and networks in general

- Examples of driving factors from IoT: growing number of interconnected things, variety of devices, types of generated data, number and type of services

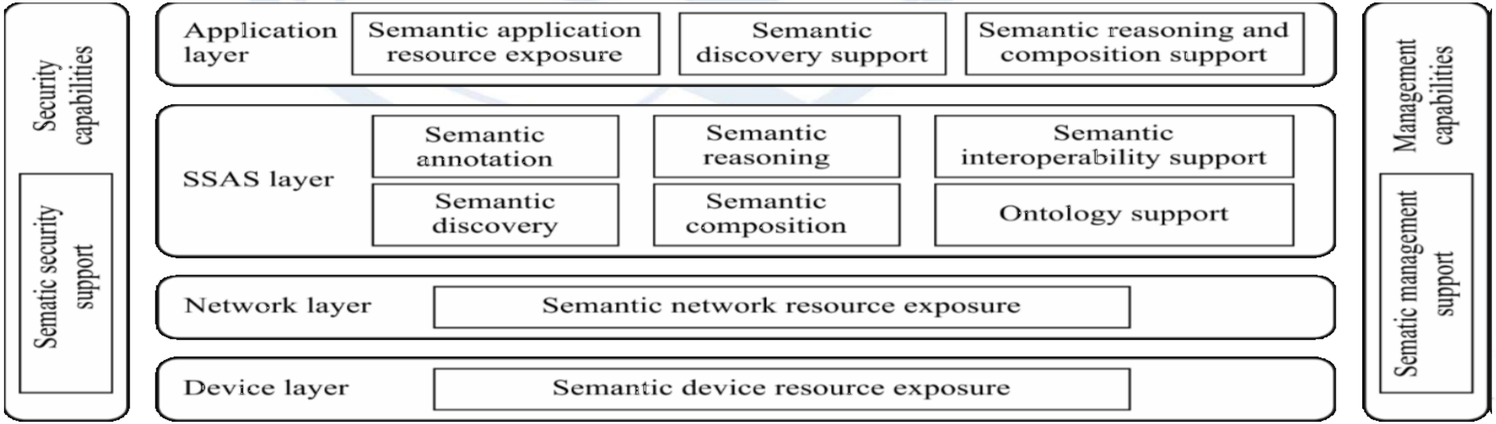
Semantics based approaches reveal outstanding features to support these requirements in IoT and networks in general

ITU-T Y.4111 “Semantics based requirements and framework of the IoT”

Examples of semantics based scenarios [Y.4111]



Global view of IoT semantics based capability framework according to the IoT Ref. Model [Y.4000/Y.2060]



The full applicability of semantics based technologies in IoT and networks is not yet there

The benefits of semantics based technologies will be realized in incremental way

Among the current issues

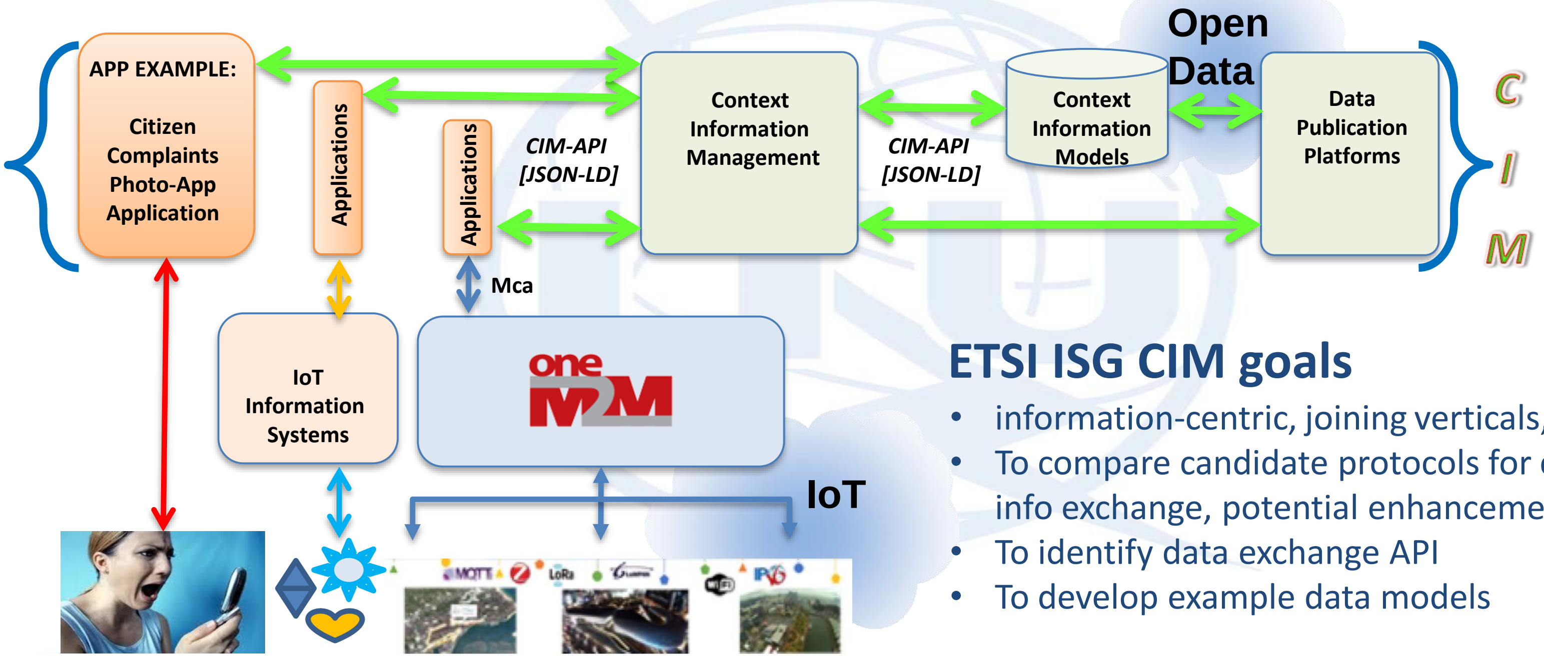
- Lack of elaborated use cases as drivers to validate the value proposition (value for the whole value chain)
- Insufficient link with network architectures
- Immaturity of tools, essential to establish semantics based bridging
- Semantic discovery of services, things and their capabilities
- Semantic metadata framework and base network ontology (cross-domain horizontal integration)
- Participation of domain specific communities (incentives for sharing) and entrepreneurs/developers interaction

Growing promising experimentations (e.g. with EC support in Europe), but further development, validation and standardization are needed

Specific standardization work is progressed in various groups, e.g. W3C, ETSI SmartM2M/oneM2M, OGC

Data sharing: ETSI ISG CIM working on a cross-domain Context Information Layer

An info-exchange layer on top of IoT platforms - especially targeting Smart City applications

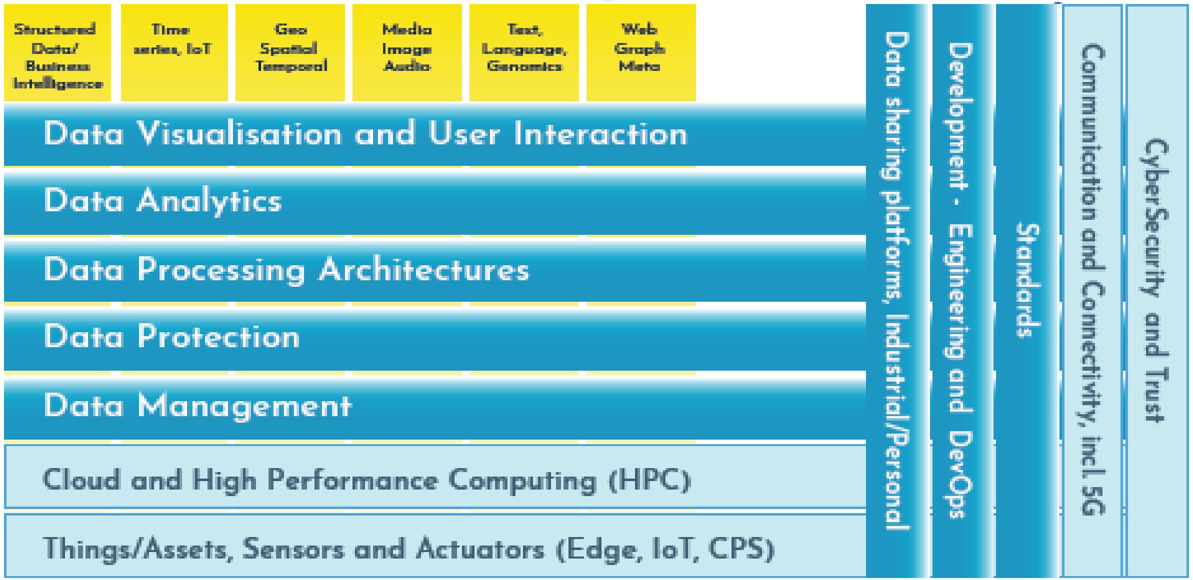


ETSI ISG CIM goals

- information-centric, joining verticals, interoperable
- To compare candidate protocols for cross-domain info exchange, potential enhancements
- To identify data exchange API
- To develop example data models

Collaboration with SDOs (ITU-T FG-DPM, oneM2M, W3C ...) & open-source implementations

Big Data Value Association – BDV Reference Model



BDV RM is structured into technical areas (capabilities) – there is no layering connotation

A key step in front of the IoT standardization work plan: Big Data-IoT architectural integration

ITU-T SG20 (IoT and SC&C)

Smart Cities as super application domain of the IoT

Integration of multiple verticals

Citizen-centric services

Still a number of technical challenges, incl. interoperability, scalability, dynamicity, security and privacy

The brain of the city

Smart City Platform

Data collection, analysis, knowledge, planning, action

The senses of the city

Social networks

Mobile applications

WorldWideWeb

Legacy Devices

IoT Devices

City data sources

Source: Dr. Levent Gürgen

Smart Cities: an incremental and participatory journey towards full support to Data Economy

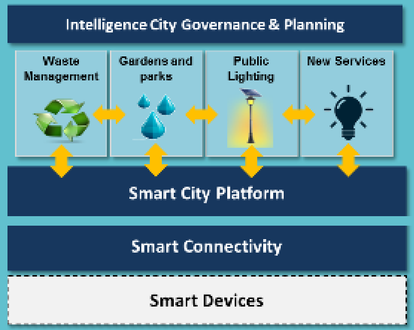
1 → 2 → 3 → 4

Efficient and Open

- Vertical solutions bringing efficiency in silos
- Historic data as open data
- Information still in vertical silos, no global picture

Truly Smart

- Horizontal platform integrating “right-time” context info from different vertical services
- Predictive and prescriptive models



Unleashing Right-time Open Data

- Right-time context info published to third parties
- Exchange of context info with systems from other domains

Support to Data Economy

- City as a platform including also 3rd party data enabling innovative business models
- Open and commercial data enabling multi-side markets

Source: FIWARE

ITU-T Focus Group on “Data Processing and Management to support IoT and Smart Cities & Communities” (FG-DPM)

ITU-T Focus Group on Data Processing and Management to support IoT and Smart Cities & Communities (ITU-T FG-DPM)

Essential tasks

- Identify challenges in IoT and smart cities for DPM
- Identify key requirements ad capabilities for DPM
- Promote the establishment of trust-based data management frameworks for IoT and SC&C
- Investigate the role of emerging technologies to support data management incl. blockchain
- Identify and address standards gaps and challenges

WG1 - Use Cases, Requirements and Applications/ Services

WG2 - DPM Framework, Architectures and Core Components

WG3 - Data sharing, Interoperability and Blockchain

WG4 - Security, Privacy and Trust including Governance

WG5 - Data Economy, commercialization and monetization

5 Working Groups

Main priority

To propose mechanisms, frameworks and guidelines for supporting security, privacy and interoperability of datasets and data management systems in IoT and smart cities

1st meeting in July 2017 (SG20 is parent SG)

1st ITU Workshop on Data Processing and Management for IoT and Smart Cities & Communities: Brussels, Belgium, 19 Feb. 2018

CCITT / ITU-T

Ongoing studies in ITU-T FG-DPM

FG-DPM Working Groups
WG1 Use Cases, Requirements and Applications/Services
WG2 DPM Framework, Architectures and Core Components
WG3 Data sharing, Interoperability and Blockchain
WG4 Security, Privacy and Trust including Governance
WG5 Data Economy, commercialization and monetization

Partial list of deliverables

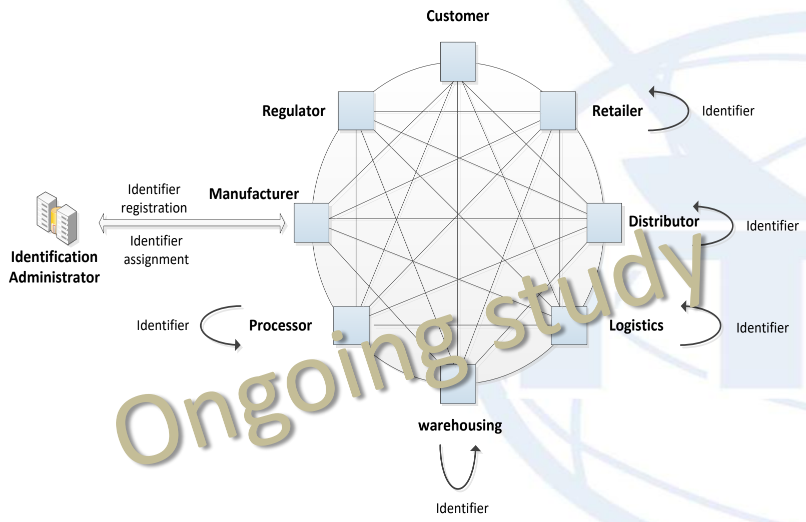
- **D1.1 Use Cases Analysis and General Requirements for DPM**
- D2.1 DPM Framework for Data-driven IoT and SC&C
- D2.2 DPM Functional Architectures
- D2.3 Data Modeling and Formats Specification for DPM
- D3.1 Framework of Open/Private Data
- D3.2 Technical Enablers for Open Data Platform
- **D3.6 Blockchain-based Data Exchange and Sharing Technology**
- D4.1 Framework of Security and Privacy in DPM
- D4.5 Data Governance Framework for IoT and SC&C
- **D5.1 Modeling of Data Economy for value creation and Stakeholders identification**
- D5.2 Business models, commercialization and monetization to support data economy
- D5.3 Data economy impact assessment, policy and sustainability implications
- Cross-WG studies: **DPM taxonomies and vocabularies**, gap analysis, DPM standardization roadmap

"Use Cases Analysis and General Requirements for DPM" (D1.1): an entry point for the whole FG-DPM work

- **Objectives**
  - Identify from DPM perspective - per each use case - ecosystem's actors and business roles, data characteristics, capabilities, requirements and other
  - **Facilitate comparison among different use cases** (**across single or multiple domains**) to enable common DPM features to be identified/adopted, and facilitate single/cross-domain applications' implementation
  - **Allow creation of new services** at little extra cost
  - **Feed other FG-DPM deliverables** (DPM framework, area-specific frameworks, others)
- **Progress so far**
  - ***"Unified DPM Use Case template"*** developed and disseminated to numerous potential DPM use cases contributors (incl. SDOs, Alliances, EU H2020 projects)
  - Numerous DPM use cases collected
  - Comparison of DPM use cases started - identifying common/use case-specific reqts
  - DPM capabilities' global picture discussion initiated (WG1, WG2, others)

Blockchain for IoT in FG-DPM and in SG20

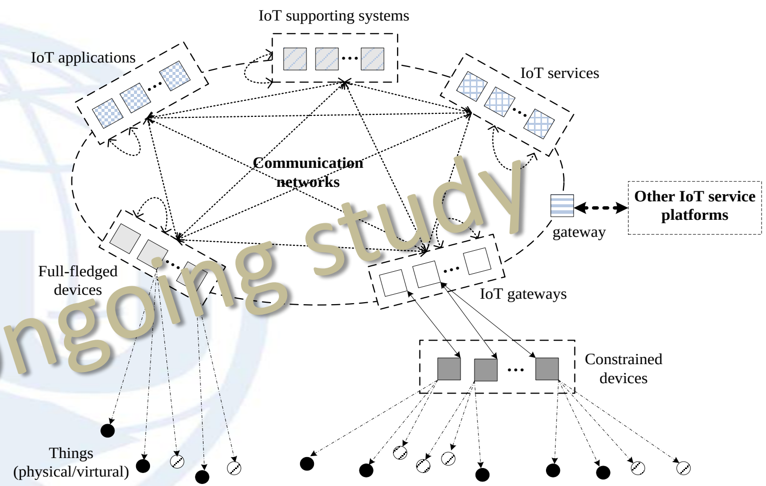
Ongoing FG-DPM D.3.6 deliverable



Blockchain-based supply chain traceability use case

Exchanging and sharing static and dynamic information among the supply chain actors

Ongoing draft Rec. Y.IoT-BoT-fw in SG20

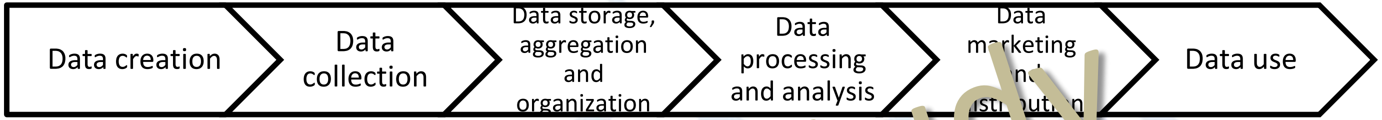


Blockchain of things as decentralized service platform

Centralized IoT service platforms may become key bottlenecks with increase of connected things and other demands (e.g., trusts and transactions),

Data Economy, commercialization and monetization (D5.1)

Data Core Activities



Data Support Activities

Data Value Chain

Use case/business model dependent: one actor can play multiple roles, more actors can participate in one single activity

Data laws, regulations and policies: formulation and enforcement of data related laws, regulations and policies in a jurisdiction (in some cases policies might be applicable to one or more organizations rather than an entire jurisdiction).

Data security and privacy services: provisioning of data related security and privacy services for implementing and enforcing data laws, regulations and policies in a jurisdiction or data policies in an organization.

ICT connectivity and infrastructure services: provisioning of ICT connectivity and infrastructure services for implementing data value chain activities (services of telecommunications operators, cloud based service providers, data centre service providers, etc.)

Enhancing DPM capabilities with AI/Machine Learning

Enhancing DPM with Machine Learning (ML) technologies

ML has potential for network design, operation and optimization

ML has potential to enable new advanced services and applications

But a number of challenges need to be addressed *[beyond trust (*)]*

- how to deal with stringent requirements of many applications (latency)
- how to ensure robust ML given small data sets and under latency constraints
- how to deal with distribution of data at different locations and diverse data formats
- usage of distributed learning to have efficient usage of scarce resources
- how to deal with (wireless) channel noise, dynamicity and unreliability
- how to ensure good tracking capabilities (explainability)
- how to exploit context information and expert knowledge (hybrid data/model-driven ML approaches)

Source: discussions in initial meetings of ITU-T FG-ML5G

(*) Trust challenges call for decision-making process transparency, controllability, verifiability , other

ITU-T FG on "Machine Learning for Future Networks including 5G" (FG-ML5G) – created in Nov 2017 , SG13 as Parent ITU-T Study Group

Initial AI studies within ITU-T SG20 [early stages of development]

- Technical Report on "Artificial Intelligence and Internet of Things" (TR.AI4IoT)
  - Aiming to explore aspects of AI within the IoT. It plans to collect and synthesize information, incl. use cases and existing standardization activities
- Draft Rec. Y.SSC-AISE-arc "Reference architecture of artificial intelligence service exposure for smart sustainable cities"
  - incl. use cases and analysis on deployment modes for AI services in cities

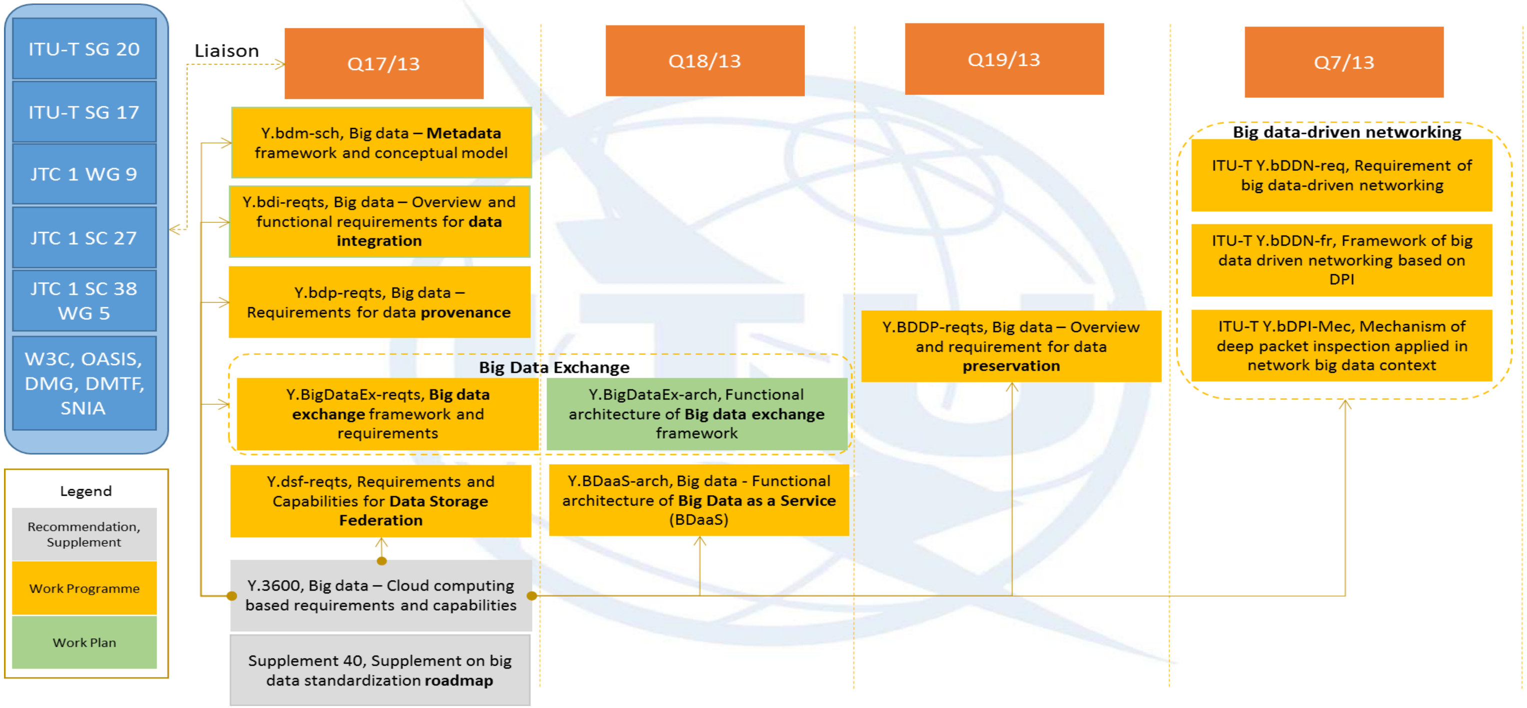
International coordination on Big Data and AI/ML standardization (for networks as well as services) is expected between overall ITU-T efforts and relevant SDOs, Alliances, Consortia

- including, but not limited to, ISO/IEC JTC1 (SC42), IEEE (Global Initiative on Ethics of Autonomous and Intelligent Systems), BDVA (Alliance for IoT Innovation)

Thank you very much for your attention

Backup information

Big Data activities in ITU-T SG13 - status Q3 2017



Overall development roadmap of Big Data in ITU-T SG13

Description of the IoT data roles identified in Y.4114

- **IoT Data provider:** it collects data from things, injects data processed within the IoT system as well as data from external sources, and provides them via the IoT Data carrier to the IoT Data consumer (optionally, the applications provided by the IoT Data application provider may execute relevant data operations with the support of the IoT Data framework provider).
- **IoT Data application provider:** **it provides applications related to the execution of IoT data operations (e.g. applications for data analysis, data pre-processing, data visualization and data query).**

  NOTE - The applications provided by the IoT Data application provider can interact with the infrastructure (e.g. storage cloud) provided by the IoT Data framework provider through the IoT Data carrier or run on the infrastructure (e.g. scalable distributed computing platforms) provided by the IoT Data framework provider.
- **IoT Data framework provider:** **it provides general IoT data processing capabilities and related infrastructure (e.g. storage and computing resources, data processing run time environment)** as required by IoT Data provider, IoT Data carrier, IoT Data application provider and IoT Data consumer for the support of data operations execution.
- **IoT Data consumer:** it consumes IoT data. Usage of the consumed data depends on the application purposes.
- **IoT Data carrier:** it carries data among IoT Data provider, IoT Data framework provider, IoT Data application provider and IoT Data consumer.

  NOTE - An actor of a concrete IoT deployment can play multiple roles. As an example, an actor executing data analysis plays the role of IoT Data application provider, but it also plays the role of IoT Data provider when it sends the results of this data analysis to other actors.

Y.4114 describes also key possible mappings from IoT business roles [Y.2060] to IoT data roles, as well as an example of deployment scenarios with respect to IoT data roles (Appendix).