DATA STORYTELLING AND BIG DATA VALUE CHAIN IN NATURAL DISASTER MANAGEMENT

e-Course on big data analytics for natural disaster management

AIDA AICET2025 Conference
Thessaloniki
14 July 2025

ANTONIO FILOGRANA

ENGINEERING INGEGNERIA INFORMATICA S.P.A.

ANTONIO.FILOGRANA@ENG.IT

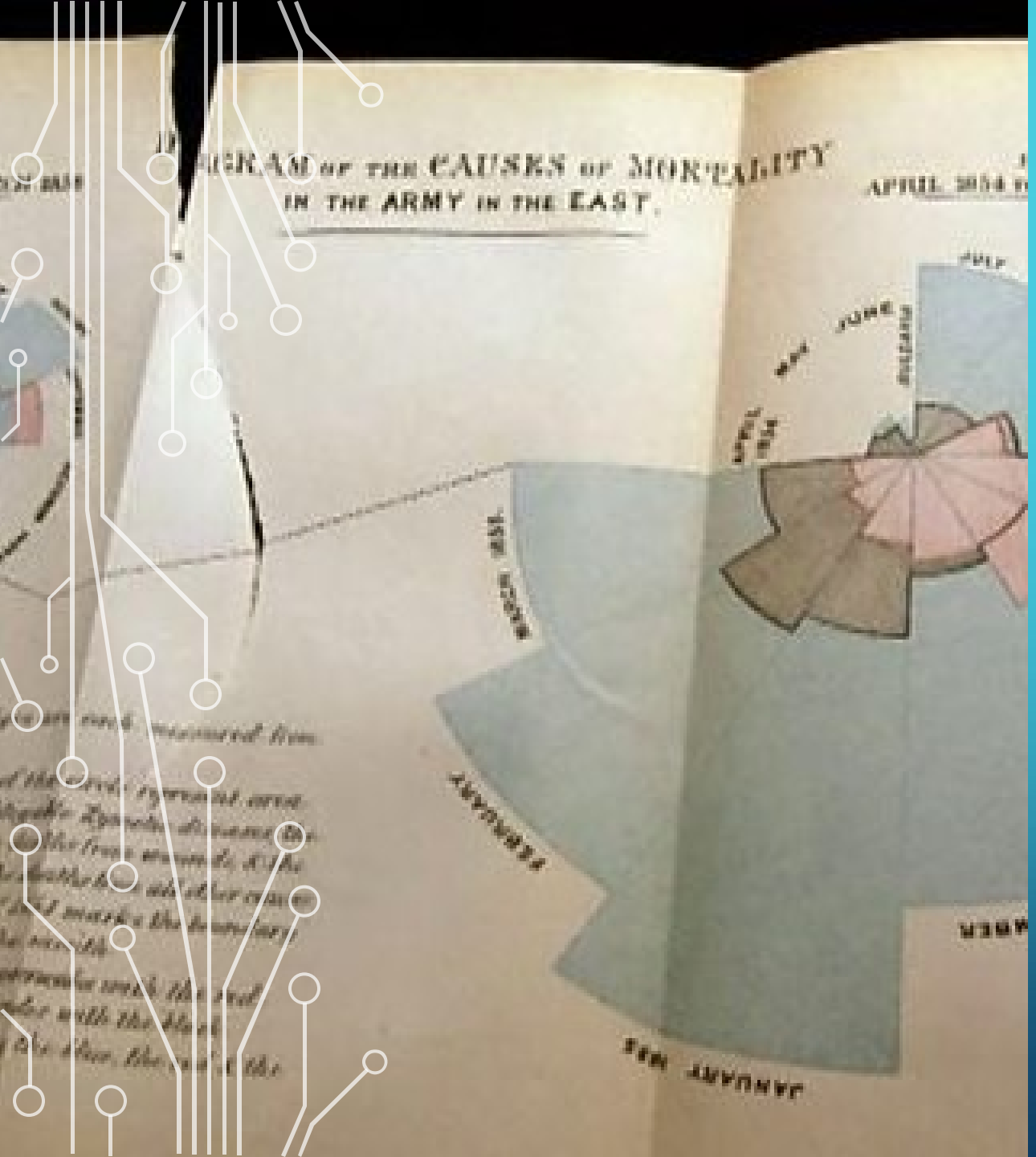
ONCE UPON A TIME…

Who is this lady?

Florence Nightingale

CRIMEAN WAR

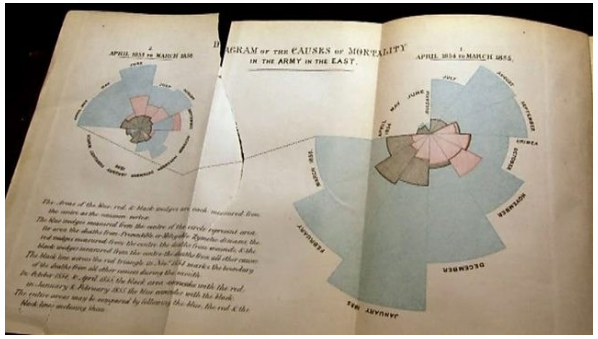
A **nurse** during the Crimean war in 1856



THE FIRST DATA STORYTELLER

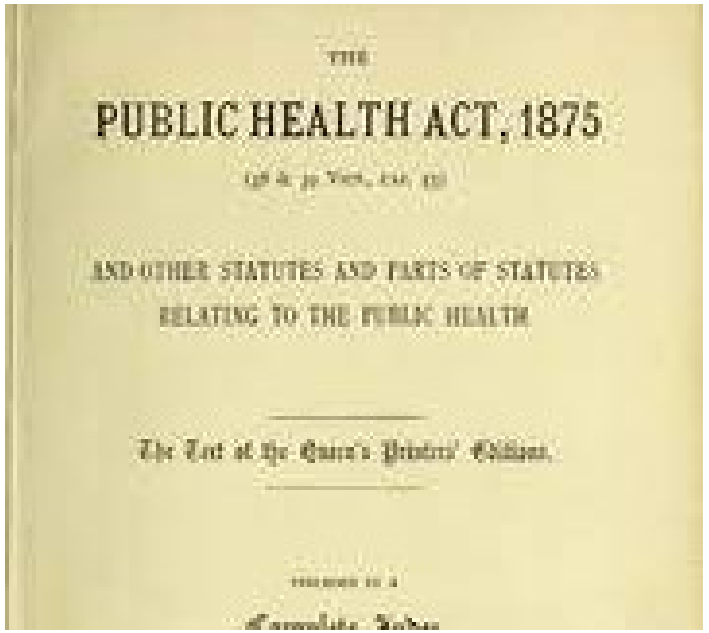
Main causes of mortality during the war: **healthcare conditions** in which the soldiers lived.

FROM DATA TO A POSSIBLE SOLUTION



solution

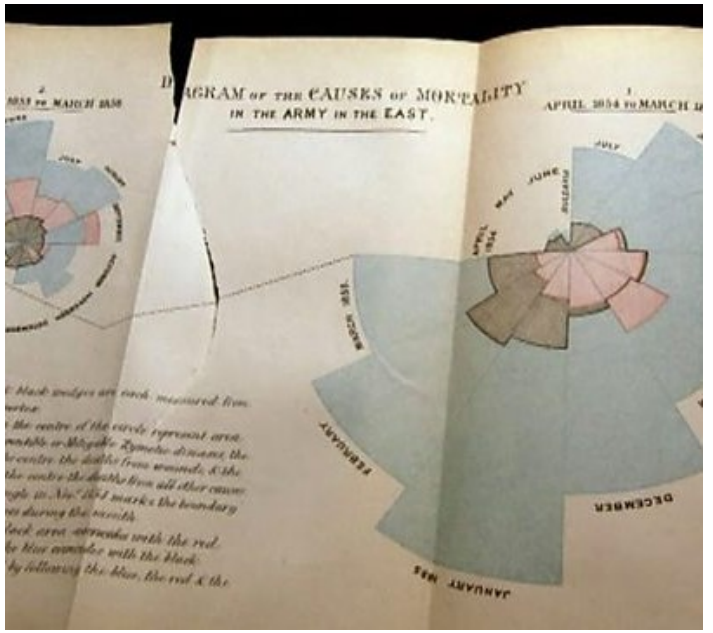
1. the sewerage network reclamation
2. clean air in the rooms
3. reduction of the overcrowding


THE

PUBLIC HEALTH ACT, 1875

AND OTHER STATUTES AND PARTS OF STATUTES RELATING TO THE PUBLIC HEALTH

The Act at the Queen's Printers' Offices.



BRITISH PUBLIC ACT


BIG DATA IS A NEW #KEYWORD ?
WHO MANAGED BIG DATA FOR THE FIRST TIME?

PARTICLE PHYSICS

1899

Cloud chamber

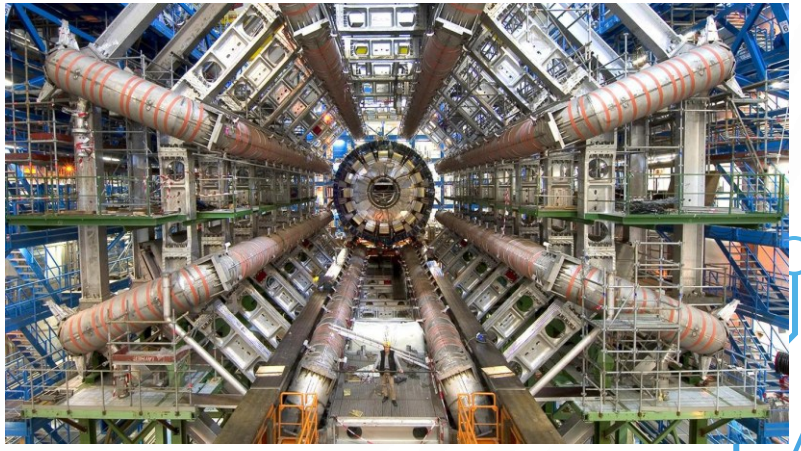
(Charles Wilson)

1962

SLAC

(Standford University)

2012



LHC

(CERN)

PARTICLE PHYSICS

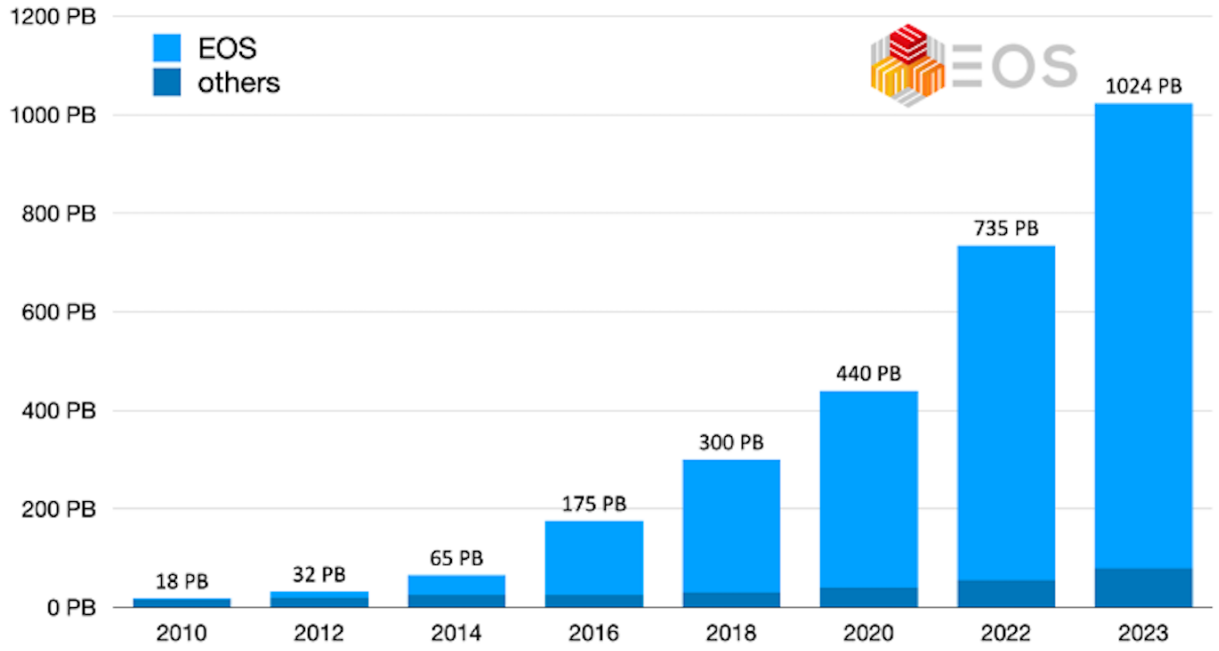
SLAC-BaBar Data Analysis System

50/400 simultaneous/total physicists, 300 Tbytes per year

HARDWARE	UNITS	End FY1999	End FY2000
Tape Silos (STK Powderhorn, 6000 tapes each)	silos	6	6
Tape Drives (STK Eagle, 20 Gbyte, 10 Mbytes/s)	drives	20	40
Disk (net capacity of RAID arrays)	Tbytes	20	56
File Servers and Data Movers (Sun)	CPUs	73	150
Interactive Servers (Sun + Linux)	CPUs	82	140
Batch Servers (Sun + Linux)	CPUs	300	900
Network Switches (Cisco 6509)	switches	5	14

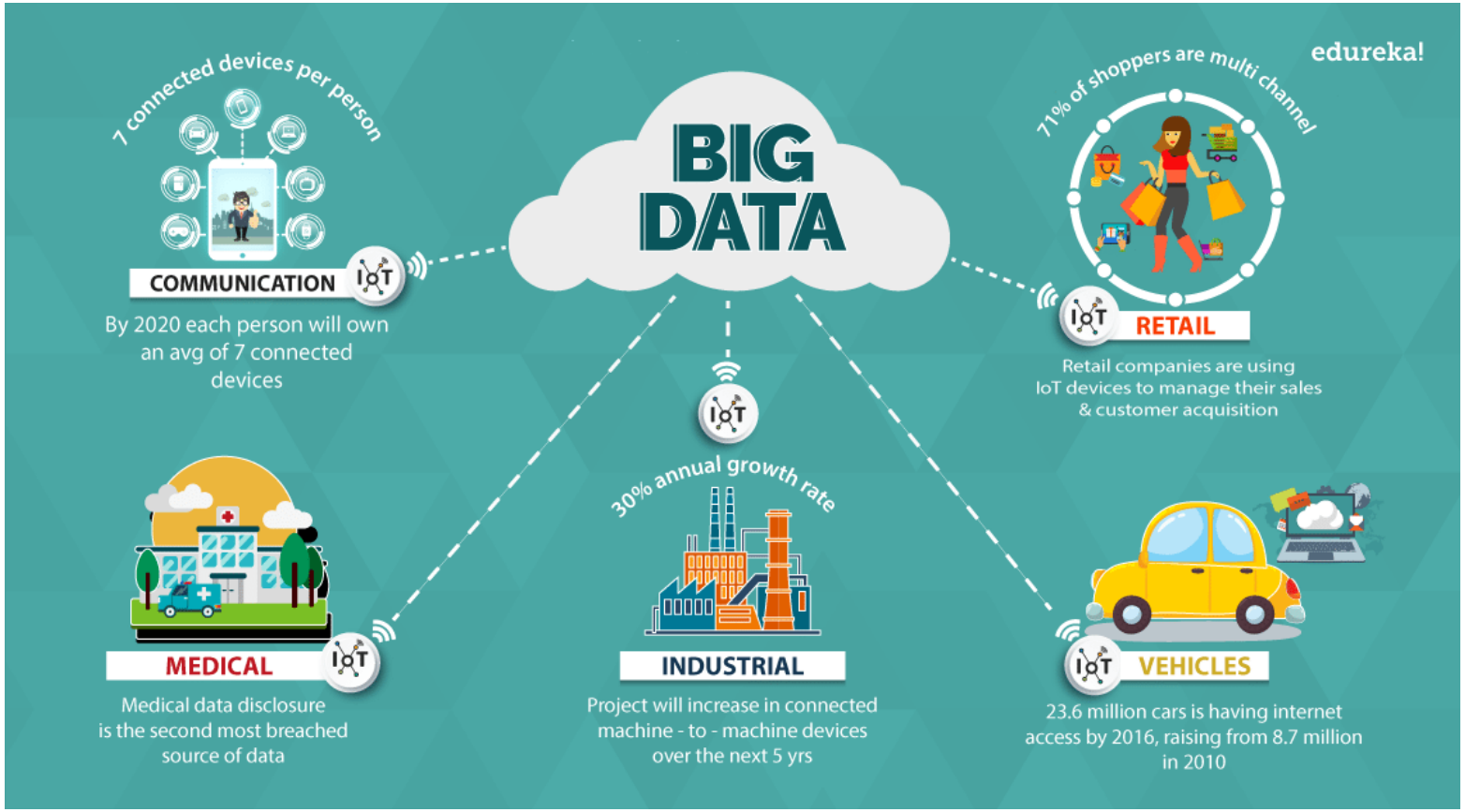
Data Analysis for SLAC Physics Richard P. Mount CHEP 2000 17

CERN



PAST BIG DATA VS NEW BIG DATA

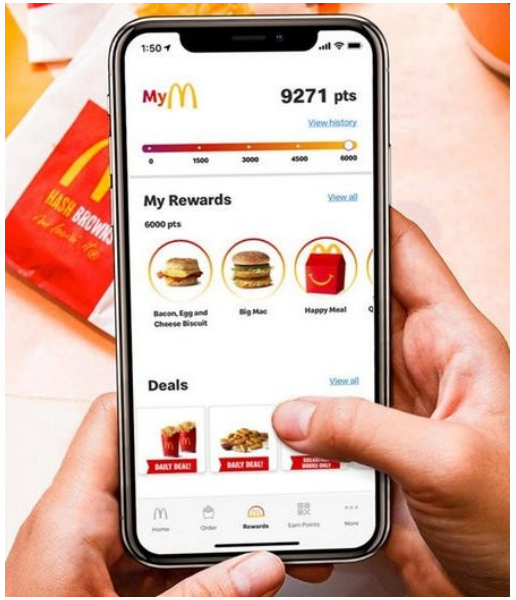
Nowadays, Big Data are heterogeneous, coming from different data sources and different domains.

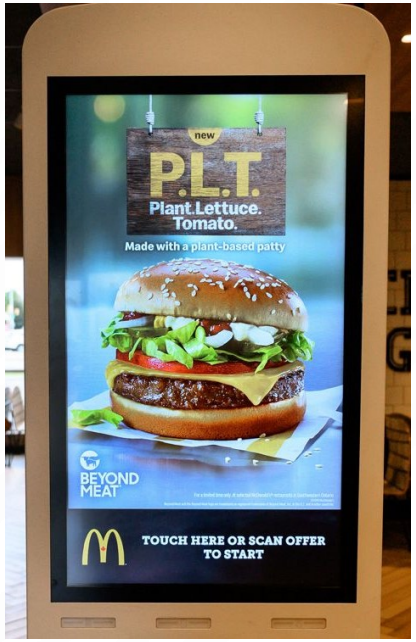


BIG DATA EXAMPLE IN HEALTHCARE

BIG DATA EXAMPLE IN MARKETING

Not only on-line data





BIG DATA IN NATURAL DISASTER MANAGEMENT

Professor Petteri Taalas
WMO Secretary-General

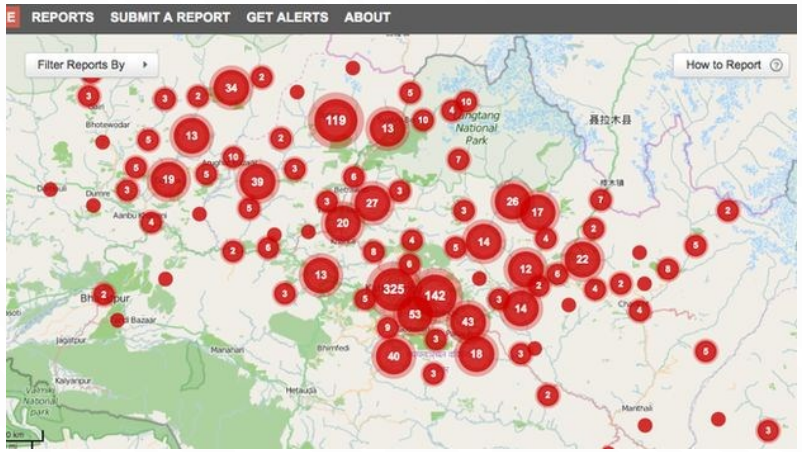
"The number of weather, climate, and water extremes are increasing and will become more frequent and severe in many parts of the world as a result of climate change"

While the number of disasters has increased by a factor of five, according to WMO, **deaths** from disasters have **decreased** by a factor of three due to an increase in early warnings and an improvement in disaster management.

BIG DATA EXAMPLES IN NATURAL DISASTER MANAGEMENT

How Big Data can help in emergency situations ?

Crisis mapping



Be prepared

Stay connected

BIG DATA & DATA STORYTELLING APPROACH

Collection of user needs

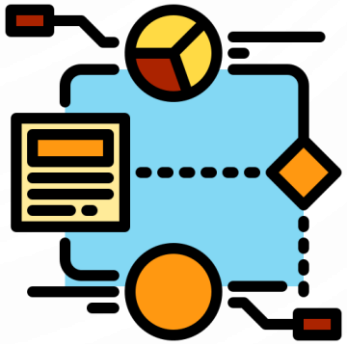
Creation of a

Data from co-creation

tivities

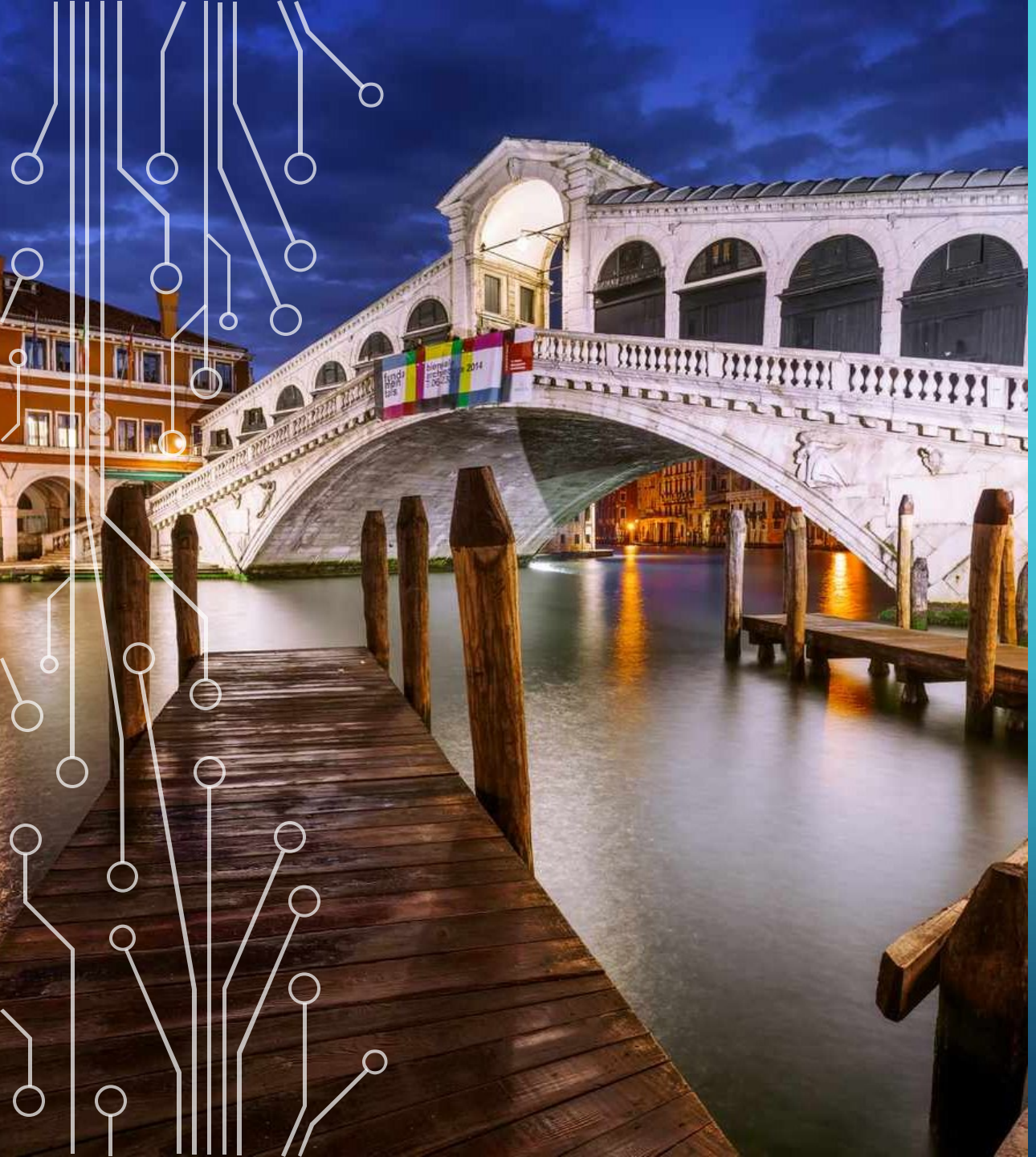
Learn, discuss and then decide

Luigi Einaudi, President of Italy (1948 – 1955)



Implementation of algorithms
to analyse those data

Visualisation of Dashboards to
take decision



CO-CREATION

Why co-creation?

As a **bridge** between the end-users and the technical team.

DATA PROPERTIES

- Findable
- Accessible
- Interoperable
- Reusable

…

- Availability
- Usability

FAIR DATA

Findable

To aid automatic discovery of relevant datasets, (meta)data should be easy to find by both humans and machines and be assigned a persistent identifier.

Accesible

Limitations on the use of data, and protocols for querying or copying data are made explicit for both humans and machines.

Interoperable

(Meta)data should use standardised terms (controlled vocabularies), have references to other (meta)data and be machine actionable.

Reusable

(Meta)data are sufficiently well described for both humans and computers to be able to understand them and have a clear and accessible data usage license.

DATA AVAILABILITY VS USABILITY

DATA ASSOCIATIONS IN EUROPE

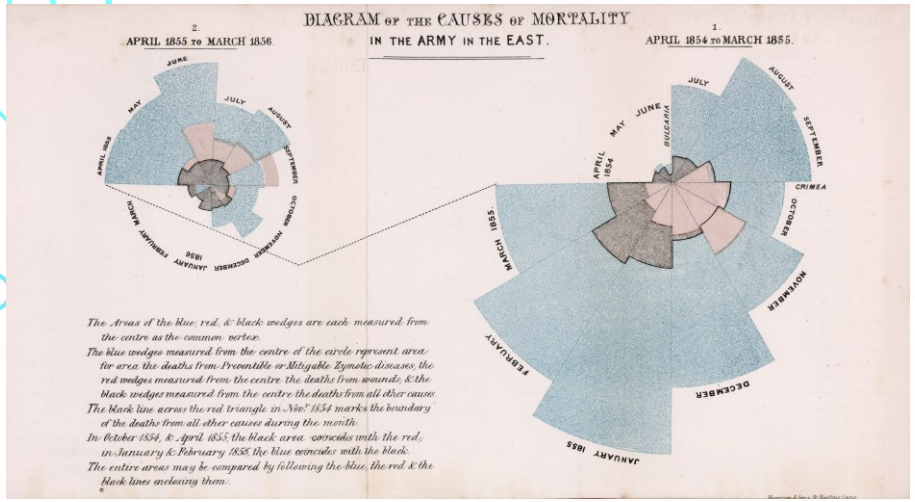
FIWARE
Open APIs for Open Minds

DATA SPACES SUPPORT CENTRE

INTERNATIONAL DATA SPACES ASSOCIATION

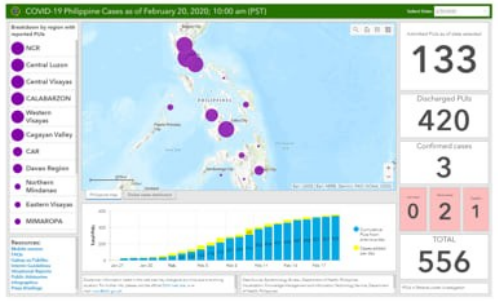
BDV BIG DATA VALUE ASSOCIATION

DATA RAPRESENTATION



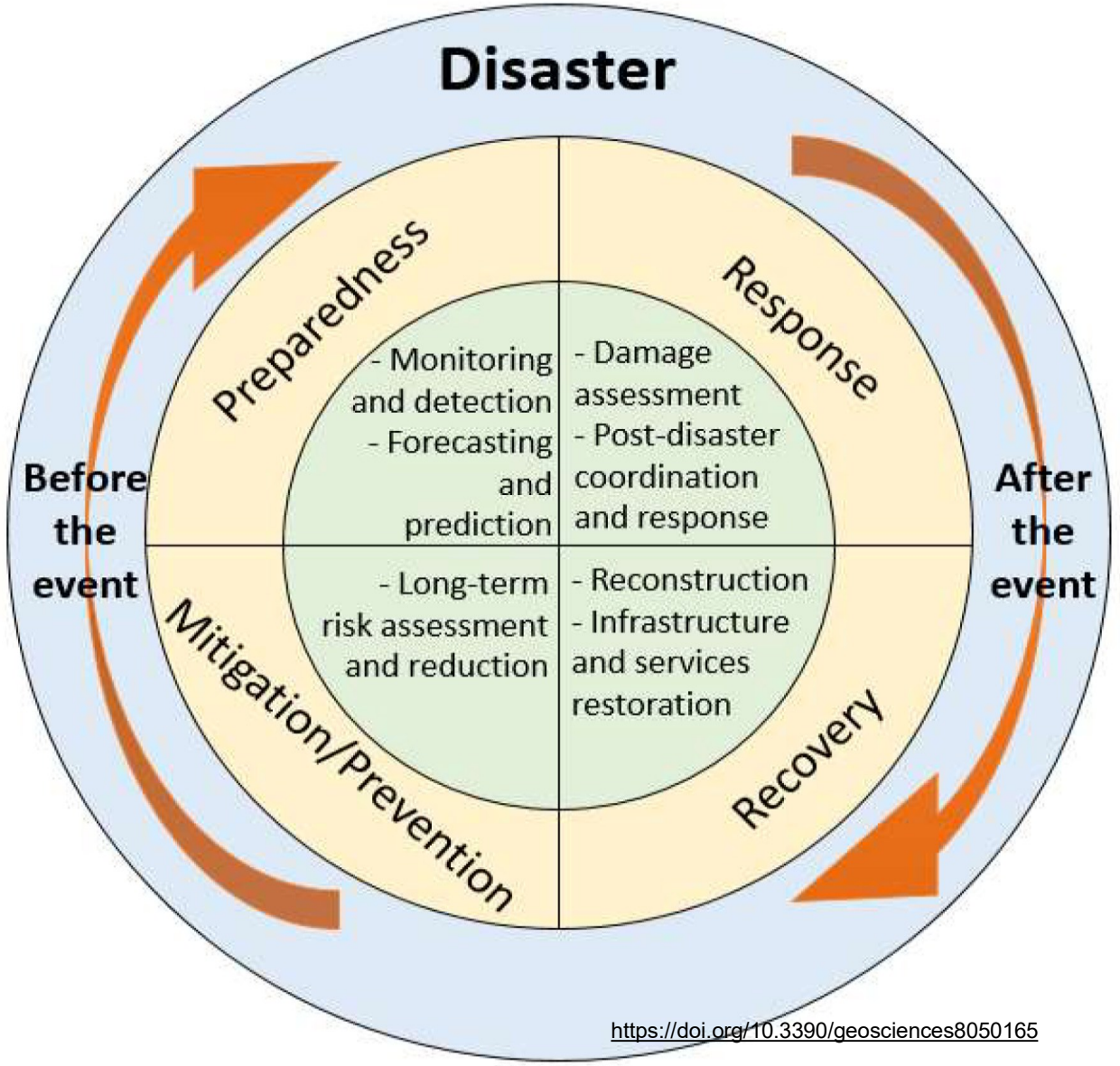
1856

2023



DATA RAPRESENTATION

BIG DATA IN NATURAL DISASTER MANAGEMENT



https://doi.org/10.3390/geosciences8050165

TECHNOLOGIES IN BIG DATA

SATELLITE IMAGES SAVE LIVES FROM SPACE

TECHNOLOGIES IN BIG DATA

Drones survey the scene

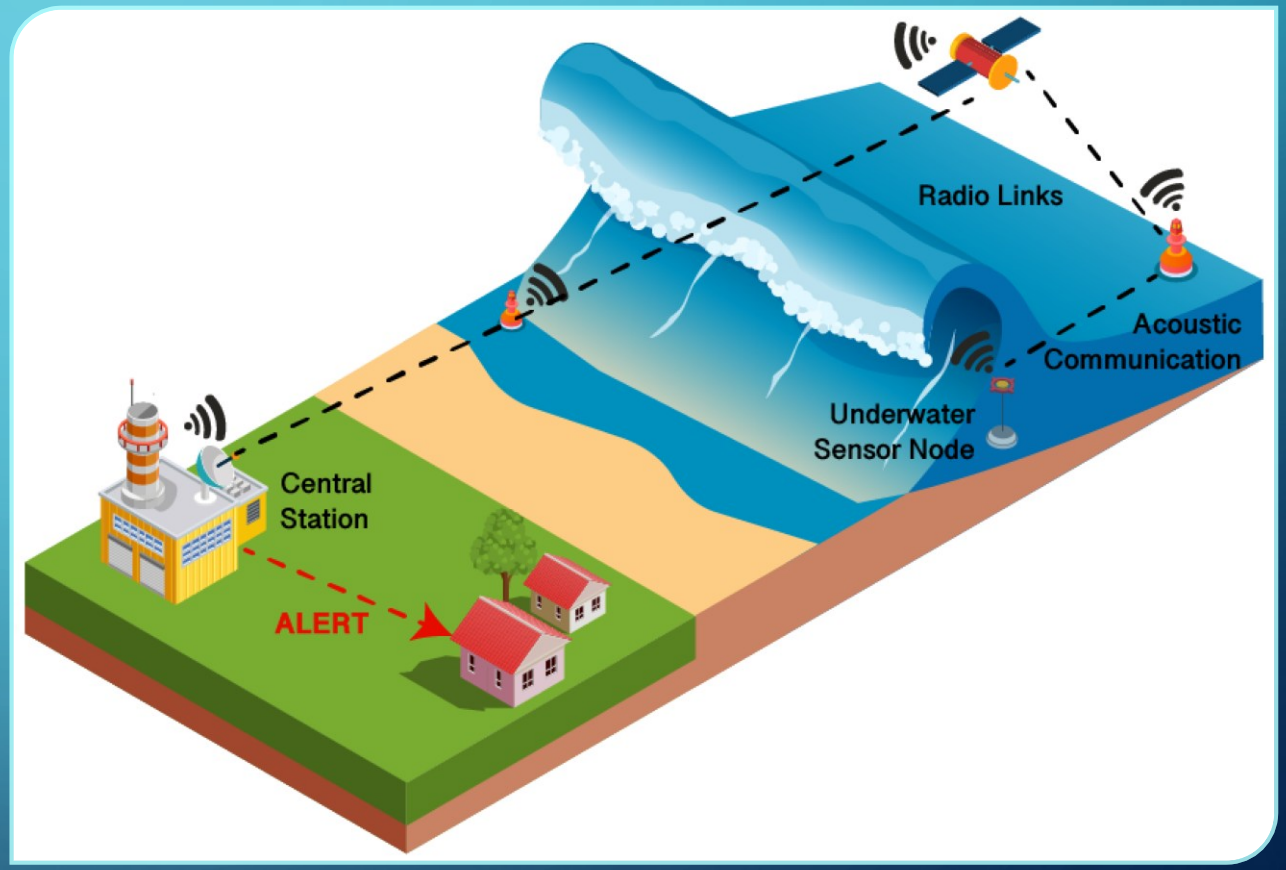
TECHNOLOGIES IN BIG DATA

SOCIAL MEDIA TURN ON THE LIGHT

TECHNOLOGIES IN BIG DATA

Sensors Sound

Early Alarms



TECHNOLOGIES IN BIG DATA

FROM STORM CLOUDS TO CLOUD COMPUTING

A Research project on the Big Data

PROJECT NAME:

TEMA
Trusted Extremely Precise Mapping and Prediction for Emergency Management

CALL:

HORIZON-CL4-2022-DATA-01-01
Methods for exploiting data and knowledge for extremely precise outcomes (analysis, prediction, decision support), reducing complexity and presenting insights in understandable way (RIA)

Project Consortium:

19 partners
from 8 European countries,
key players in the fields of data analysis, AI, modelling, drone technologies, simulation and visualization, analytics and cloud computing, as well as policy counselling
and emergency response authorities/public bodies.

Project Duration:

48 months.
Starting date: 1 December 2022

Budget:

€
11.340.223,50

TEMA

The Vision

TEMA aims to develop beyond-state-of-the-art technologies for facilitating Natural Disaster Management, by dynamically exploiting **data sources and Artificial Intelligence technologies** in order to provide an accurate assessment of an evolving crisis situation.

The **goal** is to deliver a technical solution that is supportive in disaster response and management by **bringing situational data** to relevant end-users, enabling transferability to tackle different disaster types in various geographic regions, thus providing the **relevant information** that can help make the best possible **operative decisions**.

TEMA

Case Studies

Germany

1. Central-European Regional Floods Pilot site: Bavaria

Greece

2. Mediterranean Flash Floods Pilot site: Municipality of Mantoudi-Limni-Agia Anna

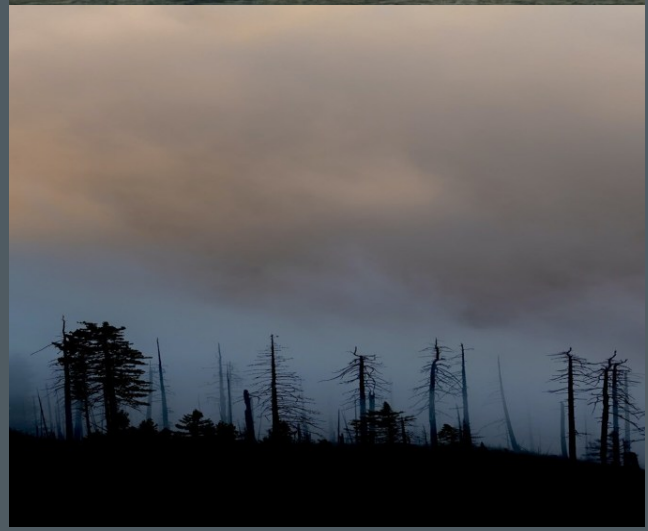
.

Italy

3. Mediterranean Forest Fires Pilot site: Montiferru (Sardinia)

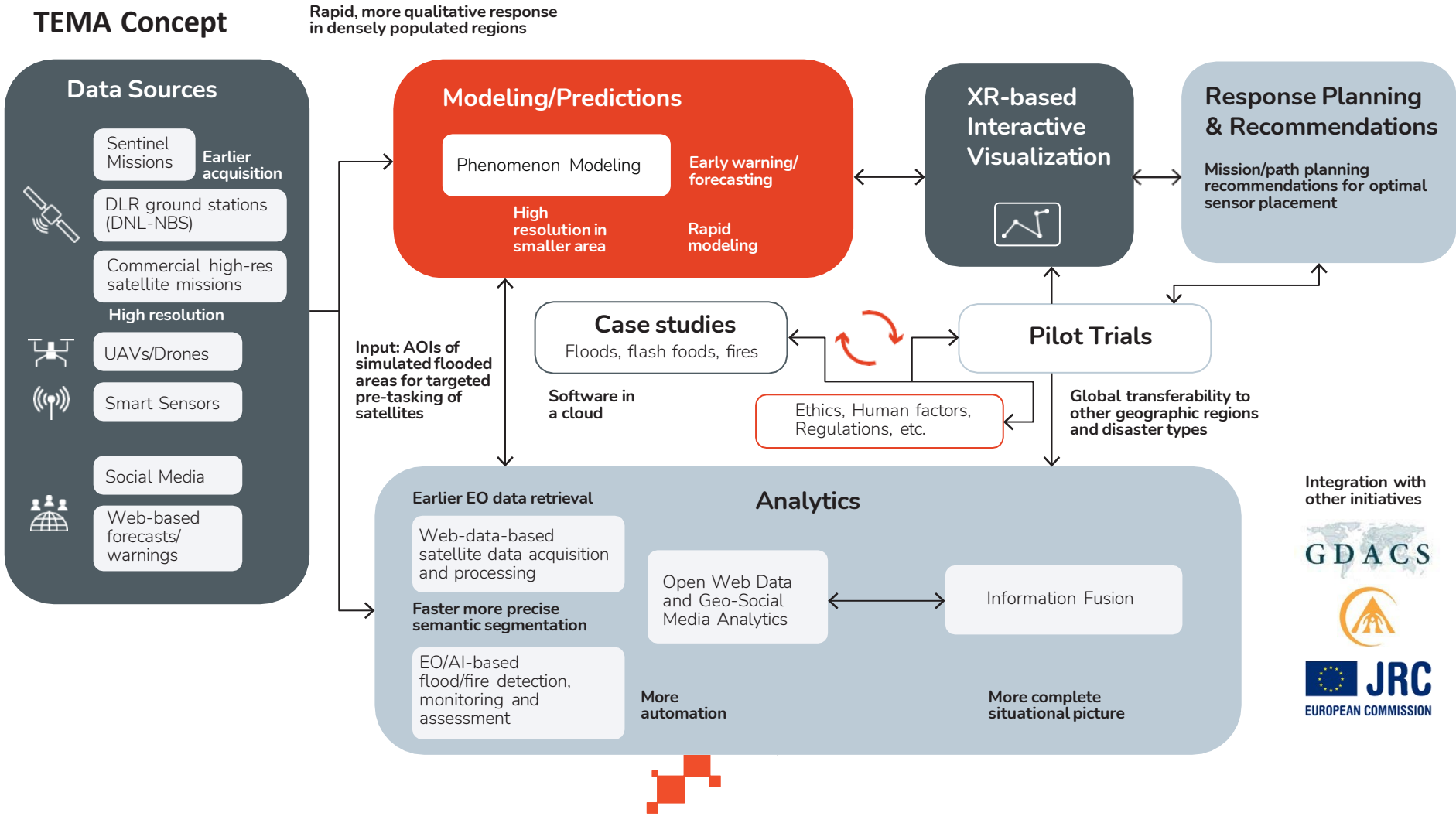
Finland

4. Finnish Forest Fires Pilot site: Kainuu area



TEMA

Concept


TEMA Concept
Rapid, more qualitative response in densely populated regions
Data Sources
Sentinel Missions
Earlier acquisition
DLR ground stations (DNL-NBS)
Commercial high-res satellite missions
High resolution
UAVs/Drones
Smart Sensors
Social Media
Web-based forecasts/ warnings
Modeling/Predictions
Phenomenon Modeling
Early warning/ forecasting
High resolution in smaller area
Rapid modeling
XR-based Interactive Visualization
Response Planning & Recommendations
Mission/path planning recommendations for optimal sensor placement
Input: AOIs of simulated flooded areas for targeted pre-tasking of satellites
Case studies
Floods, flash foods, fires
Software in a cloud
Pilot Trials
Ethics, Human factors, Regulations, etc.
Global transferability to other geographic regions and disaster types
Analytics
Earlier EO data retrieval
Web-data-based satellite data acquisition and processing
Faster more precise semantic segmentation
EO/AI-based flood/fire detection, monitoring and assessment
Open Web Data and Geo-Social Media Analytics
Information Fusion
More automation
More complete situational picture
Integration with other initiatives
GDACS
JRC
EUROPEAN COMMISSION

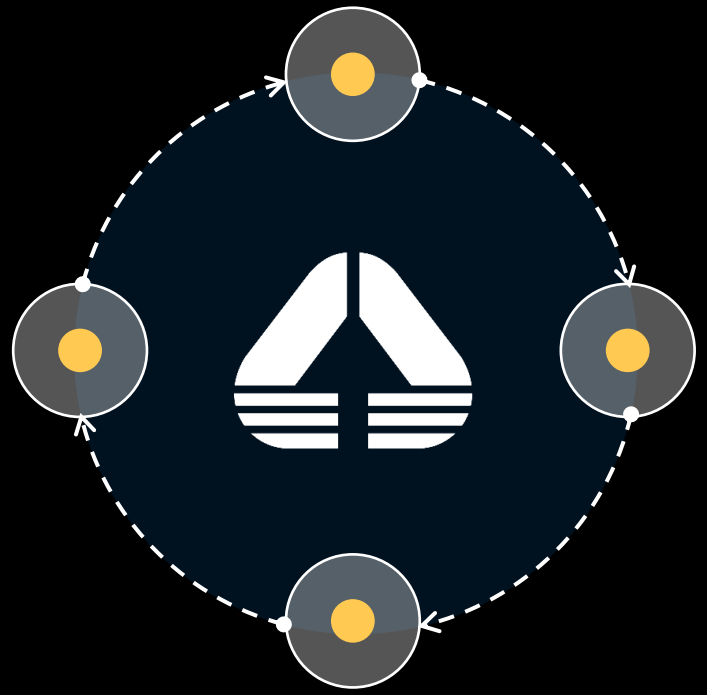
Digital Experience: Digital Enabler

Transforming Data into Knowledge through a data “ecosystem” platform

Digital Enabler (DE) is an ecosystem platform that allows to accelerate time to market overcoming major data-related business challenges

Organizations struggle to drive business decision without factoring in external 'ecosystem' data

Siloed and segregated data prevent value delivery and effective time to market



Increasingly heterogeneous data sources are pushing Big Data and AI as key business enablers

Growing security requirements incl. IoT Digital Identity and asset protection

DE is a Data Management ecosystem platform that enables new data economy business models leveraging a complete suite of accelerators

Digital Enabler
Engineering's Ecosystem Platform

Main features

Cloud native **scalable** ecosystem platform that can be **easily composed (composable)**

Independent and interoperable tools allow to accelerate the development of **data oriented vertical applications**

Based on **opensource** software, it complies with **data interoperability standards** and with the **GDPR**

Main advantages

Industry, domain & technology independent, deployable everywhere, on-prem, on public/private cloud, on hybrid clouds, **no lock-ins**

Usability and flexibility for **reducing time to market** and improving responsiveness to business demands

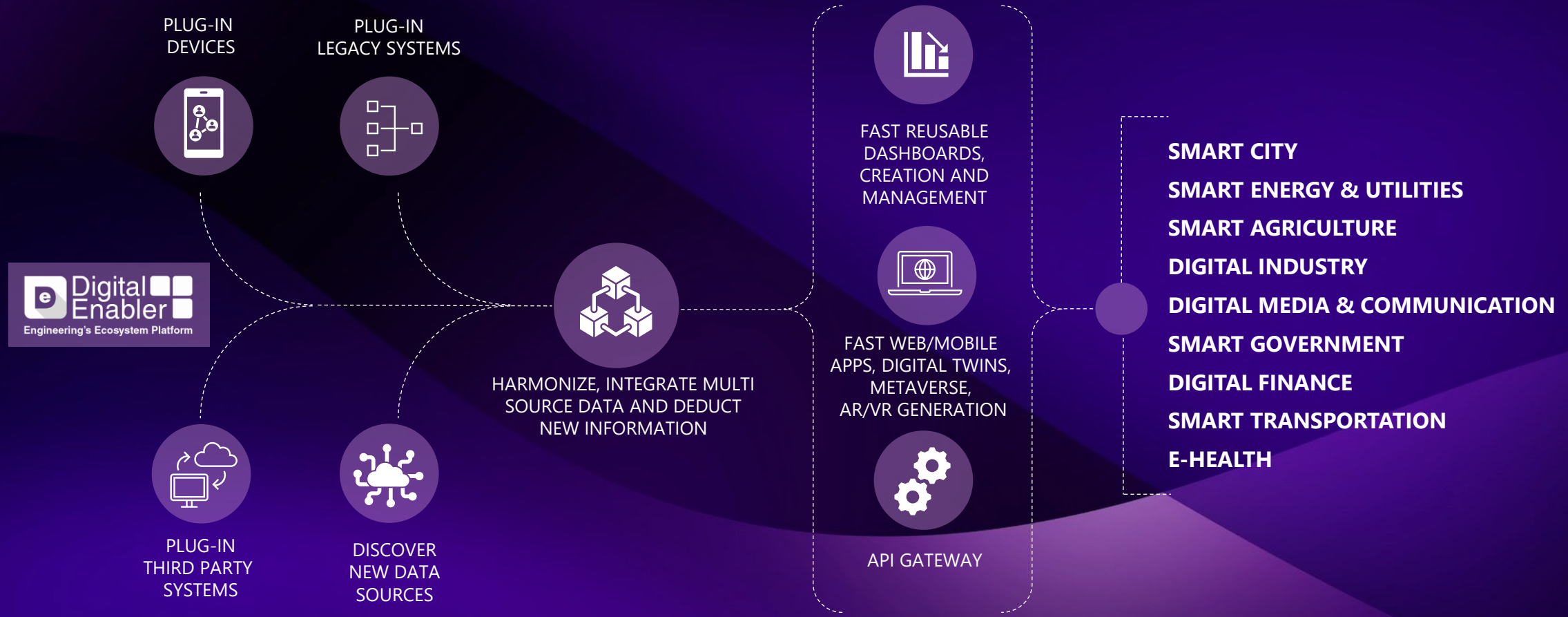
Interoperability among third party systems and increasing value of existing assets by promoting **new business solutions**

OUR ACCELERATORS

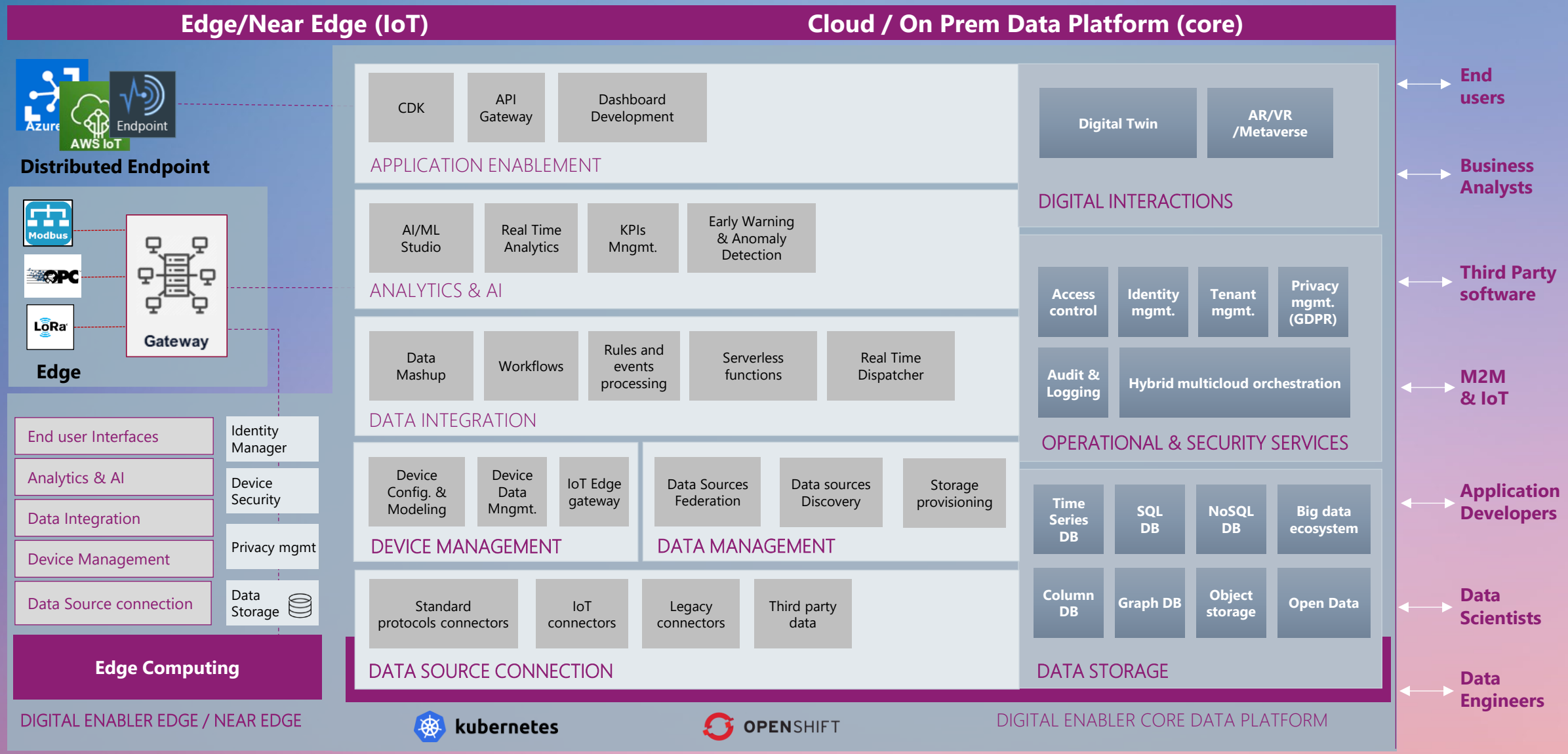
- Data discovery
- Data collection from heterogeneous sources
- Low/No code data integration
- IoT and Edge management
- Rule Engine, Adv. analytics, AI, Serverless
- Digital interaction

DE leverages a suite of 'accelerators' to gather and gain data insight to drive decisions and monitor outcomes

DE is a solution to make data easily exploitable via an integrated and efficient platform for enabling different digital ecosystem domains



Digital Enabler: functional Platform Landscape



Why choose Engineering Digital Enabler?

Microservices/containers based architecture that makes the platform scalable and composable

Not intrusive nor exclusive for existing systems thanks to functionalities and connectors that make easy the plug-in with 3rd party and legacy systems

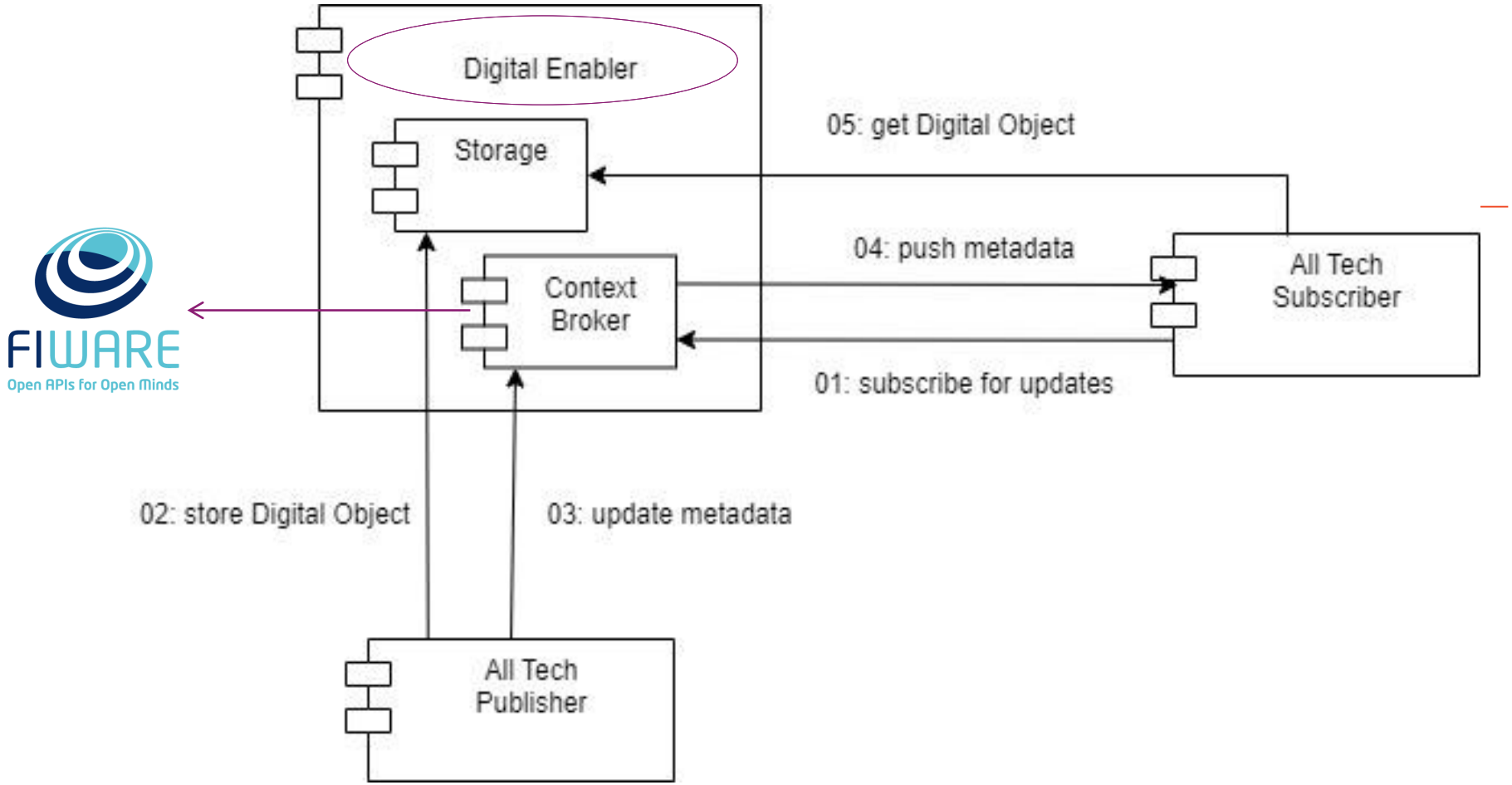
Specific expertise on customer domain and setting of **technological solution application**

Low/No code Tool suite to accelerate the development of new apps and **enablement of IoT at different levels (Edge, near Edge, Cloud)**

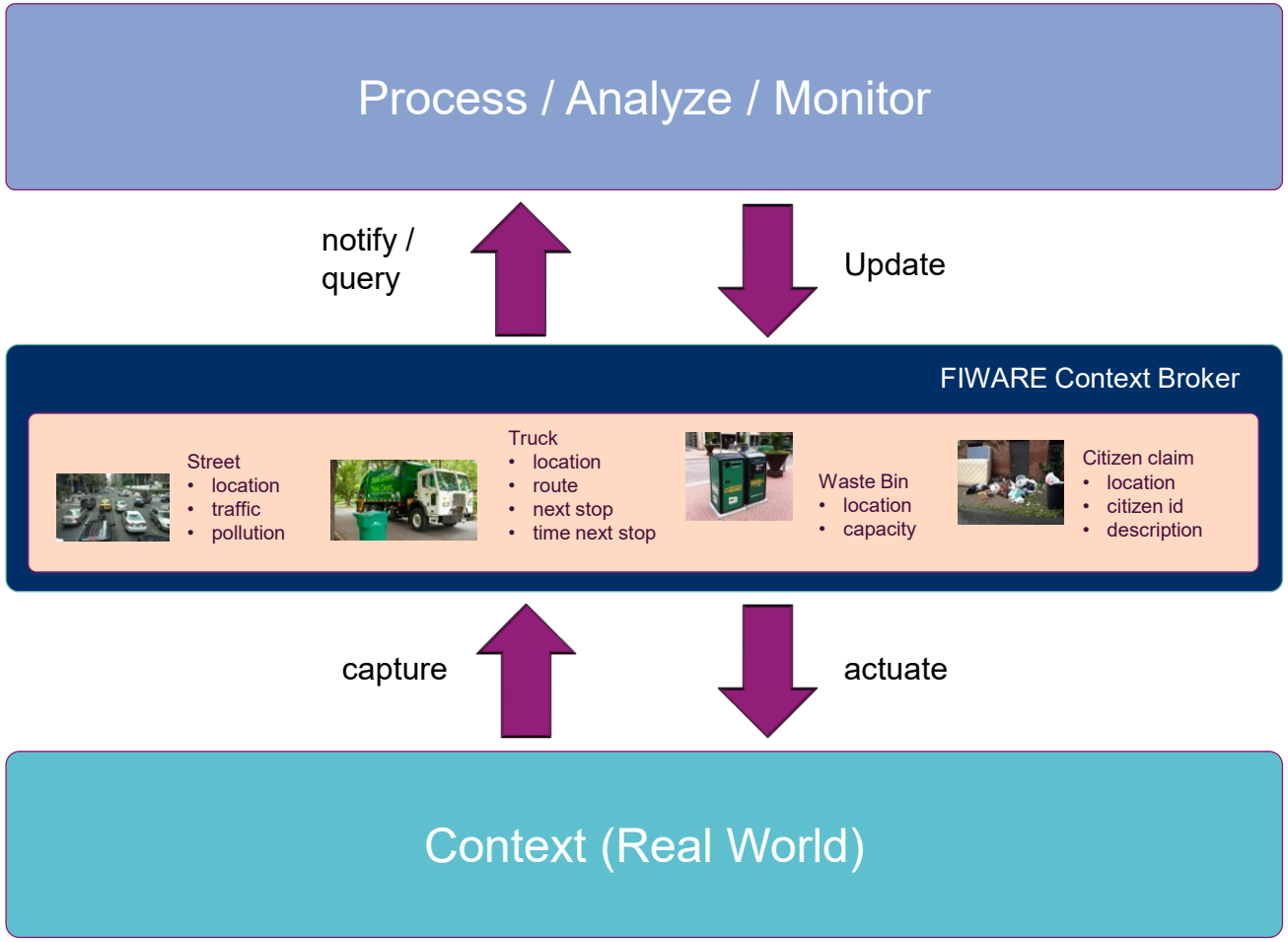
No lock-ins with vendors and technologies as DE is based on **opensource** solutions and deployable everywhere (i.e. on site, cloud, hybrid)

Designed to quickly and easily enrich and mash-up data, through low code tools to harmonize data in standard data models and deduct new information through AI

Components adopted in TEMA to support Natural Disasters management



FIWARE Platform Architecture overview



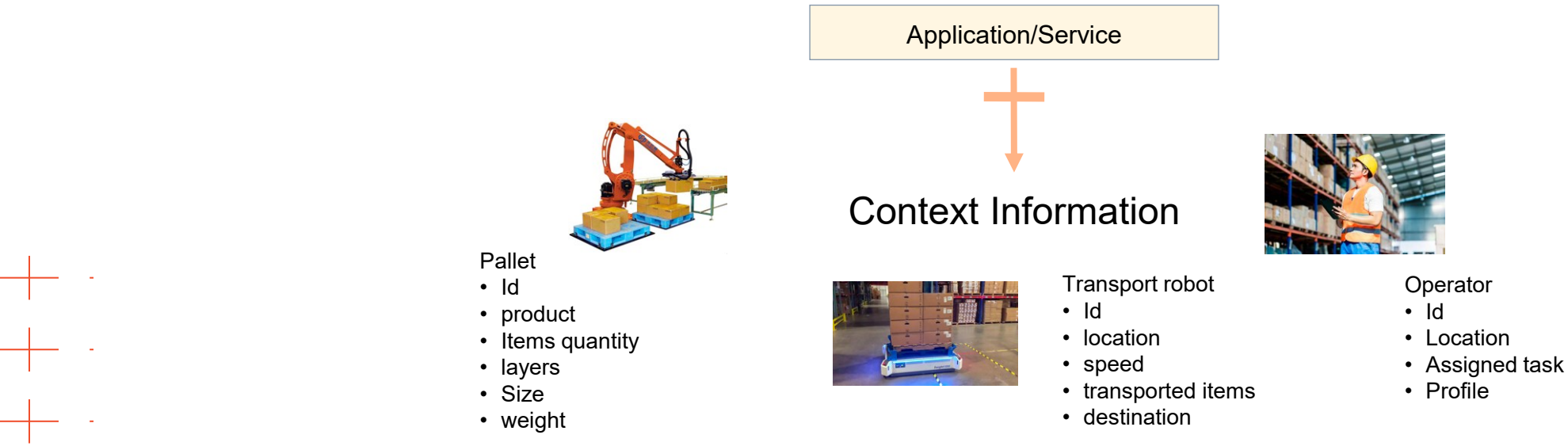
https://www.slideshare.net/slideshow/orion-context-broker-20230606/258274971

ETSI NGSI-LD: A standard API for Context Information Management

The ETSI NGSI-LD API is a simple yet powerful public, royalty-free standard API for Context Information Management

Simple: A RESTful API which any web programmer learns how to use in one day

Yet **powerful**: It supports geo-queries, Linked Data (JSON-LD), subscription/notification, …

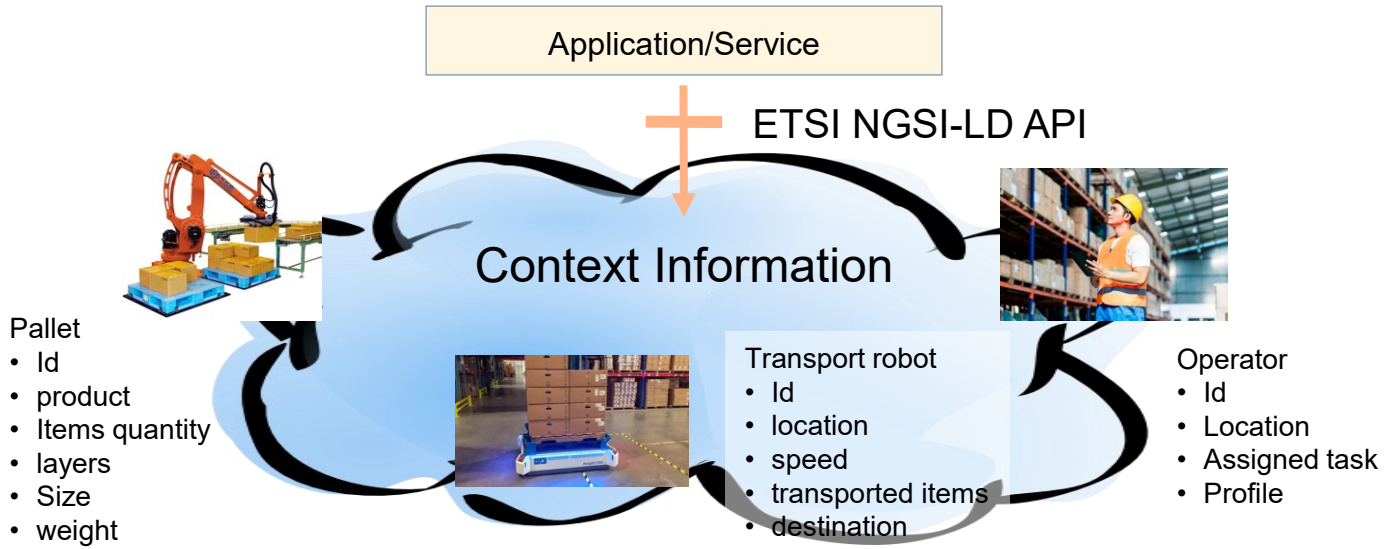


Context Broker

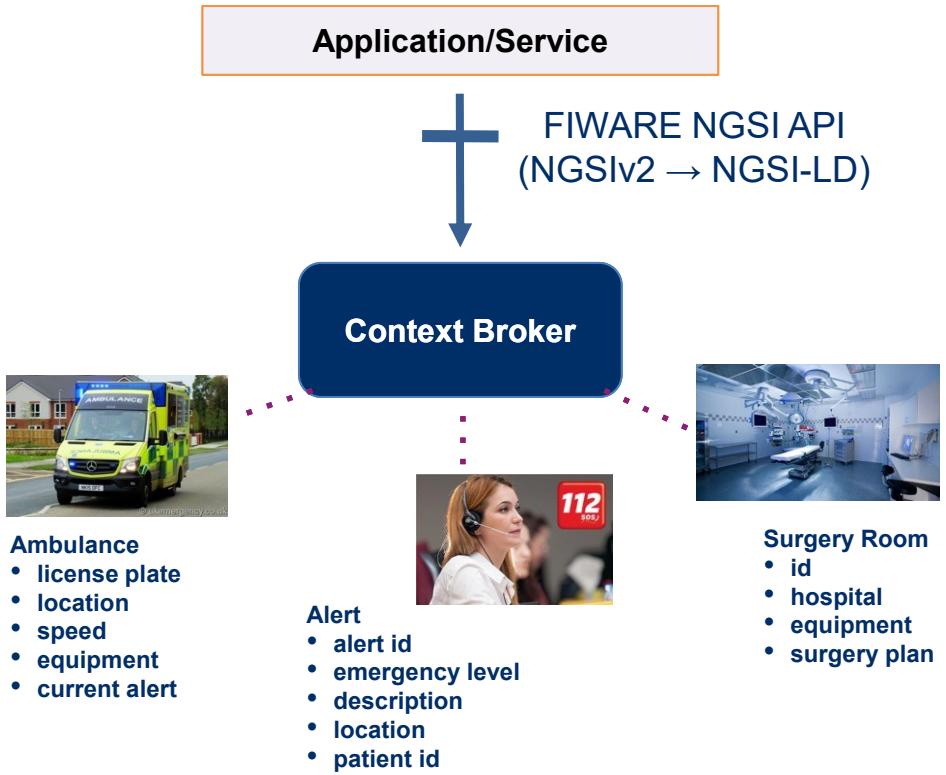
NGSI-LD API servers are usually referred as Context Brokers

A Context Broker is associated to a transport end point

A Context Broker does not necessarily hold the data you are looking for but “knows” how that data can be obtained. Strictly speaking, they provide access to data (using the NGSI-LD API)



Data Exchange API: ETSI NGSI-LD



NSGI-LD

Next Generation Service Interface-Linked Data (NGSI-LD) is an open **standard** for context information management developed to facilitate the exchange of information between applications in the context of the Internet of Things (IoT).

JSON-LD

- **JSON** is the data exchange format. However, it is not so easy for machines to read. The attributes of a JSON can have different meanings even among humans themselves.

- For example, consider a JSON that has "***name***" as an attribute. For a person that attribute could mean the person's native name. For another, it could mean the person's stage name. For yet another the username of the person.

- What has been done is to define a **JSON extension** in which it is possible to **annotate** a piece of JSON with **additional information**.

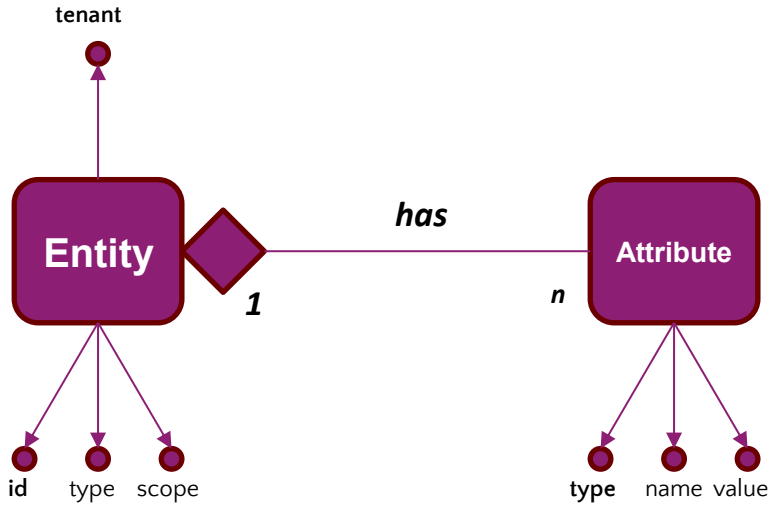
- Instead of having the *name* attribute **without context (JSON)** we can make that **attribute** a **URI** rather than a simple key-value pair. In this way, via the URI that name attribute actually represents the person's real name: **easier for people and machines to understand it**.

NGSI-LD
Entities and Attributes

NGSI-LD entity's **type** serves as a representation of the entity's **class or category**.

In NGSI-LD, the **attributes** of an entity represent the **characteristics** or **properties** of the entity.
In adherence to the Entity-Attribute-Value (EAV) model, an entity in NGSI-LD consists of:

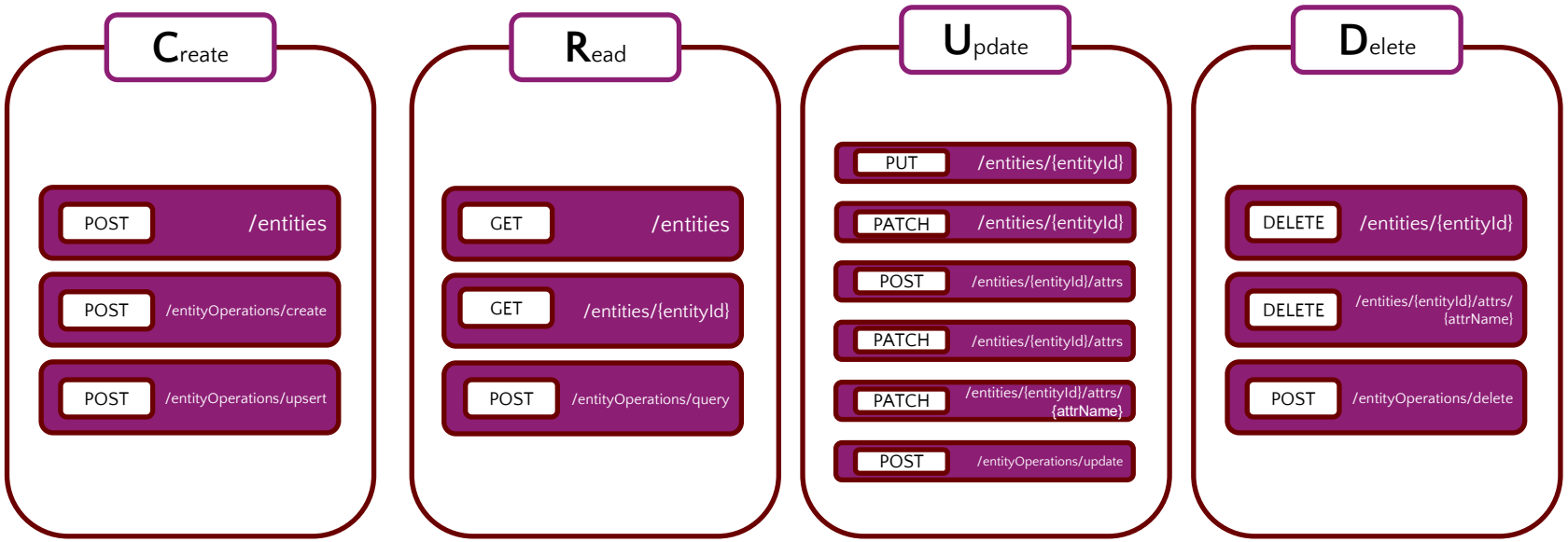
- Entity: A tangible object in the physical world, such as a car or weather station, capable of being defined and tracked.
- **Attribute**: Descriptors that define an entity's specific characteristics.
- **Value**: The current measurement or state of an attribute.



NGSI-LD – CRUD on Entities/Attributes

The Base Path of an NGSI-LD **Context Broker** is:

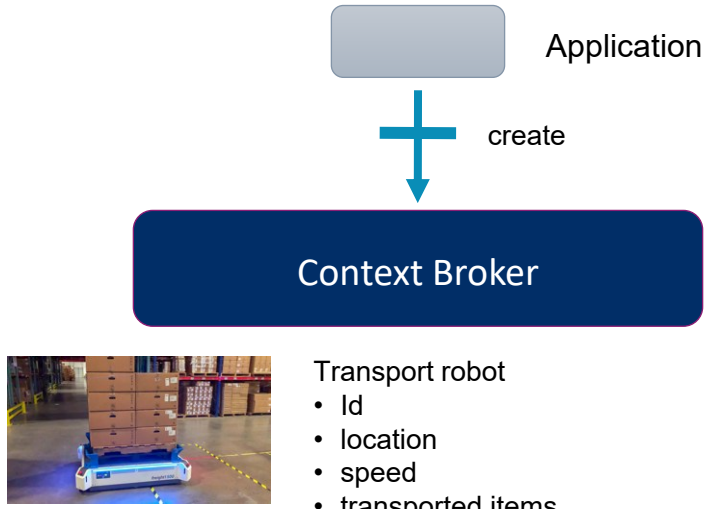
https://<YOUR_DOMAIN>/ngsi-ld/v1



Creation of entities

Application can create entities and give initial values to their attributes (properties/relationships)

Multiple entities can be created in a single request



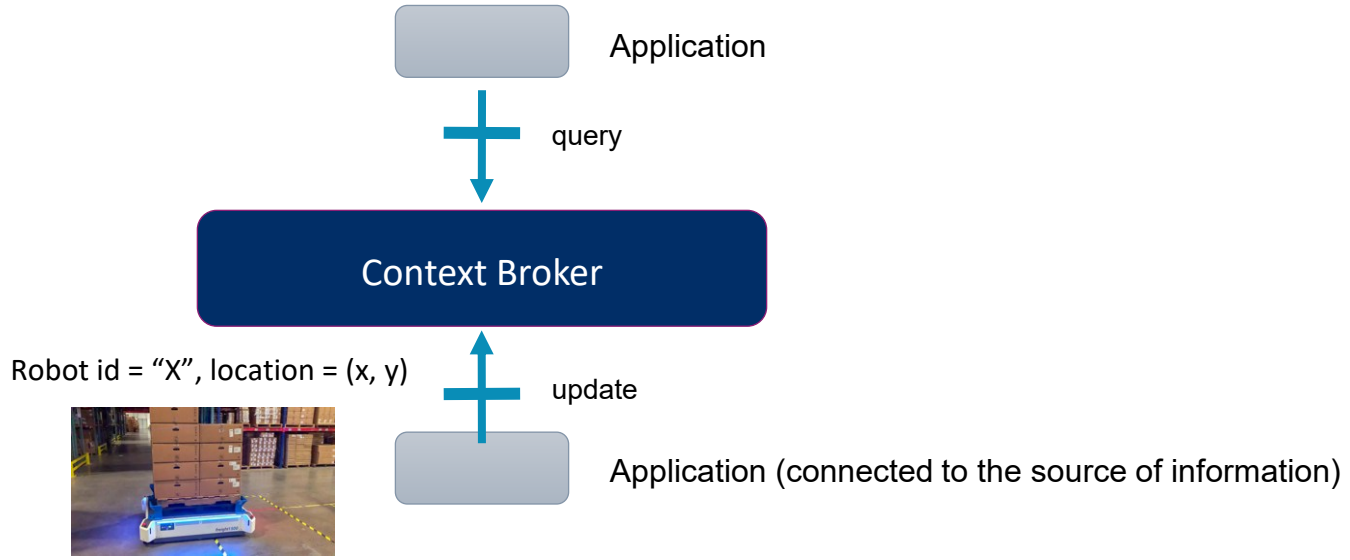
Transport robot
- Id
- location
- speed
- transported items
- destination

Updates on entities

You can update an attribute of a given entity with a request (simplest case)

You can update a set of attributes on one or more entities in a single request

Attributes can be added to a given entity type (effectively extending the underlying data model) at any time without the need to re-initiate

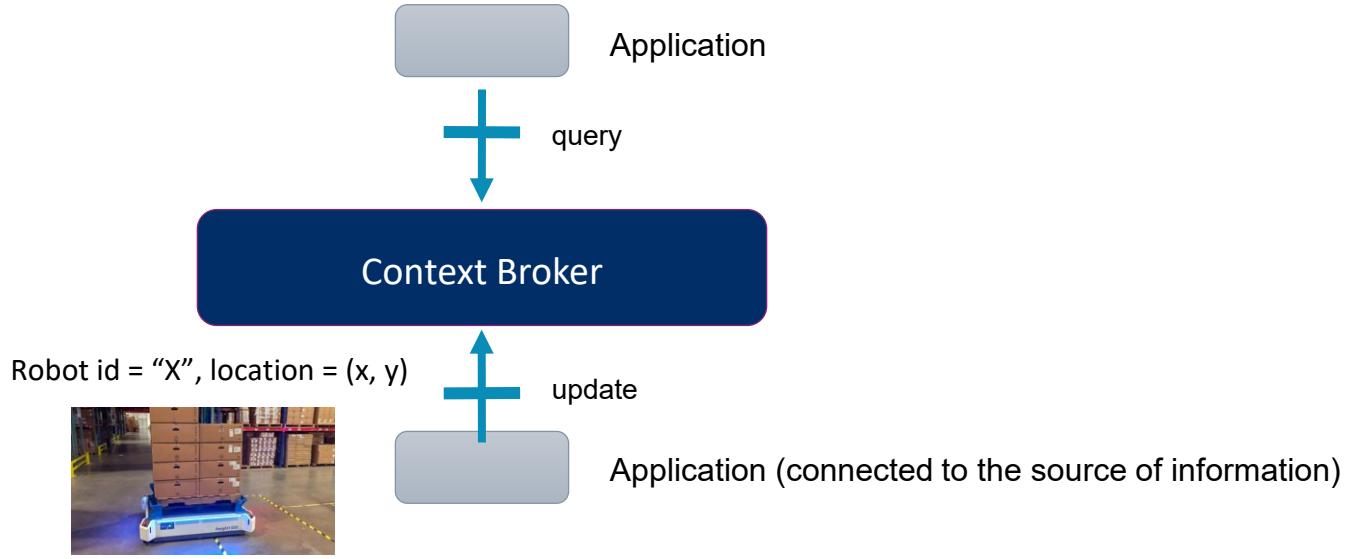


Queries on entities

You can query for the value of an attribute of a given entity with a request (simplest case)

You can query for a set of attributes on one or more entities in a single request

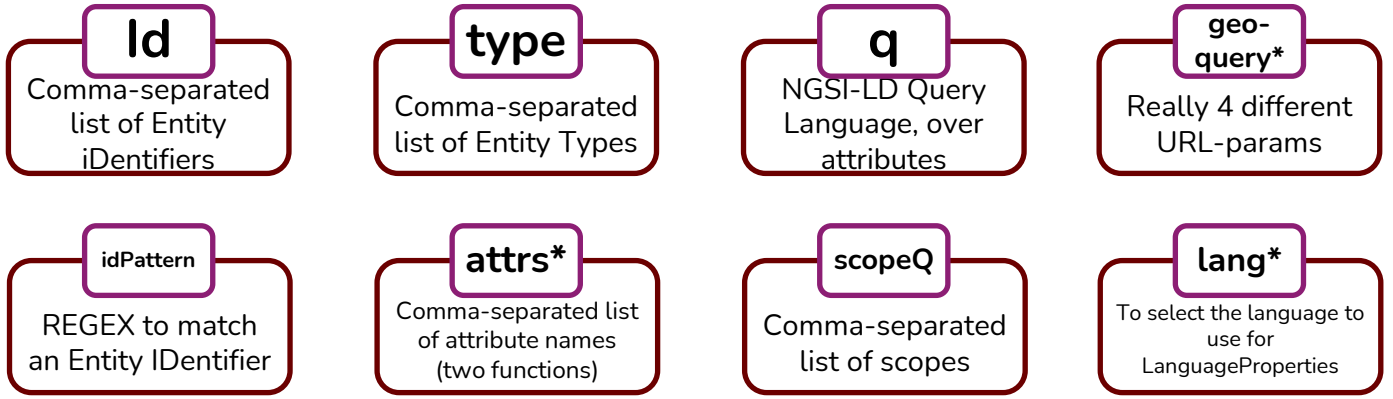
You can specify filters/conditions on your query to refine what entities you are querying information about



NGSI-LD – Options for Querying Entities

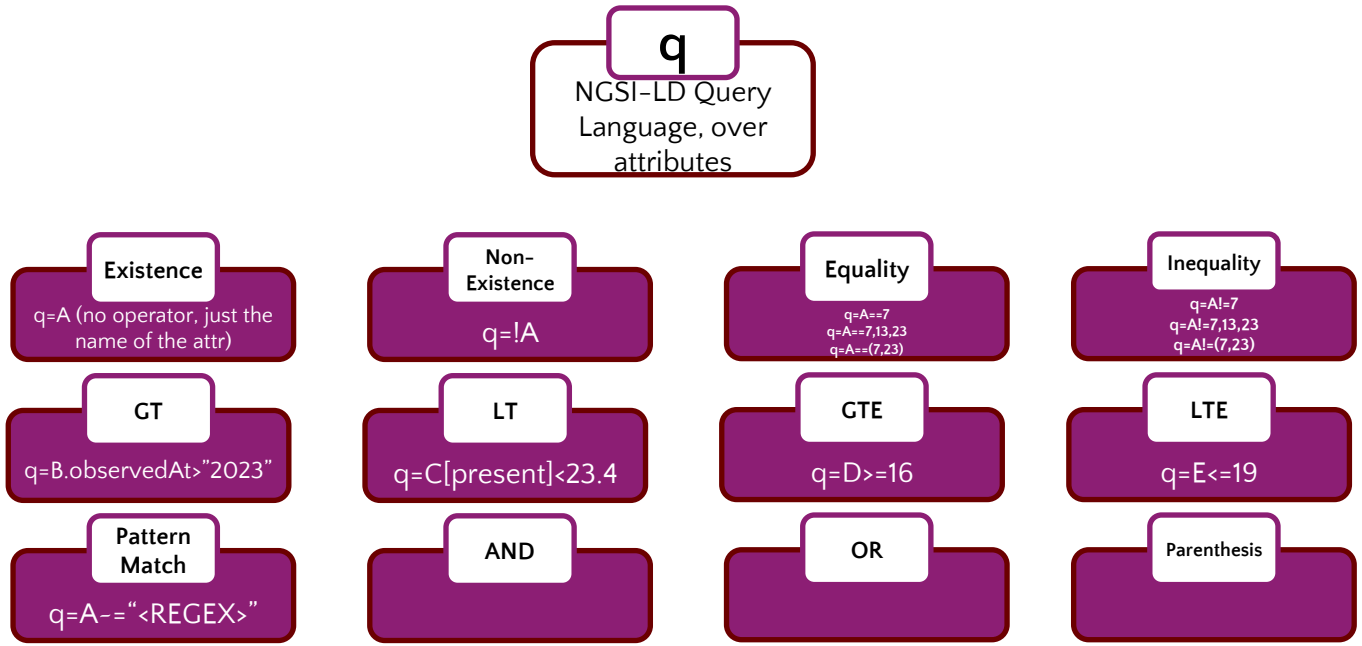
Filters are criteria or parameters used to restrict and select specific entities or information during a query of entities.

When using a “POST Query”, where they are included in the body of the request (*).



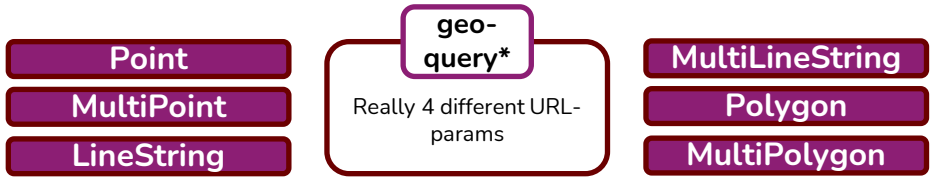
NGSI-LD – Query Language: Operators

The “NGSI-LD Query Language” is a query language used in the context of NGSI-LD to **perform advanced queries on entity attributes**.



NGSI-LD – Query Language: Operators

The “geo-query” is a functionality within NGSI-LD that allows **spatial queries to be performed to select entities based on geographical criteria**. This functionality is particularly useful when managing geographical or geospatial location data.



Each geometry type has its own “**georels**” that specifies the **geographical relationship** between the entity and the specified point. (e.g., “near;minDistance==X” for “Point”). See: https://www.rfc-editor.org/rfc/rfc7946.

Example (GET)

/entities?geoProperty=location&geometry=Point&coordinates=[1,2]&georel=near;maxDistance==100

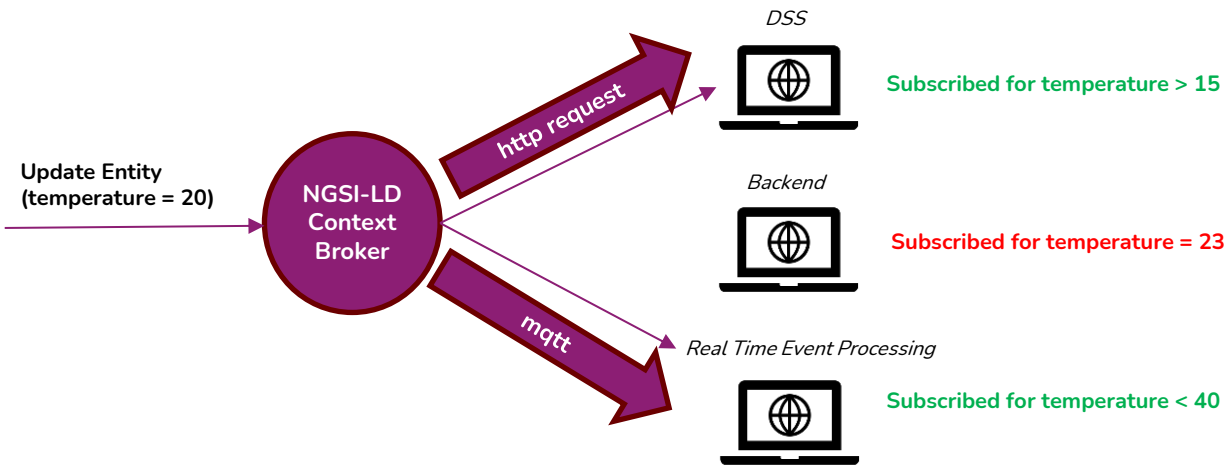
Example (POST)

```
{
  "geo-query": {
    "geometry": {
      "type": "Point",
      "coordinates": [longitude, latitude]
    },
    "georel": "near",
    "maxDistance": distance
  }
}
```

NGSI-LD – Subscriptions

Subscriptions are mechanisms that allow users **to receive real-time notifications when information associated with entities meets certain specified conditions**.

A Subscription defines “what to get notified for” and “where and in what format to send the notification”. The notification are sent by CB **via HTTP or MQTT Protocols**.



NGSI-LD – Subscriptions (Example)

```
curl -L -X POST 'http://localhost:1026/ngsi-ld/v1/subscriptions/' \
-H 'Content-Type: application/ld+json' \
-H 'NGSILD-Tenant: openiot' \
--data-raw '{
  "description": "Notify me of low feedstock on Farm:001",
  "type": "Subscription",
  "entities": [{"type": "FillingLevelSensor"}],
  "watchedAttributes": ["filling"],
  "q": "filling>0.6;filling<0.8;controlledAsset==%22urn:ngsi-
ld:Building:farm001%22",
  "notification": {
    "attributes": ["filling", "controlledAsset"],
    "format": "keyValues",
    "endpoint": {
      "uri": "http://tutorial:3000/subscription/low-stock-farm001",
      "accept": "application/json"
    }
  },
   "@context": "http://context/ngsi-context.jsonld"
}'
```

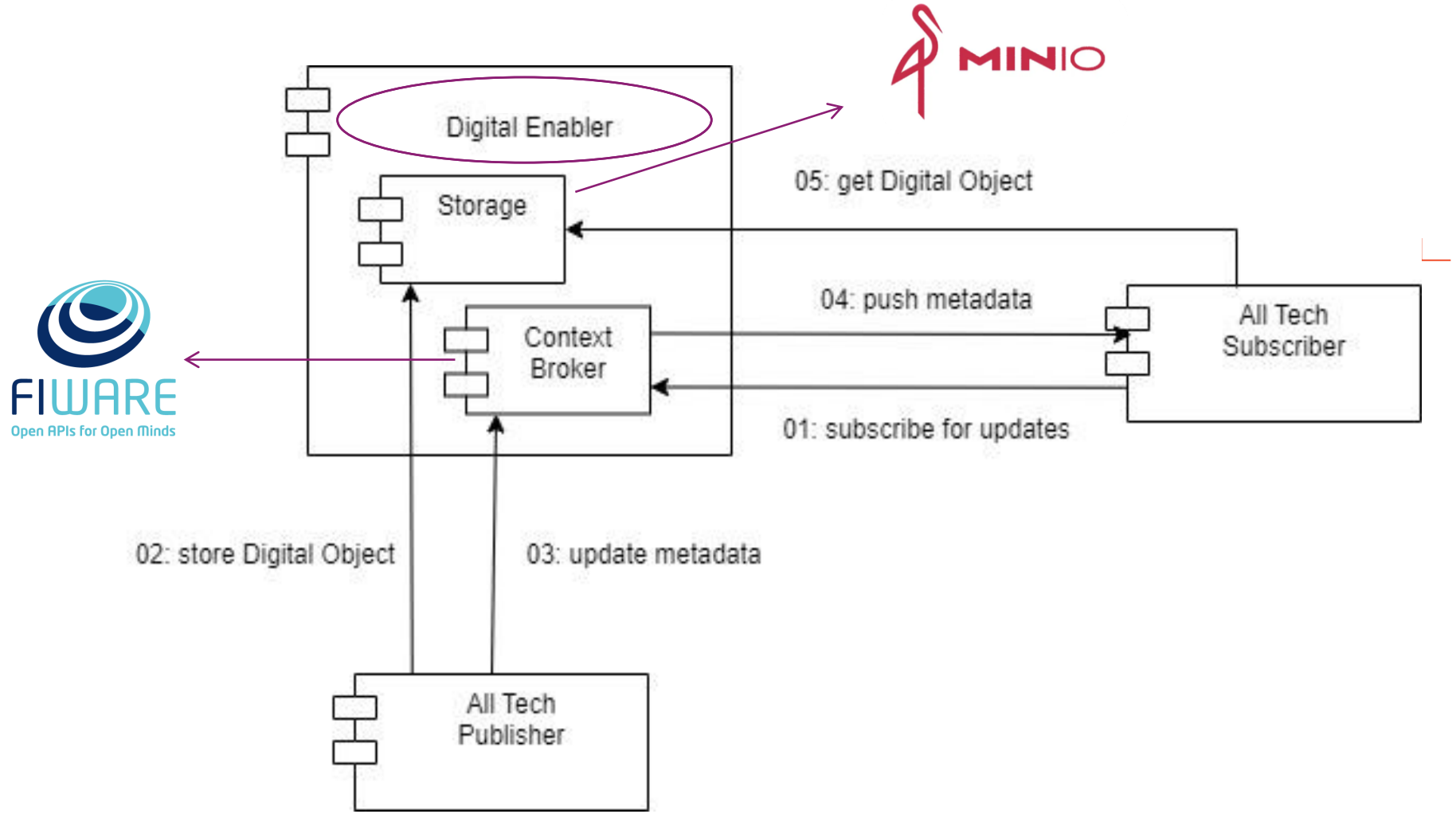
URL for Subscription

Trigger Condition(s)

Target

Data type to be reiceived

Components adopted in TEMA to support Natural Disasters management



What is MinIO

MinIO is a **high performance, software defined, distributed object storage** server, designed for peta-scale data infrastructure.

It was built from scratch with the **private cloud** as its target.

MINIO SERVER

MINIO CLIENT

MINIO SDK

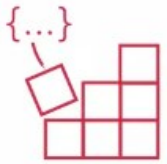
MINIO

https://www.slideshare.net/slideshow/minio-january-2020-briefing/225615613

MinIO Deployment Use Cases

Big Data/Machine Learning environments

HDFS replacements



High performance data lake/warehouse infrastructure

Cloud native applications (replacing file + block)

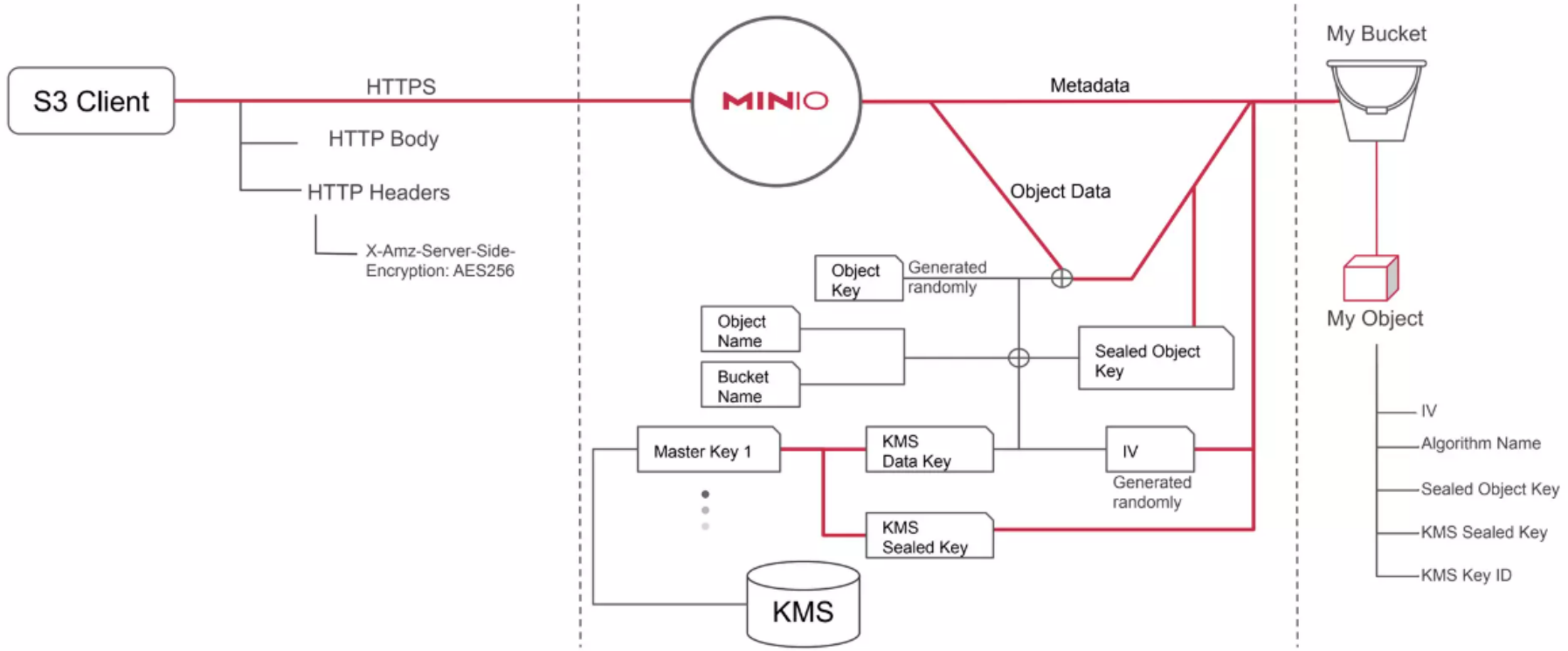
Multi-cloud environments (portability)

Endpoint for streaming workloads

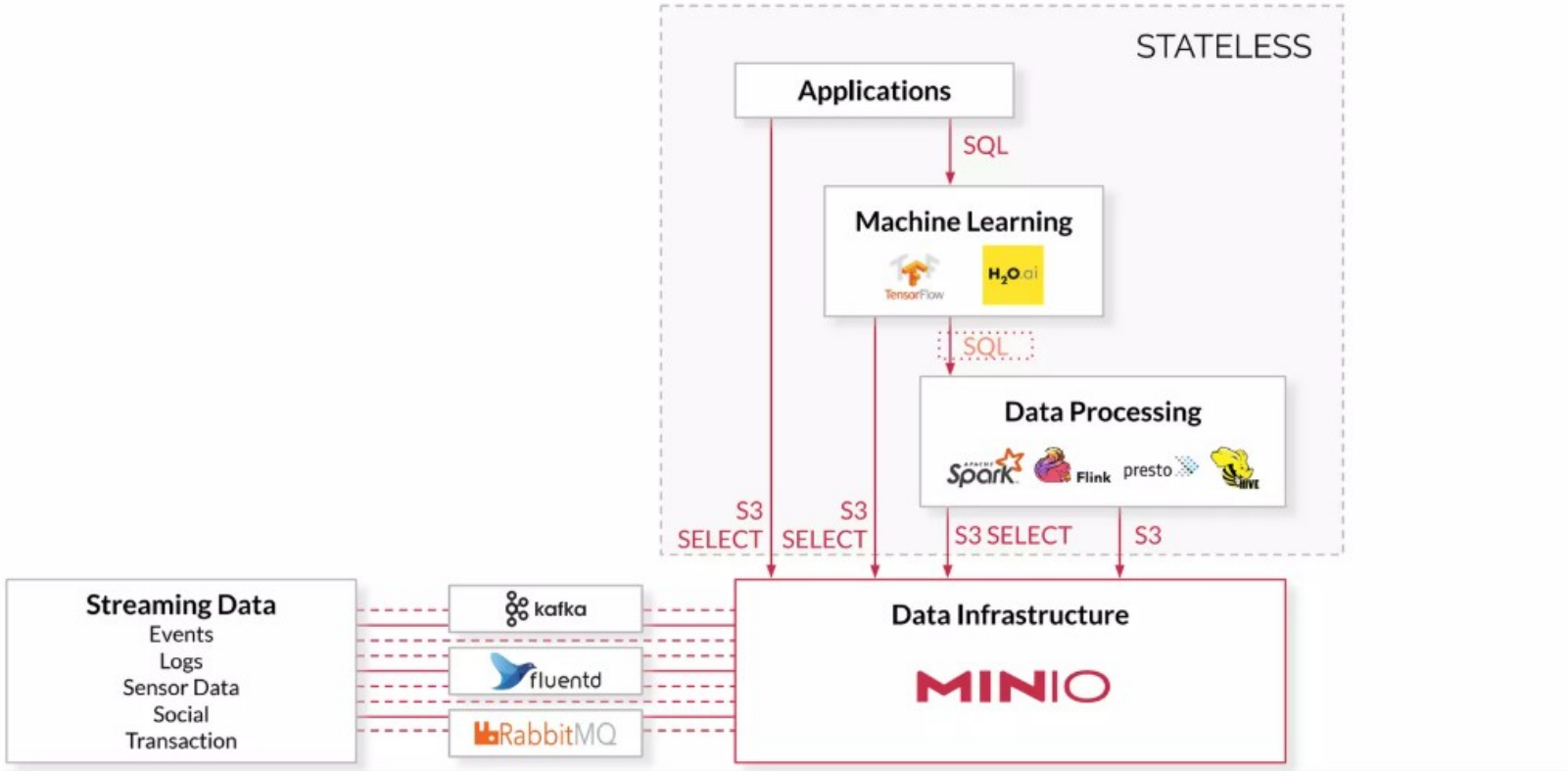
By deploying for performance, long term archival storage and disaster recovery are effectively free.

MINIO

MinIO Encryption - SSE-S3



At the Center of a Modern Architecture



MINIO

Performance

Scalability

Simplicity

MinIO Server

MinIO Client

MinIO Console

CONCLUSION

Big data and data storytelling are an **evolution** of concepts already used in the past from other scientists

Final goal of the Natural Disaster Management is to improve the **life of citizens**

"In God we trust.
All others must bring data."

- *Dr. W. Edwards Deming*

THANKS FOR YOUR TIME

ANY QUESTION?

ANTONIO FILOGRANA

ANTONIO.FILOGRANA@ENG.IT

LINKEDIN: @ANTONIOFILOGRANA