This Communication is part of a project that has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement Nº101069732

DATA MANAGEMENT PARADIGMS: DATA FABRIC & DATA MESH

IETF 117, NMRG
07/27/2023
Latest data management approaches

- **Data Fabric** — Defines a data infrastructure architecture
  - Unified view of heterogenous data served through secure and standard interface
  - Automated data integration mechanisms

- **Data Mesh** — Organizational approach
  1. Domain ownership — Domains teams are accountable for their data
  2. Data as a product — Apply product thinking. Data products must be discoverable, understandable, accessible and reusable.
  3. Self-service data infrastructure — Shared data infrastructure that helps data domains in the creation of data products
  4. Federated governance — Definition of policies for data access and privacy; central policies for data interoperability

*Data Fabric facilitates the creation of Data Products*
The almighty Knowledge Graph

• Concepts and their relations with other concepts are represented in a graph
  • Leverage ontologies and taxonomies to explicitly capture knowledge
  • Instances and categories/classes are combined in the graph

• Knowledge graph standards:
  • W3C Semantic Web → RDF, OWL
  • ETSI NGSI-LD (Property Graphs + Ontologies)

• Enabler of the data fabric paradigm
  • Data integration thanks to semantic layer
  • (Raw) Datasets are mapped to concepts
    • Also known as ontology mapping or semantic lifting

• Data mesh proposal
  • Data product as the conceptual representation of raw datasets
Use case: network interface management

- Definition of rules for mapping YANG to knowledge graph based on NGSI-LD

- Challenges:
  - Multiple YANG data models can refer to the same concepts
  - Even different elements of the same YANG data model can
  - Lack of standard ontology
    - Cover core concepts like "interface" or "input packets"
Takeaways and future challenges

- **Knowledge graphs** are key to unlocking AI
  - Facilitates data integration and provides machine-readable representation of knowledge

- YANG is widely used in network telemetry…
  - But not the only data modelling language (RFC 9232)

- Network industry should look at **ontologies**
  - NMRG can identify and analyze existing standard ontologies
  - Need to define standard ontologies aligned with work from other areas (e.g., OPSAWG, NETMOD)

- Cannot overlook the **data mesh wave**
  - Identify data domains in the network (telemetry planes, use cases)
  - Guidelines for building data products (quality, data sources)

- **Draft(s)** to discuss data management paradigms for network management
  - And do not forget data governance…
This Communication is part of a project that has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement Nº101069732

THANK YOU!

Ignacio Dominguez Martinez-Casanueva

✉️ ignacio.dominguezmartinez@telefonica.com
🌐 www.telefonica.com