Using automotive property graph-based data models in a knowledge graph

Aidan O Mahony
Principal Research Scientist
CTO Research Office

DELL Technologies

Agenda

- MOSAICrOWN overview
  - Use Case 1 (ICV) overview
  - Platform architecture
- Automotive data model
- FIWARE data model
- Dell EMC schema
- Data/Metadata split
- MOSAICrOWN policy language
- RDF-star and policy annotation

MOSAICrOWN

ORO

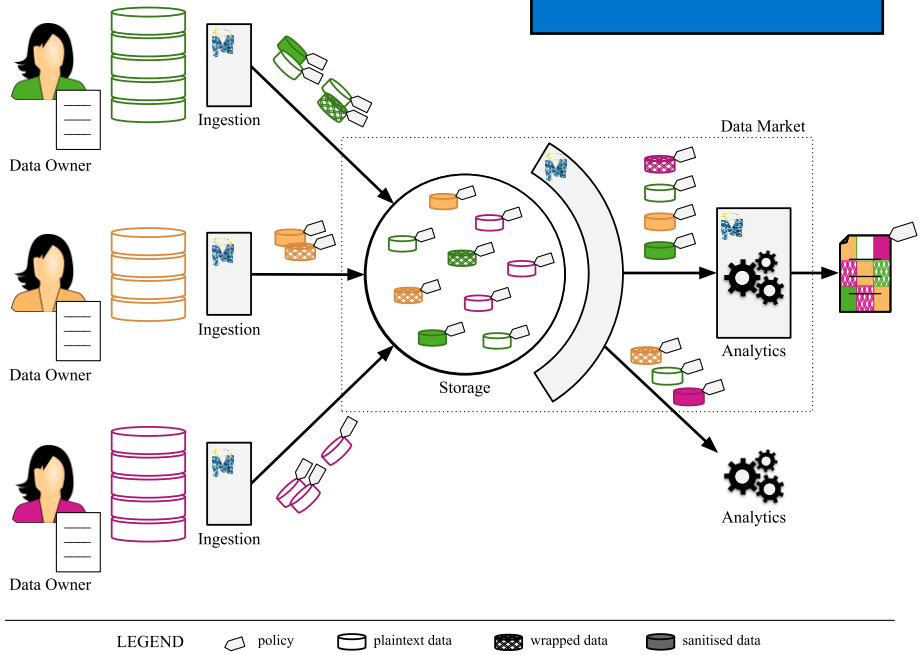
OCTO Research Office

Privacy Preserving Big Data Sharing and Analytics

Jan 2019 – Dec 2021

The Vision of MOSAICrOWN

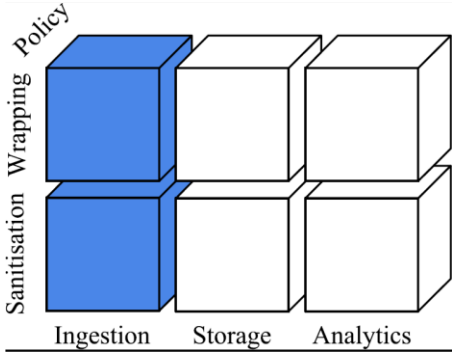
MOSAICrOWN aims to enable data sharing and collaborative analytics in multi-owner scenarios in a privacy-preserving way, ensuring proper protection of private/sensitive/confidential information. MOSAICrOWN will provide effective and deployable solutions allowing data owners to maintain control on the data sharing process, enabling selective and sanitized disclosure providing for efficient and scalable privacy-aware collaborative computations.



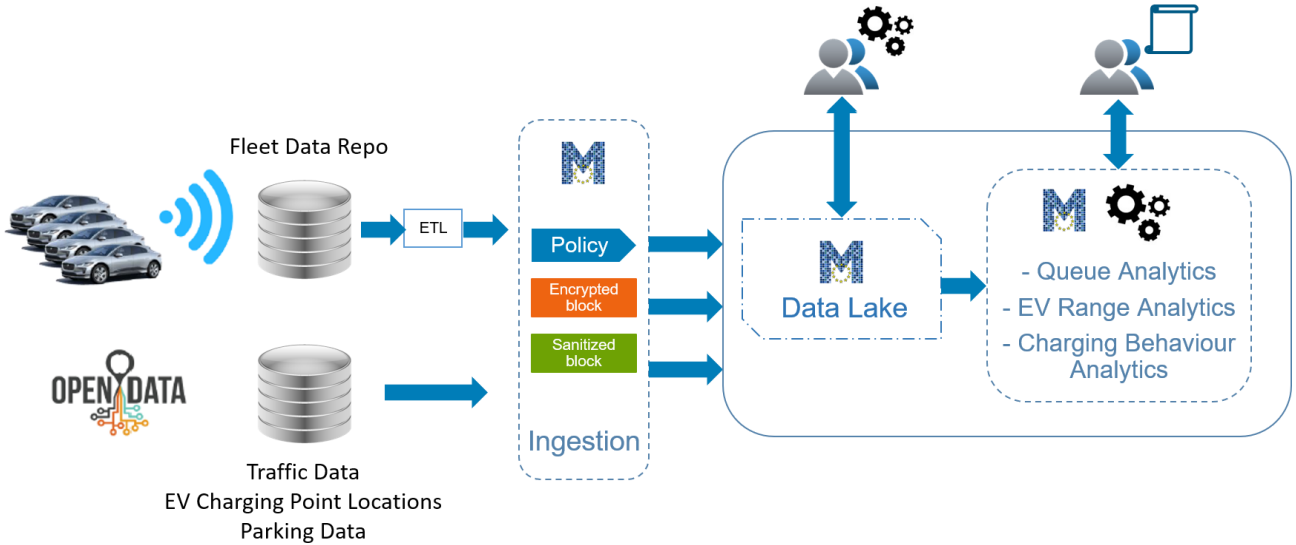
DELL Technologies

Use Case 1 Overview

ORO
OCTO Research Office



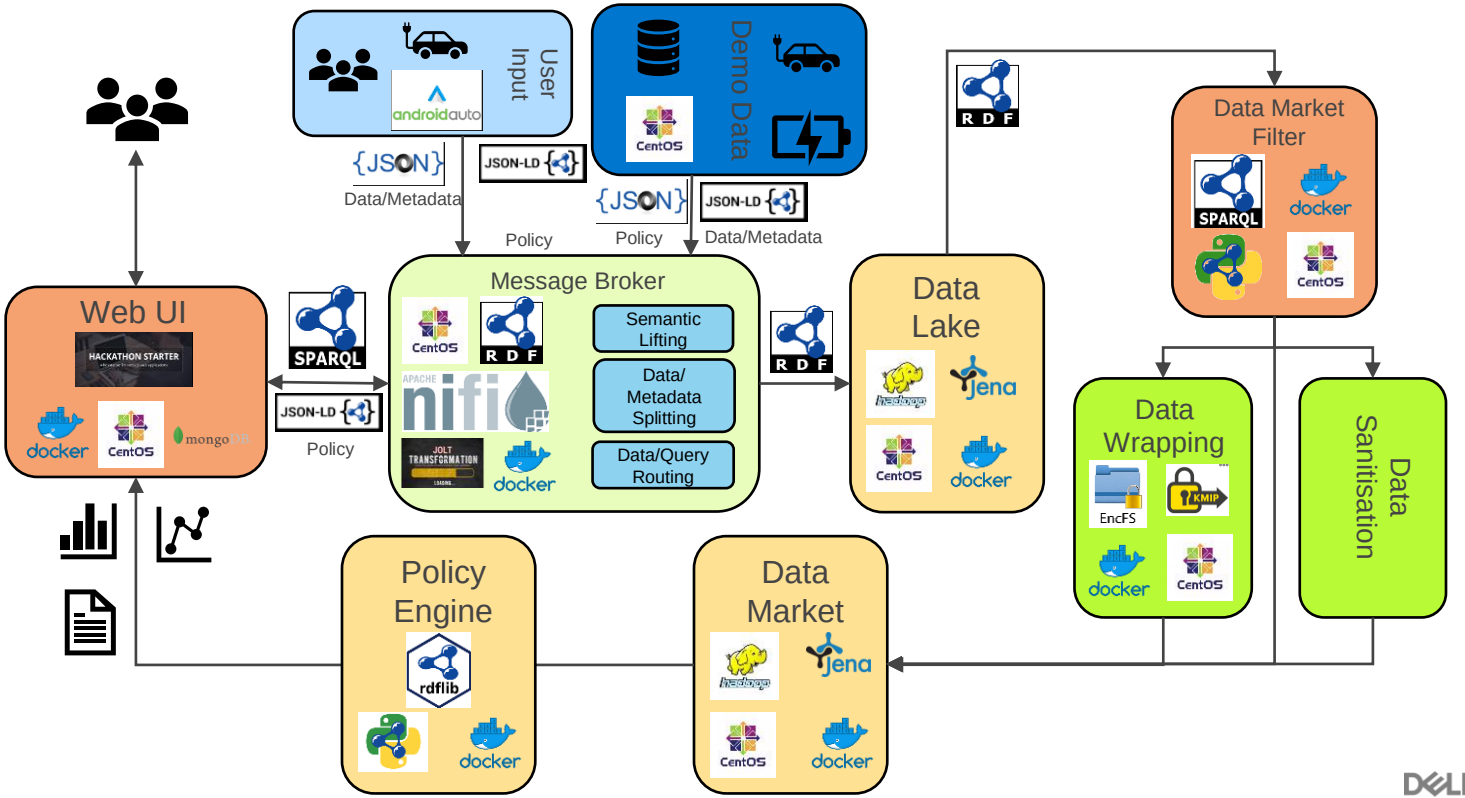
MOSAICrOWN Dimensions



DELL Technologies

MOSAICrOWN Platform Architecture

ORO
OCTO Research Office



DELL Technologies

Data Model

ORO

OCTO Research Office

VEHICLE INFORMATION

GET `/vehicle_information` Get Vehicle Information state

PARAMETERS

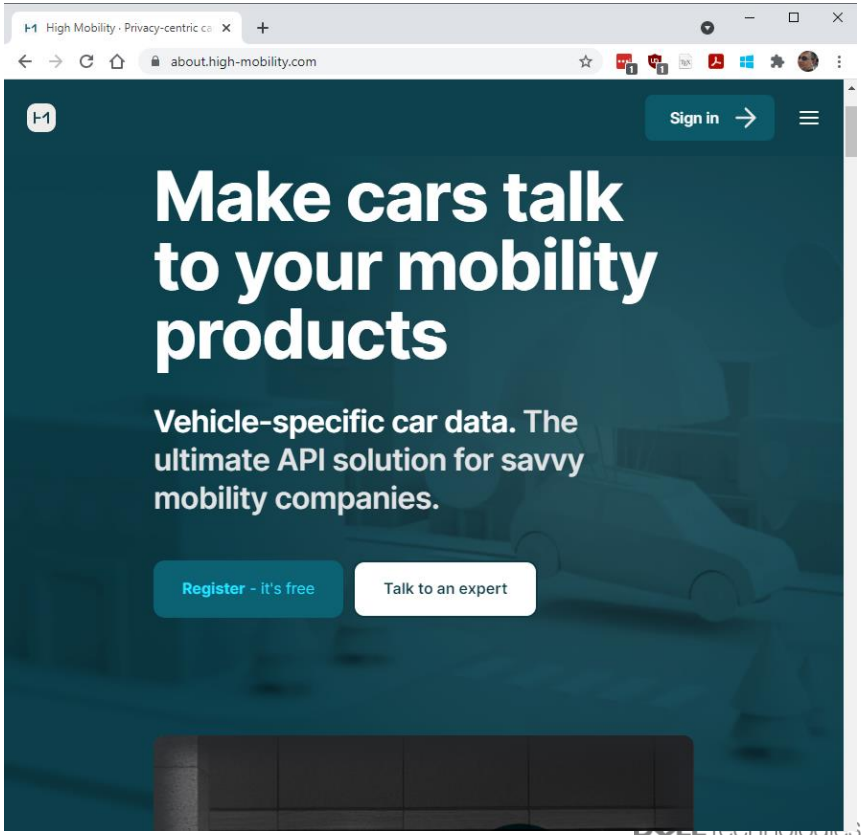
NAME	DESCRIPTION
Authorization * required	Authorization
string	
(header)	

RESPONSES

Response content type: application/json

Code	Description
200	Vehicle Information State Example Value | Model

```
{
  "brand": {
    "data": {
      "value": "bmw"
    }
  },
  "vin": {
    "data": {
```



FIWARE Vehicle Data Model

ORO
OCTO Research Office


```
{
  "id": "urn:ngsi-ld:Vehicle:vehicle:WasteManagement:1",
  "type": "Vehicle",
  "category": {
    "type": "Property",
    "value": [
      "municipalServices"
    ]
  },
  "vehicleType": {
    "type": "Property",
    "value": "lorry"
  },
  "name": {
    "type": "Property",
    "value": "C Recogida 1"
  },
  "vehiclePlateIdentifier": {
    "type": "Property",
    "value": "3456ABC"
  },
  "refVehicleModel": {
    "type": "Relationship",
    "object": "urn:ngsi-ld:VehicleModel:vehiclemodel:econic"
  },
  "location": {
    "type": "GeoProperty",
    "value": {
      "type": "Point",
      "coordinates": [
        -3.164485591715449,
        40.62785133667262
      ]
    },
    "observedAt": "2018-09-27T12:00:00Z"
  },
  "areaServed": {
    "type": "Property",
    "value": "Centro"
  },
  "serviceStatus": {
    "type": "Property",
    "value": "onRoute"
  },
  "cargoWeight": {
    "type": "Property",
    "value": 314
  },
  "speed": {
    "type": "Property",
    "value": 50,
    "observedAt": "2018-09-27T12:00:00Z"
  },
  "serviceProvided": {
    "type": "Property",
    "value": [
      "gargabeCollection",
      "wasteContainerCleaning"
    ]
  },
  "@context": [
    "https://smartdatamodels.org/context.jsonld",
    "https://uri.etsi.org/ngsi-ld/v1/ngsi-ld-core-context.jsonld"
  ]
}
```


DELL Technologies

Dell Schema

ORO
OCTO Research Office

- FIWARE data model resulted in blank nodes in graph
- Needed to store metadata in RDF graph
- Also needed to account for data/metadata split.

```
{
    "@context":
    {
      "icv": "http://dellemc.com:8080/icv/",
      "id": "@id",
      "type": "@type",
      "value": "http://dellemc.com:8080/icv/hasValue",
      "object": {
        "@id": "http://dellemc.com:8080/icv/hasObject",
        "@type":"@id"
      },
      "Property": "http://dellemc.com:8080/icv/Property",
      "Relationship": "http://dellemc.com:8080/icv/Relationship",
      "DateTime": "http://dellemc.com:8080/icv/DateTime",
      "Date": "http://dellemc.com:8080/icv/Date",
      "Time": "http://dellemc.com:8080/icv/Time",
      "createdAt": {
        "@id": "http://dellemc.com:8080/icv/createdAt",
        "@type": "DateTime"
      },
      "modifiedAt": {
        "@id": "http://dellemc.com:8080/icv/modifiedAt",
        "@type": "DateTime"
      },
      "observedAt": {
        "@id": "http://dellemc.com:8080/icv/observedAt",
        "@type": "DateTime"
      },
      "datasetId": {
        "@id": "http://dellemc.com:8080/icv/datasetId",
        "@type": "@id"
      },
      "instanceId": {
        "@id": "http://dellemc.com:8080/icv/instanceId",
        "@type": "@id"
      },
```

DELL Technologies

Metadata/Data Split

ORO
OCTO Research Office


```
{
    "id": "urn:ngsi-ld:Vehicle:7CKRXPHSZE",
    "@context": [
        {
            "id": "@id",
            "type": "@type",
            "dataCreated": {
                "@id": "http://purl.org/dc/terms/modified",
                "@type": "http://www.w3.org/2001/XMLSchema#dateTime"
            }
        },
        "http://dellemc.com:8080/icv/schema.json"
    ],
    "type": "Vehicle",
    "category": "Private",
    "vin": "7CKRXPHSZE",
    "privacyPolicy": "http://dellemc.com/policy/leastPrivatePolicies",
    "name": "GOHGTXXM",
    "modelName": "Mercedes-Benz EQC",
    "modelYear": 2013,
    "colourName": "orange",
    "numberOfDoors": 2,
    "numberOfSeats": 4,
    "gearbox": "Automatic",
    "displayUnit": "Km",
    "driverSeatLocation": "Left",
    "dataset": {
        "id": "http://172.17.0.2:50070/webhdfs/v1/data/c9a2ba28-e1da-4f42-8581-ffefa847436e",
        "dateCreated": "2021/08/31 12:41:45.854"
    }
}
```


DELL Technologies

MOSAICrOWN Policy Language

ORO
OCTO Research Office

```
"@context": [
  "http://www.w3.org/ns/odrl.jsonld",
  "http://localhost:8000/ns/mosaicrown/namespace.jsonld"
],
"@type": "Set",
"uid": "https://dellemc.com/policy/leastPrivatePolicies",
"permission": [
  {
    "uid": "https://dellemc.com/policy/
       leastPrivatePolicies_perm",
    "assignee": "https://dellemc.com/user/fleetmanager",
    "target": ["https://dellemc.com/icv/licensePlate",
               "https://dellemc.com/icv/vin",
               "https://dellemc.com/icv/name"
              ],
     "action": ["odrl:read","odrl:use","odrl:write","odrl
        :sell","odrl:sellReport"],
     "purpose": ["statistical", "marketing"]
  }
]
```

- The automotive scenario includes three main parties
  - Data owners (drivers) ingesting their data into the data market
  - Consumers accessing data in the data market
  - The data market provider offering storage and computation services to data owners and consumers.
- Subjects, transformations, datasets, metadata, operations, and purposes were identified as basic elements of the policy model that also needs to be captured by the policy language.

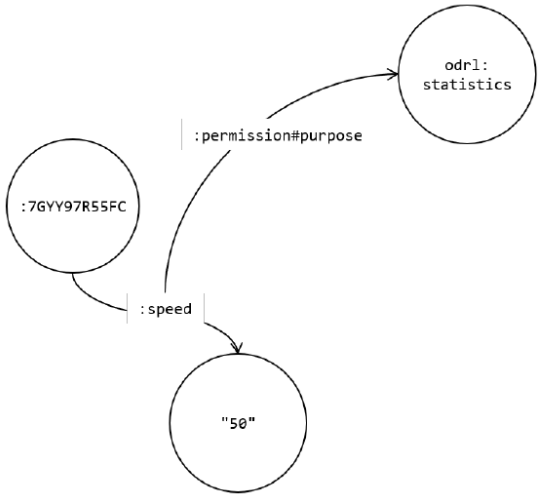
DELL Technologies

RDF-Star and Policy Annotation

- RDF-Star annotation is added to indicate that at the speed of 50 kilometers per hour the driver of the vehicle allows the speed to be used for odrl:statistical purposes, but not odrl:marketing.
- This indicates that the data owner allows for their speed to be aggregated for analysis, but not for marketing.

```
:7GYY97R55FC :speed "50"^^xs:integer .

<<:7GYY97R55FC :speed "50"^^xs:integer>> :permission#purpose odrl:statistics .
```



Conclusion

- Existing data models required extending to best model an ICV
- Data/Metadata splitting allowed for us to search data (textual/non-textual) via graph like queries
- RDF-star could allow for the annotation of policy to metadata for policy enforcement


DELL Technologies