OGC Testbed-14
Machine Learning for Earth Observation (ML/EO)

Final Demonstration meeting
ESA/ESRIN, January 2019

Kevin Heffner, Tom Landry

CRIM

OGC®
Making location count.

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

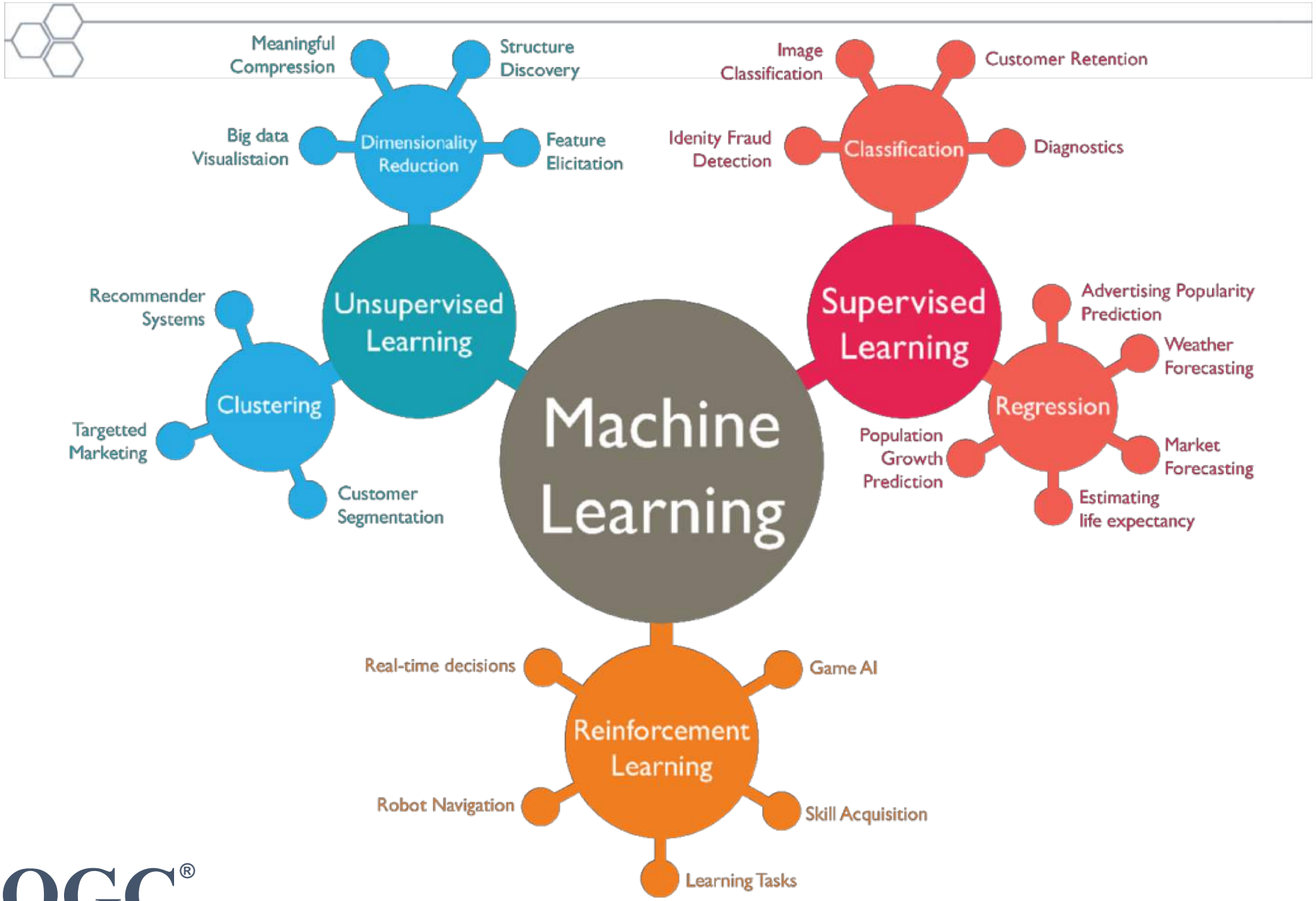
TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

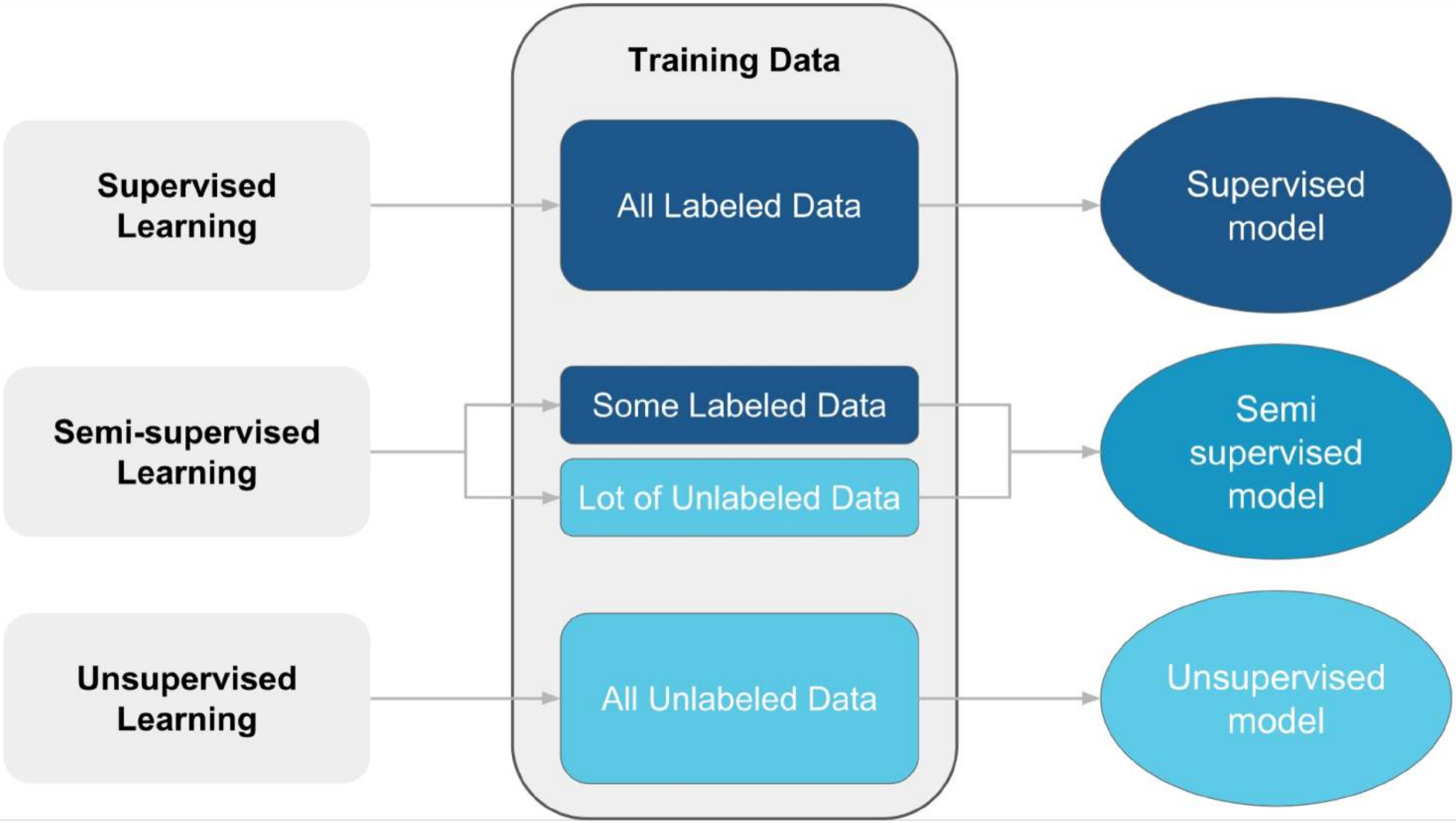
TAXONOMY OF MACHINE LEARNING



OGC®

SOURCE: DHL, ARTIFICIAL INTELLIGENCE IN LOGISTICS, 2018 (PDF, 45 PP., NO OPT-IN).

Supervised, Semi-Supervised & Unsupervised ML



Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

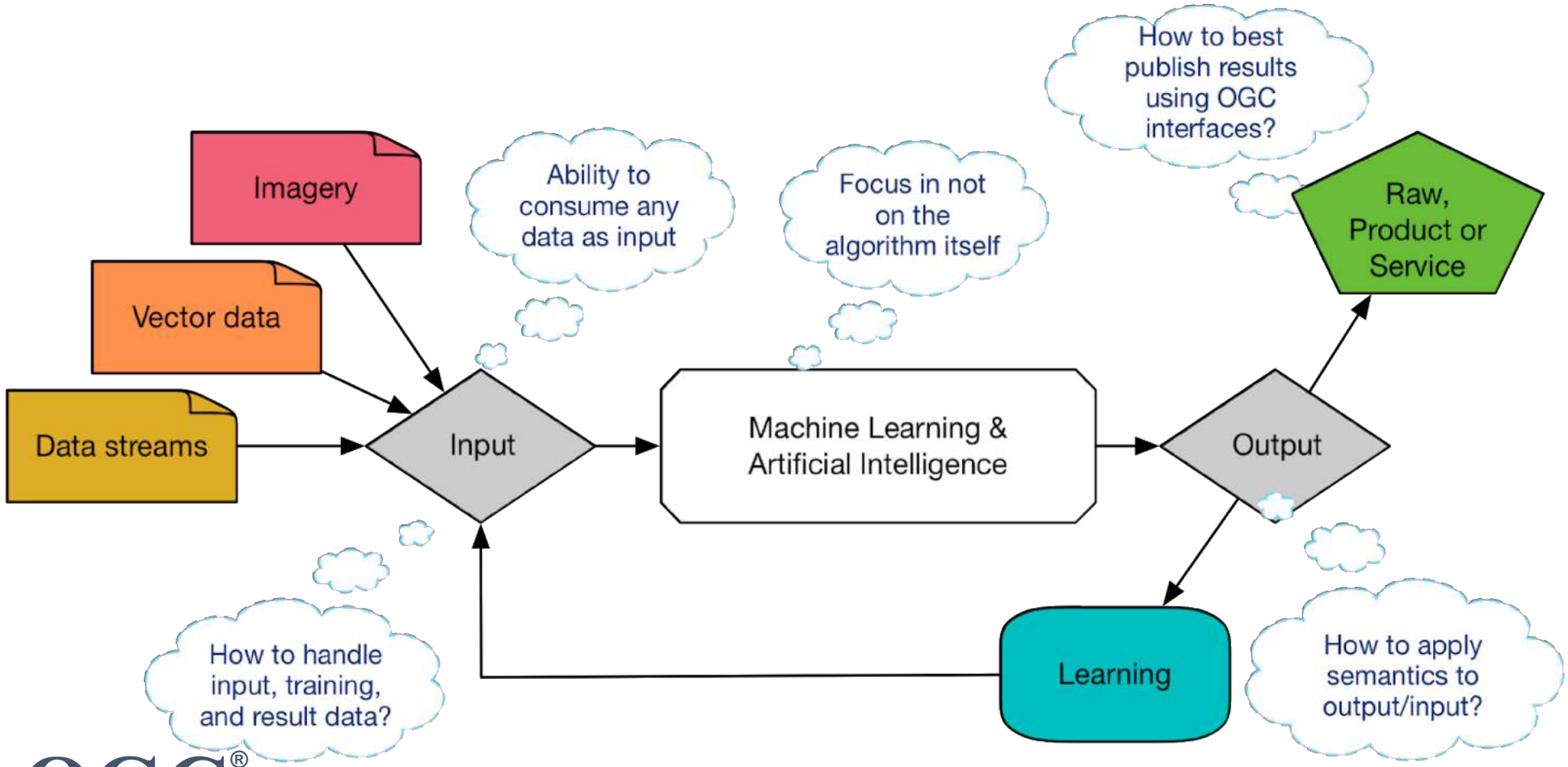
Conclusions

OGC®

TB-14 ML/EO Task: Motivation

Modeling, Portrayal, and Quality of Service (MoPoQ) Thread

Question: What is best approach to support ML and AI using OWS?



OGC®

TB-14 ML/EO Task: Participants and sponsors

International Effort: NA: CAN, USA EUR: GBR, GER SEA:TWN

Sponsors

SatCen

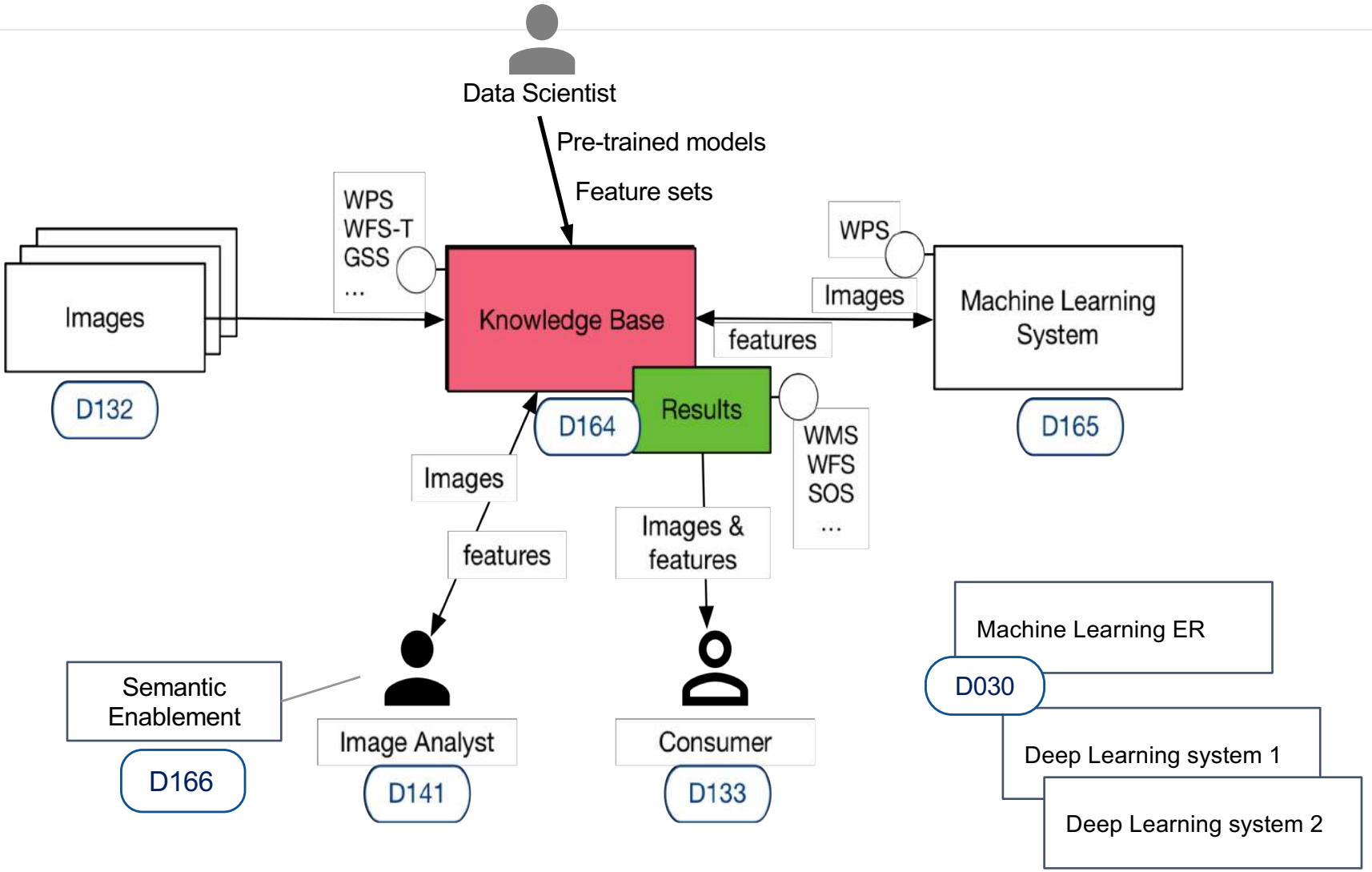
[dstl]

Participants

Tom Landry, *CRIM*
Martin Sotir, *CRIM*
Cullen Rombach, *Image Matters*
Cameron Brown, *Envitia*
Neil Kirk, *Envitia*
Benjamin Pross, *52 North*
Chih-Wei Kuan, *Feng Chia University*

OGC®

TB-14 ML/EO Task: Architecture



OGC®

TB-14 ML/EO Task: Deliverables

D030: Machine Learning ER

D132: Image repository and feature store

D133: Client to Knowledge Base (KB)

D141: Machine Learning (ML) validation client

D164: ML Knowledge Base

D165: ML System (MLS)

D166: Semantic Enablement of ML *(not in CFP)*

OGC®

TB-14 ML/EO Task: Summary of Use cases

Image Analyst

- ***Annotates*** images, i.e. links new features to classes
- ***Trains*** new model using annotated images
- ***Validates*** model output; re-trains model, if necessary

Consumer

- ***Executes*** MLS on geospatial data (with trained, validated model)
- ***Consults*** MLS outputs: images and features

Data Scientist *(Not in original CFP)*

- ***Provides*** pre-trained models from their own training feature sets
- ***Packages*** reference MLS and data samples

OGC®

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept (POC)

TB-14 ML/EO Deep Learning Applications

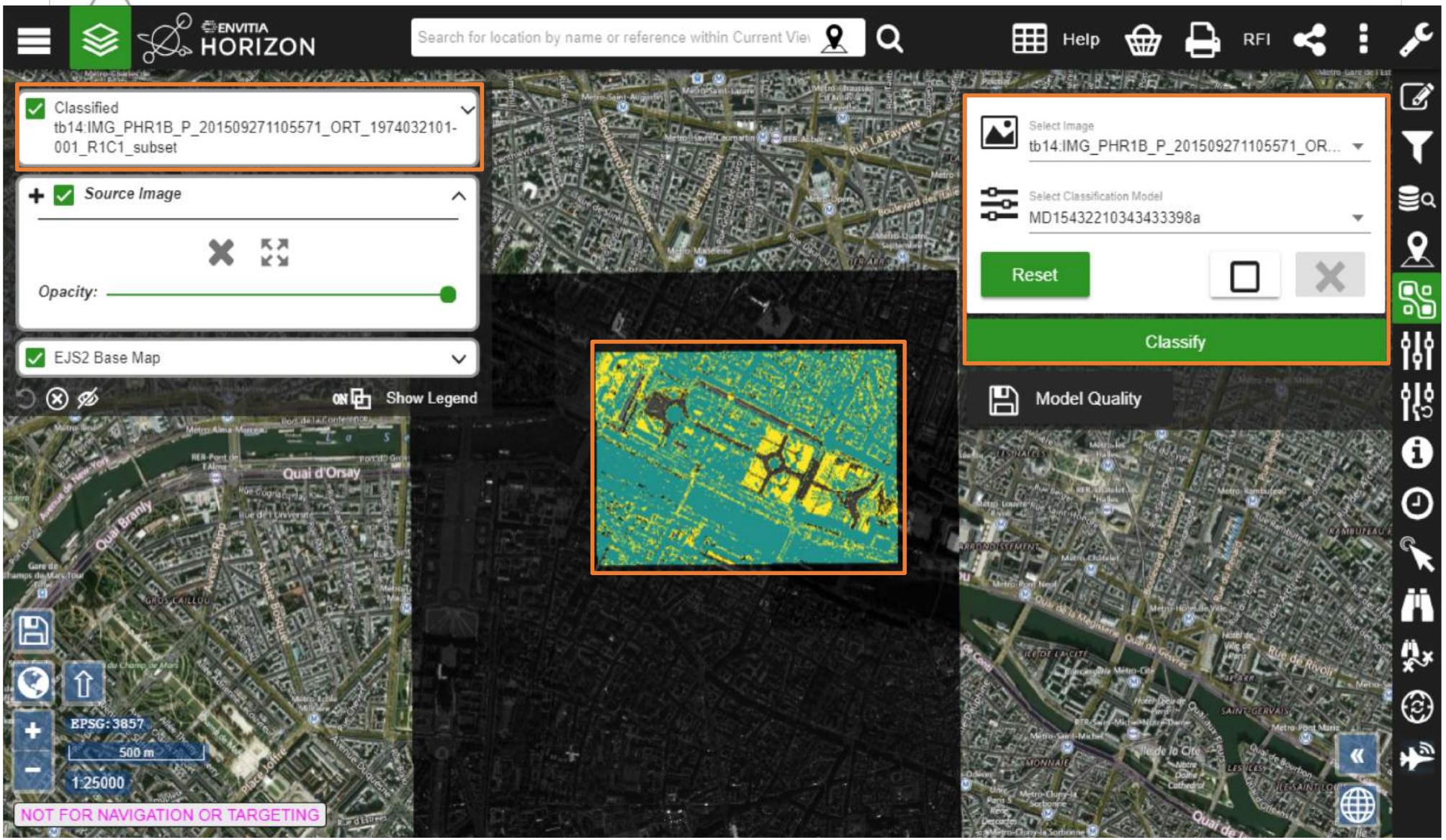
TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

TB-14 ML/EO POC: Classification



TB-14 ML/EO POC: Controlled Vocabulary Manager

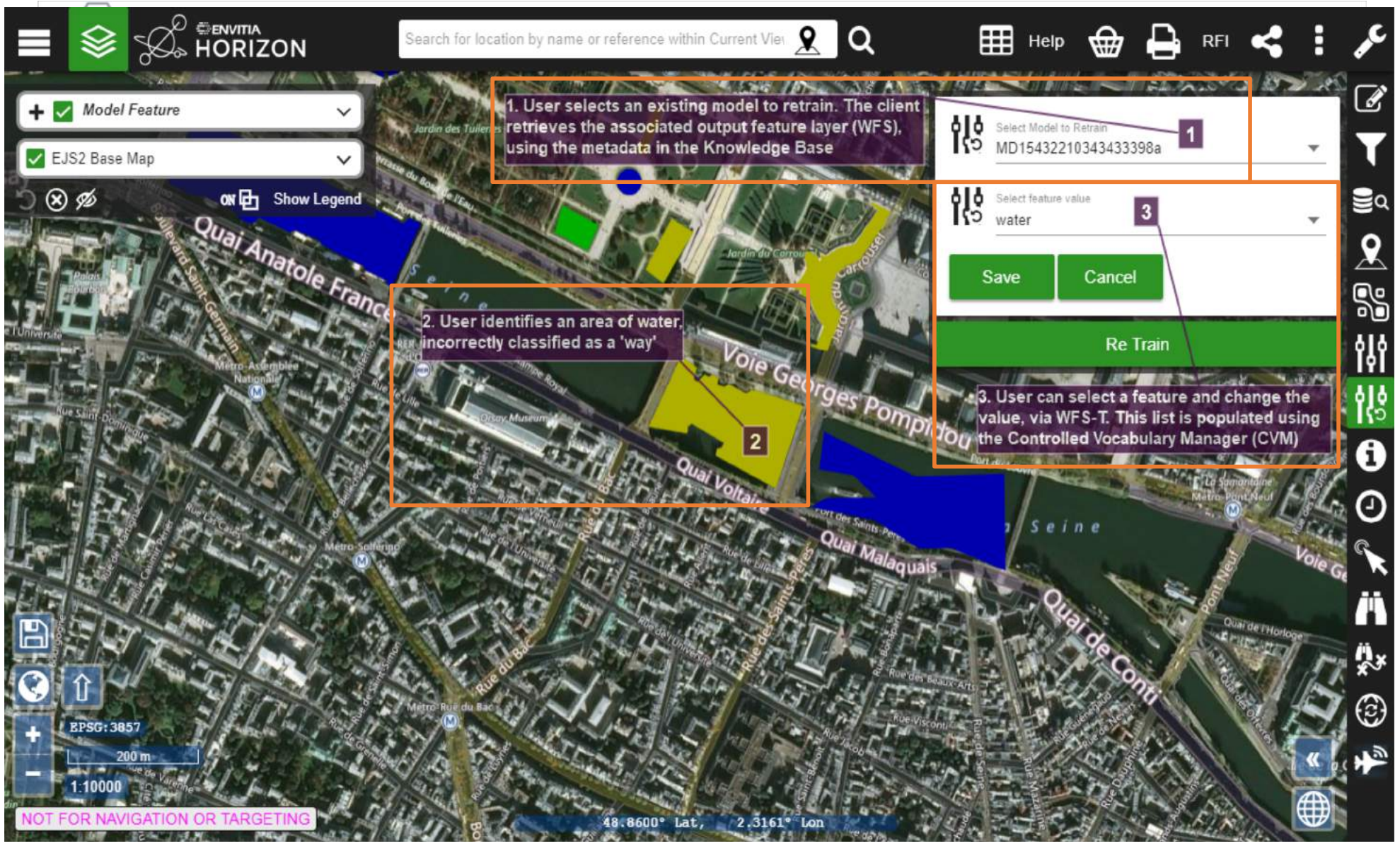
D166 Semantic Enablement

- Combination of NIEM and NSG
- Land cover classes vs targets


```
{
  "results": [
      {
          "id": "4d2fda7b59b1092b2b028f8ddb3d2e01",
          "uri":
"http://api.nsgreg.nga.mil/codelist/ConveyanceType/automobile",
          "type": [
              "individual",
              "concept"
          ],
          "primary": false,
          "prefix": "codelist",
          "version": "1.0",
          "label": "automobile",
          "description": "Definition: A motor vehicle generally with four
wheels that carries a small number of passengers.  Description: [None
Specified]",
          "namespace": "http://api.nsgreg.nga.mil/codelist/ConveyanceType/",
          "localName": "automobile",
          "qname": "codelist:automobile",
          "category": [
              "Individual",
              "additionalProperties"
          ]
      }
  ]
}
```


OGC®

TB-14 ML/EO POC: Validation and Re-training



OGC®

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

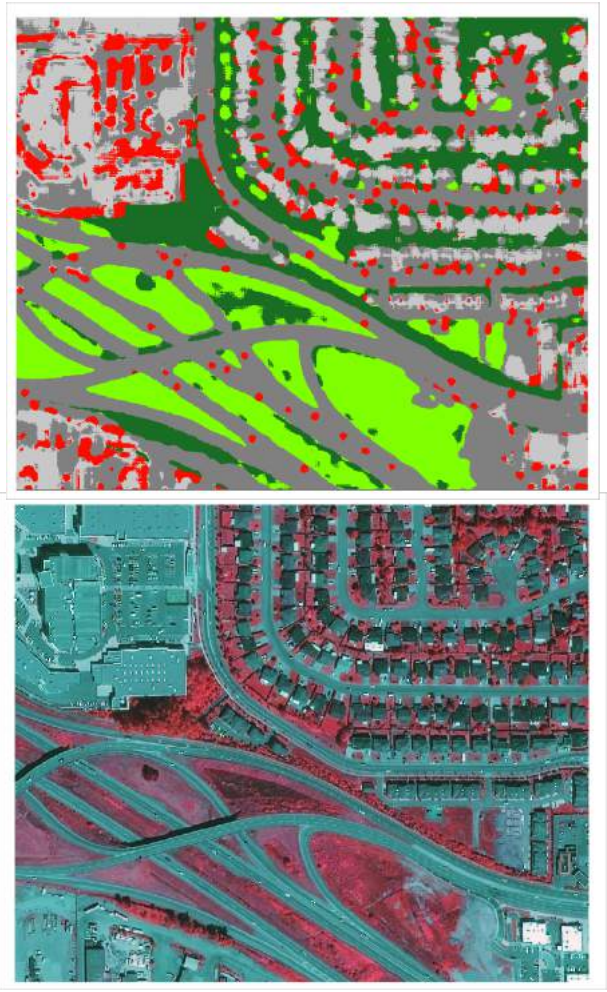
TB-15 ML Thread

Conclusions

OGC®

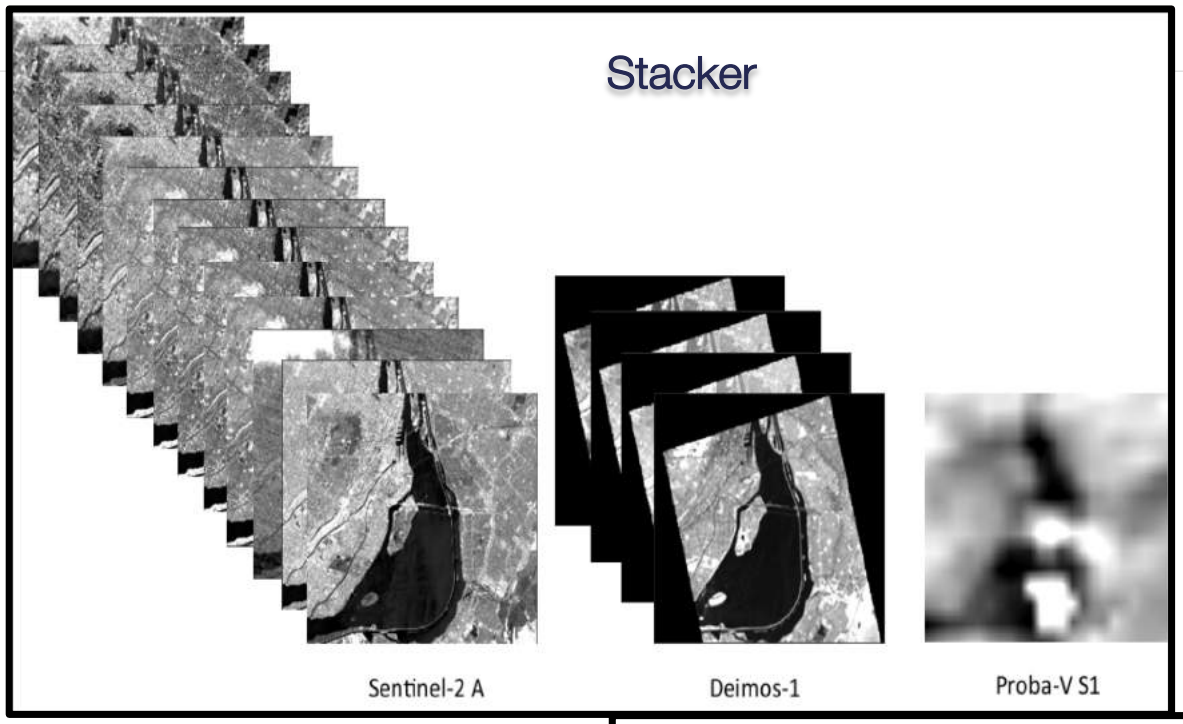
TB-14 ML/EO Task DL Apps: Car Detector

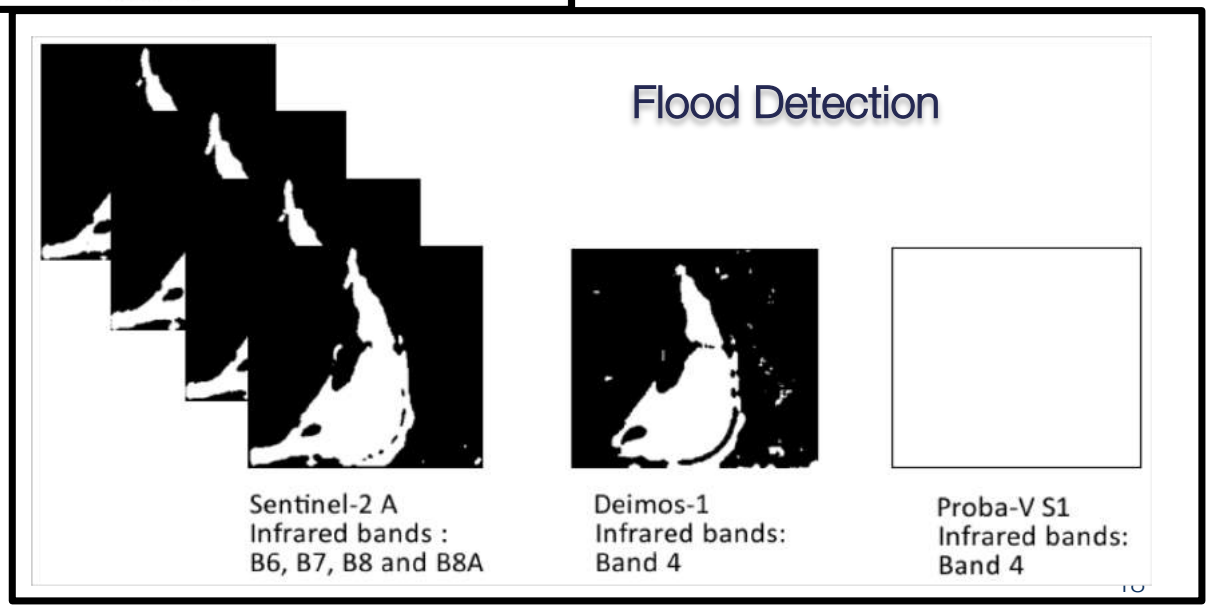
- **5 classes**
  - **cars**, trees, roads, houses, buildings
  - probability for each class at each px
- **Small training set in Vancouver**
  - about 1000 points per class
- **MLS as standalone Docker**
  - Based on PyTorch
  - Easy to use in CWL workflows
  - Still need vectorization of detections



OGC®

TB-14 ML/EO Task DL Apps: Flood Detection

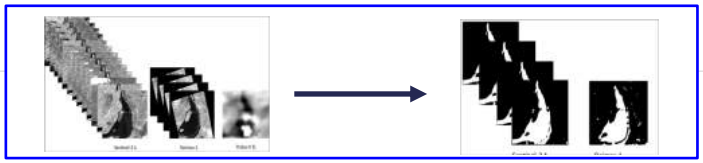


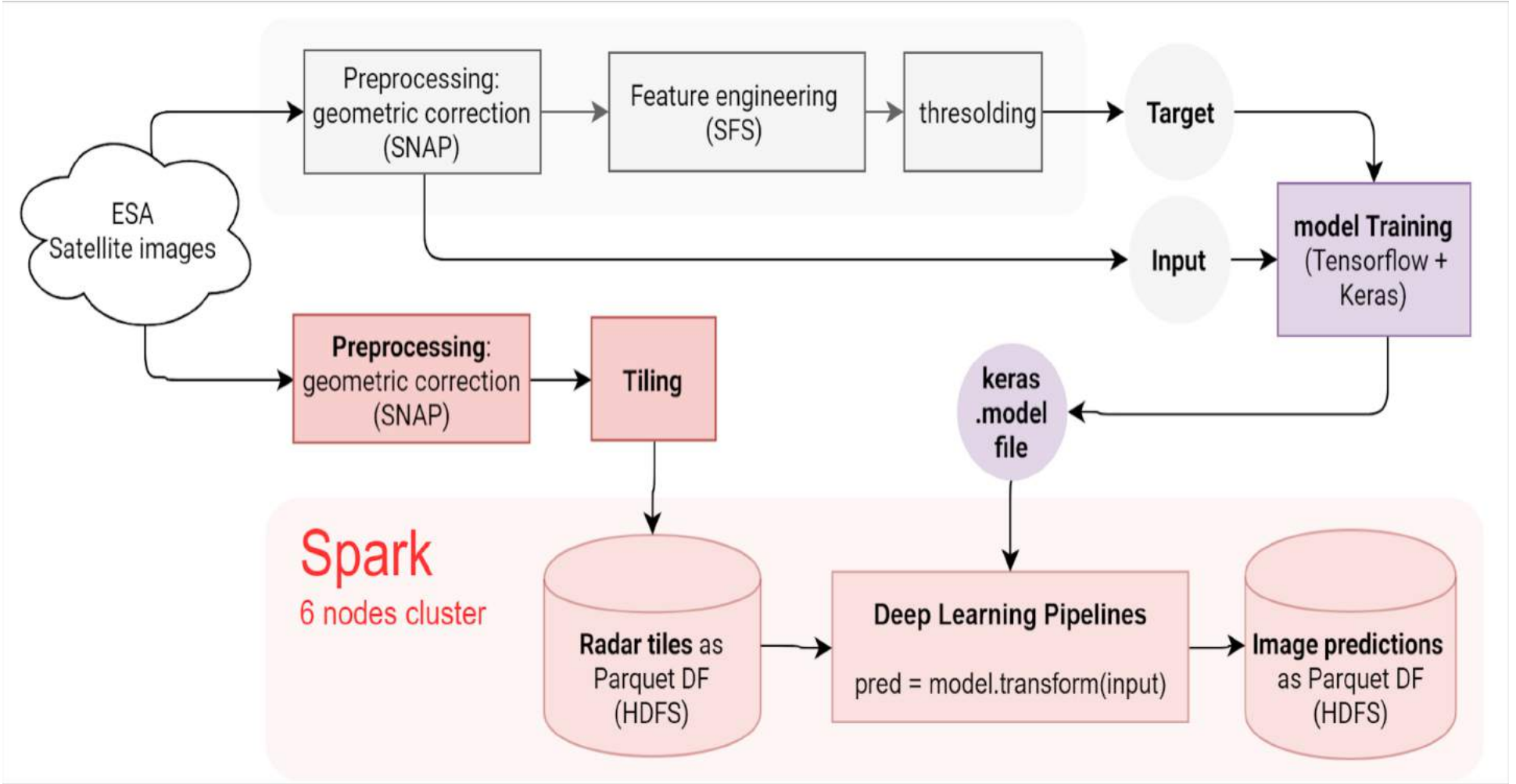


OGC®

TB-14 ML/EO Task DL Apps: Flood Detection

WORKFLOW DESCRIPTION





Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

TB-14 ML/EO: Findings & Recommendations

- **Opportunities**
  - Relatively few standards found in AI and ML
  - ML applications often built as workflows (pipelines, graphs)
  - Models should be discoverable, like applications and data
- **Recommended Standards, Specifications & Extensions**
  - WPS-T 2.0 REST/JSON for specifying processes
  - WFS 3.0 for management of features
  - CSW for Knowledge Base and Controlled Vocabulary Manager
  - CSW-ISO 19115 and CSW-ebRIM application profiles
- **Suggested Future Work**
  - Tasks
    - Temporal, semantic and computational enablements
    - Use of CSW and NextGen services
  - Deliverables
    - Application packages and workflows, Point clouds

OGC®

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

TB-15: ML Thread

Scenarios

- Cloud-free mosaicing, land cover classification
- Forest supply management
- Lake-river differentiation
- Linked data harvest
- Web service discovery

Data: Optical Imagery, LiDAR, Annotated Data, Road & Hydro Networks, DEM, Places, Ontologies, etc.

Standards: WPS, WFS, CSW, WCS, WMS

OGC®

Agenda

Machine Learning Taxonomy

TB-14 ML/EO Task Overview

TB-14 ML/EO Proof of Concept

TB-14 ML/EO Deep Learning Applications

TB-14 ML/EO Findings & Recommendations

TB-15 ML Thread

Conclusions

OGC®

Conclusions

- **TB-14 ML/OE was a challenging, complex task !**
  - Many topics, concerns & deliverables
  - Difficult to converge on design early on

- **Achievements**
  - The disaster use case helped convey the potential value of the work
  - Successfully developed a functional Proof-Of-Concept
  - Good synergy from a diverse group of participants

- **Bottom Line**
  - TB-14 ML provides a good foundation for future work
  - ML now is a first class thread in TB-15
  - Good value delivered to sponsors

OGC®

Thank you for your attention !

OGC®

BACKUP SLIDES

OGC®

Concept: POC operations

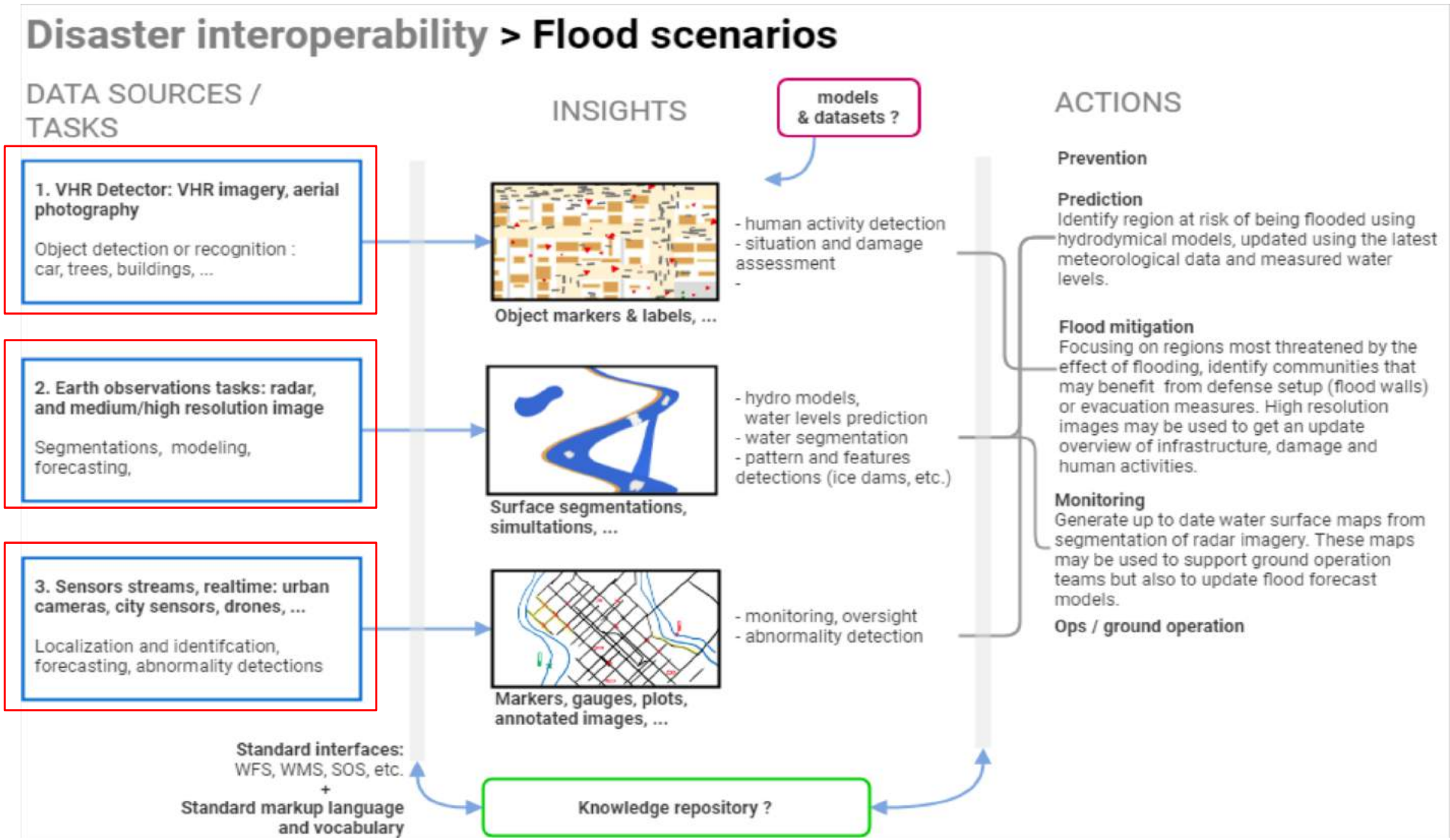
- WCS GetCapabilities, GetCoverage, DescribeCoverage
- WFS GetCapabilities, GetFeature, Update/Delete/Insert
- WMS GetMap
- WPS Execute
  - TrainML, RetrainML, ExecuteML
- ML Knowledge Base (KB). OpenAPI to Get/Store:
  - Models, Images, Features, Metadata
- Controlled Vocabulary Manager (CVM). Search by:
  - Id, uri, namespace, prefix, facet

ML/EO Task: ER by the numbers

- International effort
  - 5 participating institutions, 3 sponsors. Total 6 countries.
- 230 commits on GitHub, 9 contributors, 9 versions
- 79 pages in 9 sections
- 6 normative documents referenced
- 22 figures, 32 citations
- Recommended future work
  - 5 tasks
  - 9 deliverables
  - 3 ER
- Covers 6 deliverables
  - Plus 2 supporting Deep Learning contributions

OGC®

Concept: disaster scenario



OGC®

Metadata for DL system

```
check_point = {}
check_point['state_dict'] = model.state_dict()  # weights
check_point['model_config'] = self.model_config # informations about the
architectures
check_point['config'] = config # Training configuration
check_point['epoch'] = current_epoch # Training epoch
check_point['iteration'] = current_iteration # Training iteration
check_point['accuracy'] = current_accuracy # Accuracy
check_point['optimizer_state_dict'] = self.optimizer.state_dict()  #
optimizer params
check_point['criterion_config'] = self.criterion_config # cost function
check_point['optimizer_config'] = self.optimizer_config  # optimizer def
```

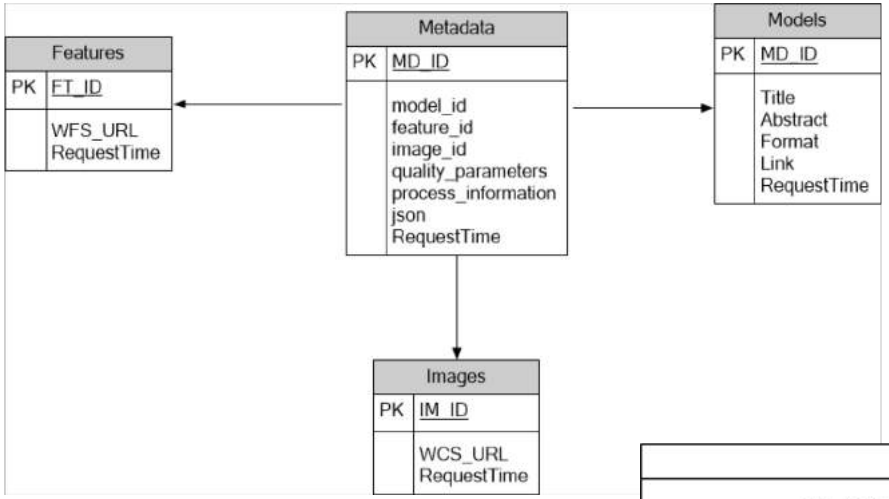
OGC®

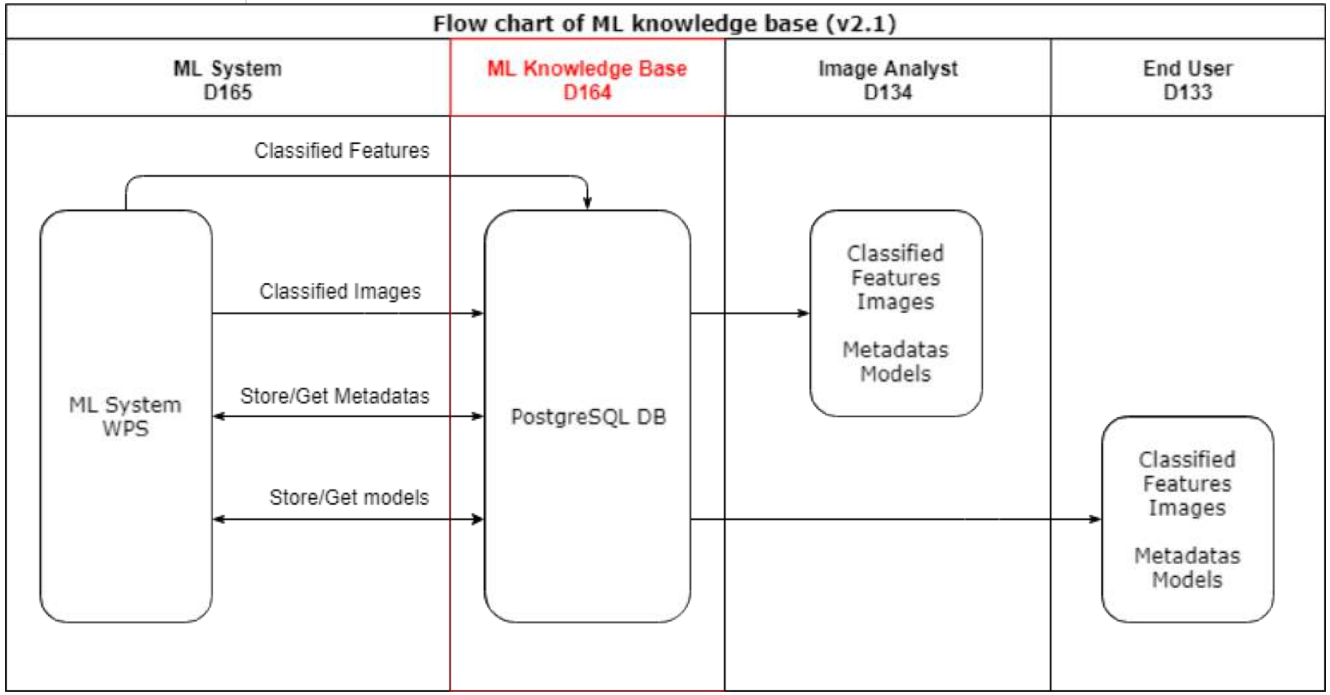
GML application schema

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:ogc-tb14-ml=
"http://schemas.opengis.net/testbed14/machine-learning" xmlns:gml=
"http://www.opengis.net/gml/3.2" targetNamespace=
"http://schemas.opengis.net/testbed14/machine-learning" elementFormDefault=
"qualified">
    <xs:import namespace="http://www.opengis.net/gml/3.2" schemaLocation=
"http://schemas.opengis.net/gml/3.2.1/gml.xsd"/>

    <xs:element name="ClassifiedFeature" type="ogc-tb14-
ml:ClassifiedFeatureType" substitutionGroup="gml:AbstractFeature">
        <xs:annotation>
            <xs:documentation>A feature representing a real-world object that
has been classified by a Machine Learning image classification
algorithm.</xs:documentation>
        </xs:annotation>
    </xs:element>

    <xs:complexType name="ClassifiedFeatureType">
        <xs:complexContent>
            <xs:extension base="gml:AbstractFeatureType">
                <xs:sequence>
                    <!-- The feature identifier in the Feature Store. -->
                    <xs:element name="Id" type="xs:long"/>
                    <!-- The image classfication category, ideally from a
controlled vocabulary such as the NAS or NIEM. -->
                    <!-- Implemented here as xs:string for simplicity. Could
use dedicated GML property for codelists instead. -->
                    <!-- Similar to properties tb14:category and tb14:term on
tb14:labels feature type in current Feature Store WFS. -->
                    <xs:element name="classificationType" type="xs:string"/>
                    <!-- The geometric footprint of the feature. -->
                    <xs:element name="the_geom" type=
"gml:GeometryPropertyType"/>
                </xs:sequence>
            </xs:extension>
        </xs:complexContent>
    </xs:complexType>
</xs:schema>
```

OGC®

KB database and sequences





OGC®

Findings: other OGC standards, extensions

- Application Packaging best practices for systems
- Workflow best practices
- WPS, WMS and WFS experiments
  - model interop and transparency
- WMTS and Vector Tiles as inputs to ML systems
- Open Modelling Interface (OpenMI)

OGC®