Government of Canada | Gouvernement du Canada

WIGOS Metadata: How to work with XSD and XML

Tom Kralidis
Senior Systems Scientist
Data Management
Meteorological Service of Canada

WMO Expert Team on World Data Centres
04 October 2017

Canada

WIGOS Metadata: How to work with XSD and XML

- XML
- XSD
- UML
- WMDR Development
- WOUDC implementation
  - WFS
  - template

eXtensible Markup Language (XML)

- Circa 1998 (1.0)
  - profile of SGML
- Self Describing
- Structured
  - tags
  - elements
  - attributes
- Nested
- Human Readable
- Portable, non-proprietary
- Interoperable
- Broad, robust support

XML Example

```
<?xml version="1.0"?>
<!-- WMO ET-WDC Meeting 2017-10-02 -->
<Meeting codeName="wmo-et-wdc">
        <Date>
                <start>2017-10-01</start>
                <end>2017-10-04</end>
        </Date>
        <Location>
                <name>Kjeller</name>
                <adminstrativeArea>Skedsmo</adminstrativeArea>
                <country>Norway</country>
        </Location>
        <OnlineResource xmlns:xlink="http://www.w3.org/1999/xlink"
xlink:href="https://sites.google.com/site/wmoetwdc/2017/2017-10-02-
et-wdc-meeting"/>
</Meeting>
```

XML Schema Definition (XSD)

- Circa 2001
  - XML-based followup to DTD
- Validation
- Data types
- Cardinality
- Custom Types
- Custom Vocabulary
- Namespaces!
- Reusable

XSD Example

```
<?xml version="1.0"?>
<!-- WMO ET-WDC Meeting 2017-10-02 -->
<schema xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:foo="http://example.org/foo">
<complexType name="MeetingType">
                    <sequence>
                              <element name="Date" type="foo:DateType"/>
                              <element name="Location" type="foo:LocationType"/>
                              <element name="OnlineResource"
type="OnlineResourceType"/>
                    </sequence>
                    <attribute name="codeName" default="cool meeting"/>
          </complexType>
          <complexType name="DateType">
                    <sequence>
                              <element name="start" type="date"/>
                              <element name="end" type="date"/>
                    </sequence>
          </complexType>
</schema>
```

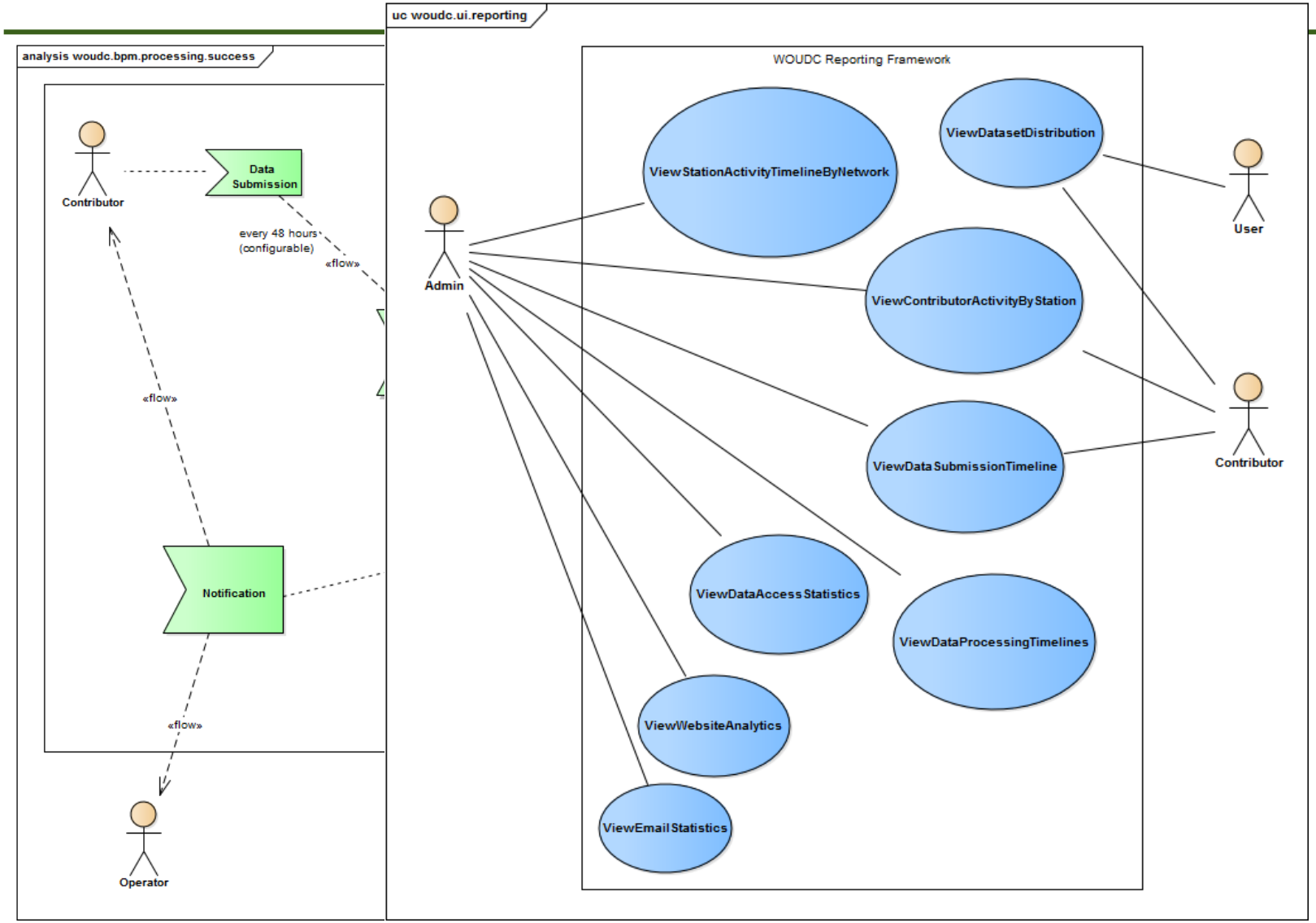
XML Example

```
<foo:Meeting xmlns:foo="http://example.org/foo"
xmlns:gml="http://www.opengis.net/gml/3.2" codeName="wmoetwdc"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://example.org/foo
http://example.org/fooschema.xsd">
        <foo:Date>2017-10-02</foo:Date>
        <foo:Location>
                <gml:Point srsName="EPSG:4326">
                        <gml:pos>11.047504 59.975000</gml:pos>
                </gml:Point>
        </foo:Location>
        <foo:OnlineResource>https://sites.google.com/site/wmoetwd
c/2017/2017-10-02-et-wdc-meeting</foo:OnlineResource>
</foo:Meeting>
```

Unified Modelling Language

- Circa mid-1990s
- Design
- Conceptual
- Implementation Agnostic
- Various capabilities for Systems Design
  - Structure
    - Classes
  - Behaviour
    - Use case
  - Interaction
    - Sequence
- Can generate XSDs, Java/Python classes, DB schemas, etc.

UML Examples

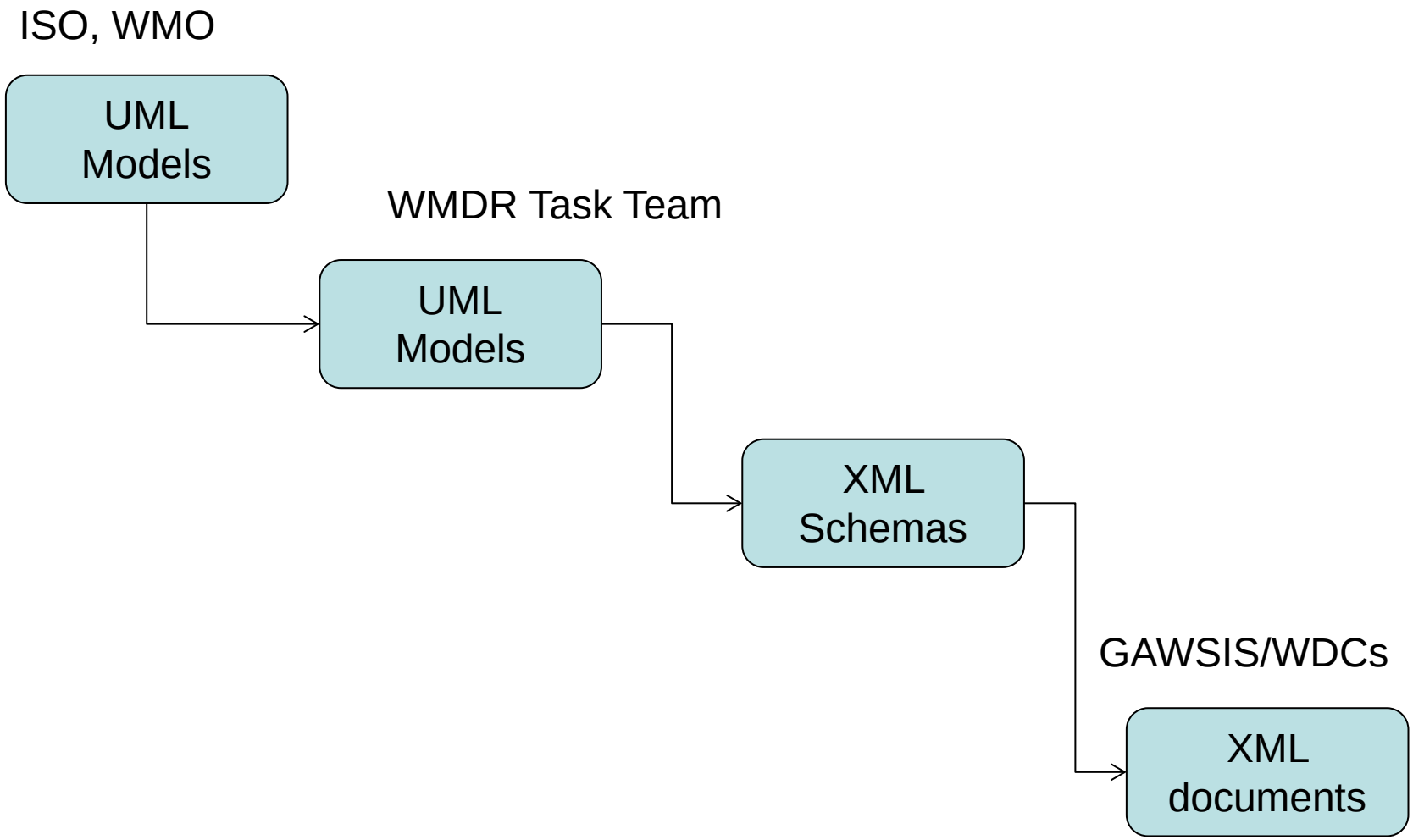


UML, XSD and XML

- UML: design/modelling
- XSD: physical implementation of model as schema
- XML: instance of XSD

- UML not required to create XSD
- XSD not required to create XML

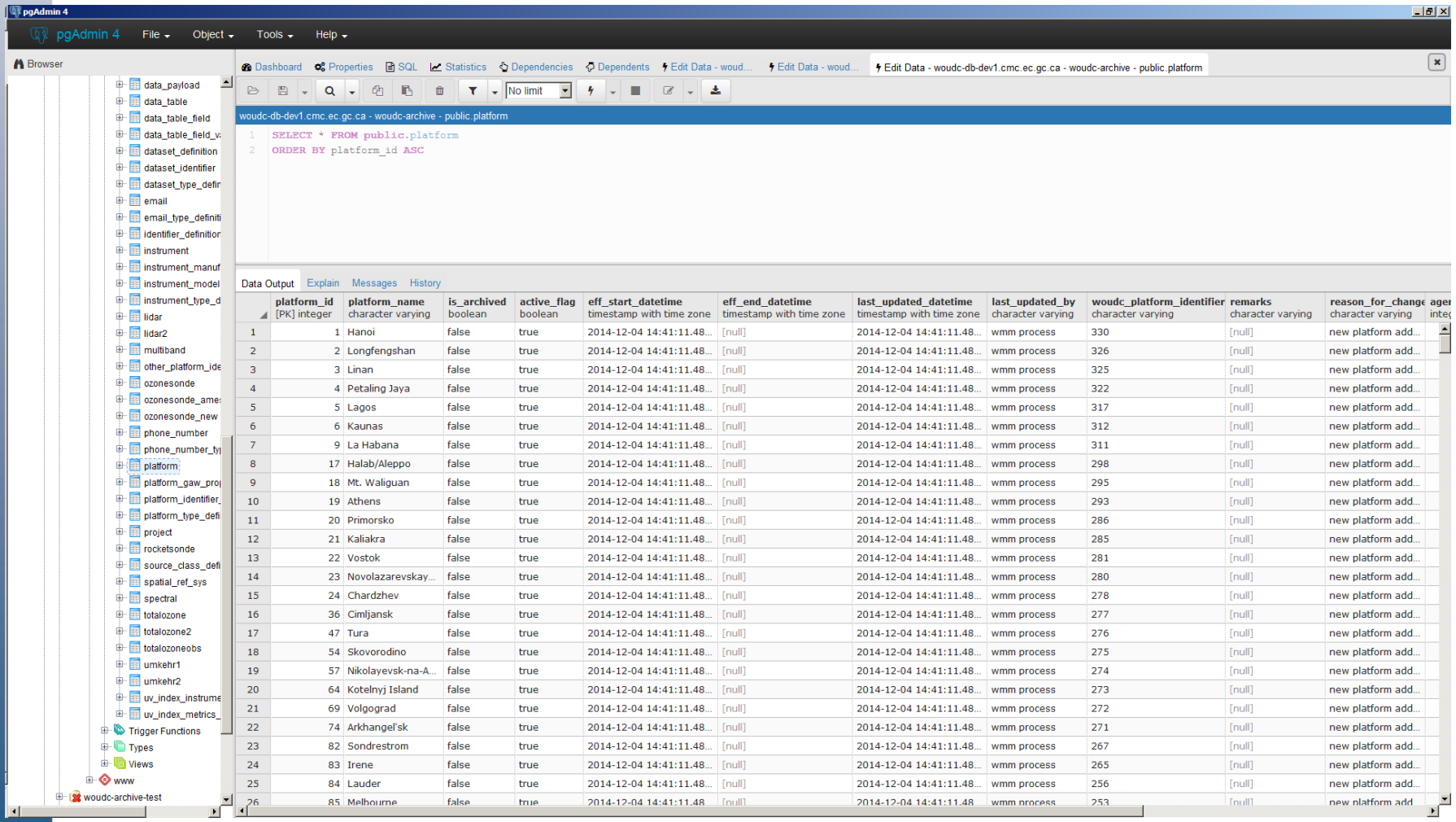
WMDR Development



WOUDC Implementation

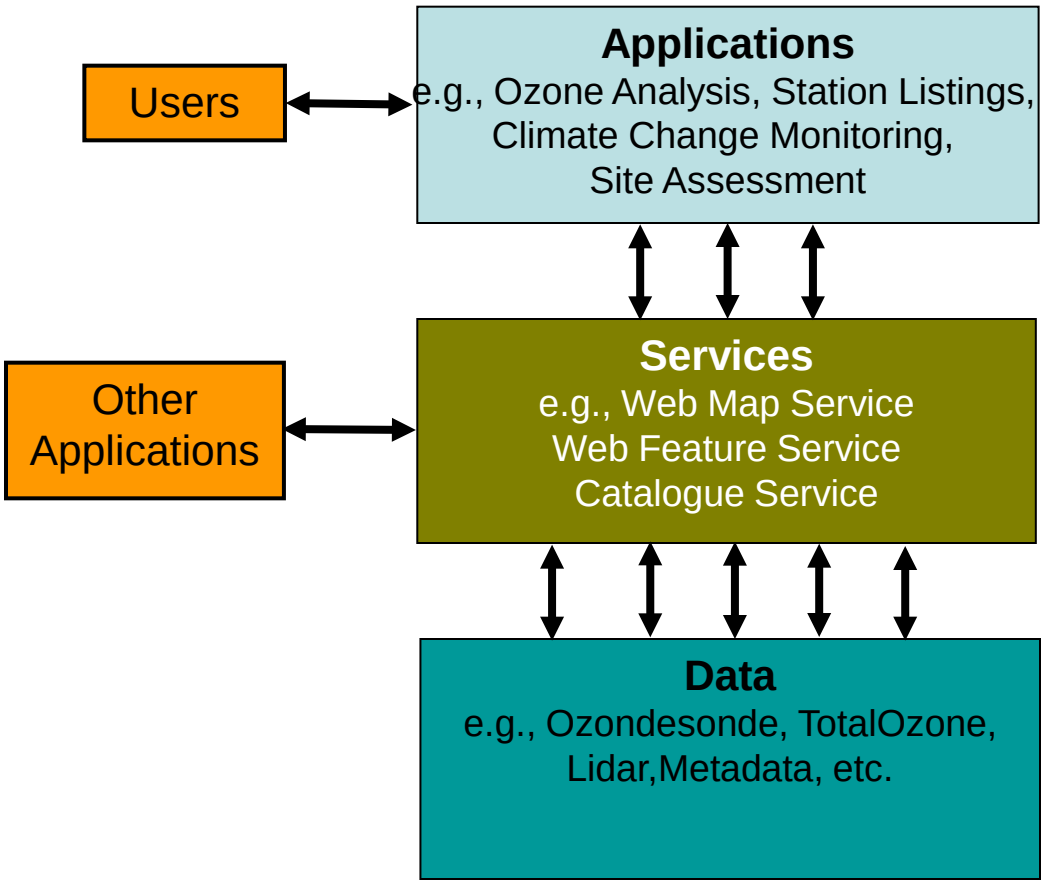
- Data model
- Architecture
- Web Services
- Templating

Data Model



platform_id [PK] integer	platform_name character varying	is_archived boolean	active_flag boolean	eff_start_datetime timestamp with time zone	eff_end_datetime timestamp with time zone	last_updated_datetime timestamp with time zone	last_updated_by character varying	woudc_platform_identifier character varying	remarks character varying	reason_for_change character varying
1	Hanoi	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	330	[null]	new platform add...
2	Longfengshan	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	326	[null]	new platform add...
3	Linan	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	325	[null]	new platform add...
4	Petaling Jaya	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	322	[null]	new platform add...
5	Lagos	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	317	[null]	new platform add...
6	Kaunas	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	312	[null]	new platform add...
7	La Habana	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	311	[null]	new platform add...
8	Halab/Aleppo	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	298	[null]	new platform add...
9	Mt. Waliguan	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	295	[null]	new platform add...
10	Athens	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	293	[null]	new platform add...
11	Primorsko	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	286	[null]	new platform add...
12	Kaliakra	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	285	[null]	new platform add...
13	Vostok	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	281	[null]	new platform add...
14	Novolazarevskay...	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	280	[null]	new platform add...
15	Chardzhev	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	278	[null]	new platform add...
16	Cimljansk	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	277	[null]	new platform add...
17	Tura	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	276	[null]	new platform add...
18	Skovorodino	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	275	[null]	new platform add...
19	Nikolayevsk-na-A...	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	274	[null]	new platform add...
20	Kotelnyj Island	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	273	[null]	new platform add...
21	Volgograd	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	272	[null]	new platform add...
22	Arkhangel'sk	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	271	[null]	new platform add...
23	Sondrestrom	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	267	[null]	new platform add...
24	Irene	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	265	[null]	new platform add...
25	Lauder	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	256	[null]	new platform add...
26	Melbourne	false	true	2014-12-04 14:41:11.48...	[null]	2014-12-04 14:41:11.48...	wmm process	253	[null]	new platform add...

Architecture



For Example…

A community website provides station locations for ozone measurements

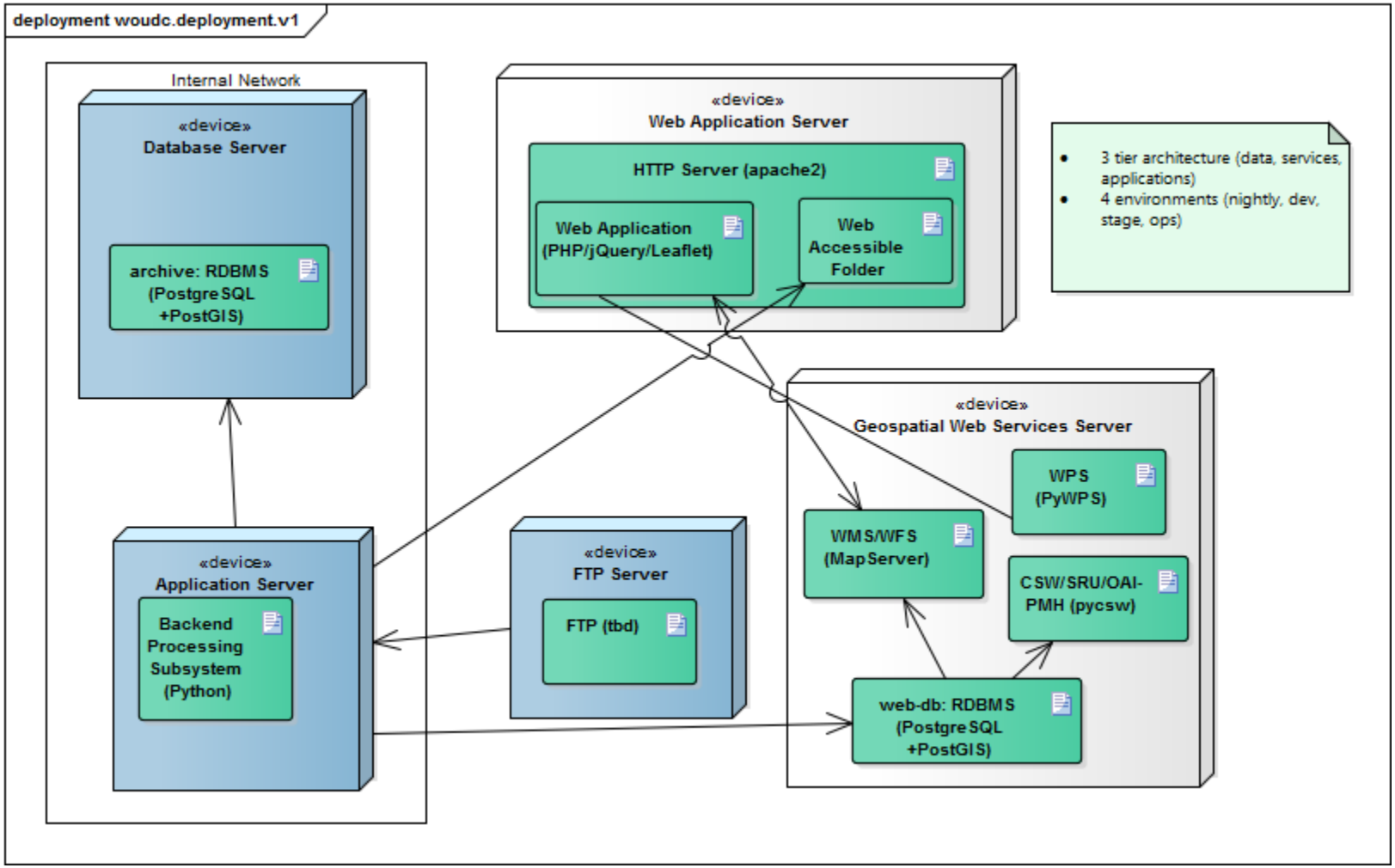
Uses services

Based on Data Contributors

Data / Standards / Technology Matrix

	Service / Interface	Format / Encoding	Software
Discovery Metadata	OGC CSW SRU OAI-PMH OpenSearch	ISO 19115 WMO Core Metadata Profile	pycsw
Station/ Instrument Metadata	OGC WMS OGC WFS	Maps WIGOS CSV KML OGC GML	MapServer
Observations	OGC WMS OGC WFS	CSV KML Shapefiles GeoJSON	MapServer
Processing (Format Transformation, Data Validation)	OGC WPS	WPS, ISO 19115 (Data Quality)	PyWPS

Deployment


deployment woudc.deployment.v1
Internal Network
«device»
Database Server
archive: RDBMS
(PostgreSQL
+PostGIS)
«device»
Application Server
Backend
Processing
Subsystem
(Python)
«device»
FTP Server
FTP (tbd)
«device»
Web Application Server
HTTP Server (apache2)
Web Application
(PHP/jQuery/Leaflet)
Web
Accessible
Folder
«device»
Geospatial Web Services Server
WPS
(PyWPS)
WMS/WFS
(MapServer)
CSW/SRU/OAI-
PMH (pycsw)
web-db: RDBMS
(PostgreSQL
+PostGIS)
3 tier architecture (data, services, applications)
4 environments (nightly, dev, stage, ops)

WOUDC Implementation Pipeline

```
<wmdr:WIGOSMetadataRecord xsi:schemaLocation="http://def.wmo.int/wmdr/2016 http://schemas.wmo.int/wm
  <gml:description>Metadata record of WOUDC stations</gml:description>
  <gml:identifier codeSpace="http://wigos.wmo.int">http://woudc.org/data/stations</gml:identifier>
  <gml:name>stations</gml:name>
- <wmdr:facilityDefinition>
  - <wmdr:ObservingFacility gml:id="observingFacility-455">
      <gml:identifier codeSpace="http://wigos.wmo.int">http://wigos.wmo.int/0-20008-0-ARG</gml:identifier>
      <gml:name>Chisinau</gml:name>
    - <wmdr:onlineResource>
      - <gmd:CI_OnlineResource>
        - <gmd:linkage>
            <gmd:URL>http://woudc.org/data/stations/?id=455</gmd:URL>
          </gmd:linkage>
        - <gmd:protocol>
            <gco:CharacterString>WWW:LINK</gco:CharacterString>
          </gmd:protocol>
        - <gmd:name>
            <gco:CharacterString>Chisinau</gco:CharacterString>
          </gmd:name>
        - <gmd:description>
            <gco:CharacterString>Station listing for Chisinau</gco:CharacterString>
          </gmd:description>
        </gmd:CI_OnlineResource>
```

http://geo.woudc.org/ows?service=WFS&version=1.1.0&request=GetFeature&typename=stations&outputformat=wmdr:MetadataRecord

pygeometa

- Python tool to generate geospatial metadata
  - https://github.com/geopython/pygeometa
- Discovery metadata
  - ISO 19115/19139
  - ISO North American Profile
  - WMO Core Metadata Profile
- Template-based
- Inputs
  - Configuration files
- Flexible workflow / pipelines
  - Create static files
  - Integration into (Python) application

pygeometa

- Proof of Concept (courtesy Dominic Lowe)
  - Configuration: https://github.com/domlowe/pygeometa/blob/master/examples/wigos/wigos.mcf
  - Template: https://github.com/domlowe/pygeometa/blob/master/pygeometa/templates/wigos/main.j2