WORLD-LEADING DISCOVERIES AT A CRITICAL TIME

OCEAN NETWORKS CANADA

OCEANS 2.0 API
PROGRAMMATIC ACCESS TO ONC DATA
TRAINING SESSION

Ryan Ross| July 27, 2017

AN INITIATIVE OF University of Victoria

WORLD-LEADING DISCOVERIES AT A CRITICAL TIME

OCEAN NETWORKS CANADA

Ocean Networks Canada enhances life on Earth by providing knowledge and leadership that deliver solutions for science, society, and industry.

AN INITIATIVE OF University of Victoria

ONC API

Programmatic access to ONC's Data Archive

The Empower Project

The Web Services
- Discovery
- Data Delivery

The Client Libraries
- Python
- MATLAB
- R *

The Documentation
- Samples
- Client libraries
- Use Cases

Demos

Testing Plan

Road Ahead

ONC API

Empower Project Overview

- A Research Platform for User-Defined Oceanographic Data Products
- 2 Year project funded by CANARIE
- Two Phases

API

Provide researchers with programmatic access to the Oceans 2.0 data products through a specially designed Application Programming Interface (API)

Sandbox

Enable researchers to define, test, use and share processing code for user-defined data products in a custom-designed programming environment.

ONC API

API Phase Overview

User-Centred design approach

1) Identify existing and future users
2) Develop hypothesis based on domain knowledge and current usage
3) Develop questionnaires to validate hypothesis and elicit responses about:
    a) Research focus
    b) Data needs
    c) Analysis needs
    d) Analytical tools
    e) Usage and performance expectations
    f) Publishing needs
4) Perform interviews and capture responses
5) Analyze responses and organize into functional groups or clusters of requirements
6) Develop personas to characterize user goals and behaviours
7) Develop use cases to capture the functional needs and requirements of the personas
8) Develop user stories to inform the functional design
9) Build the API in an iterative cycle of sprints with user feedback and testing throughout

ONC API

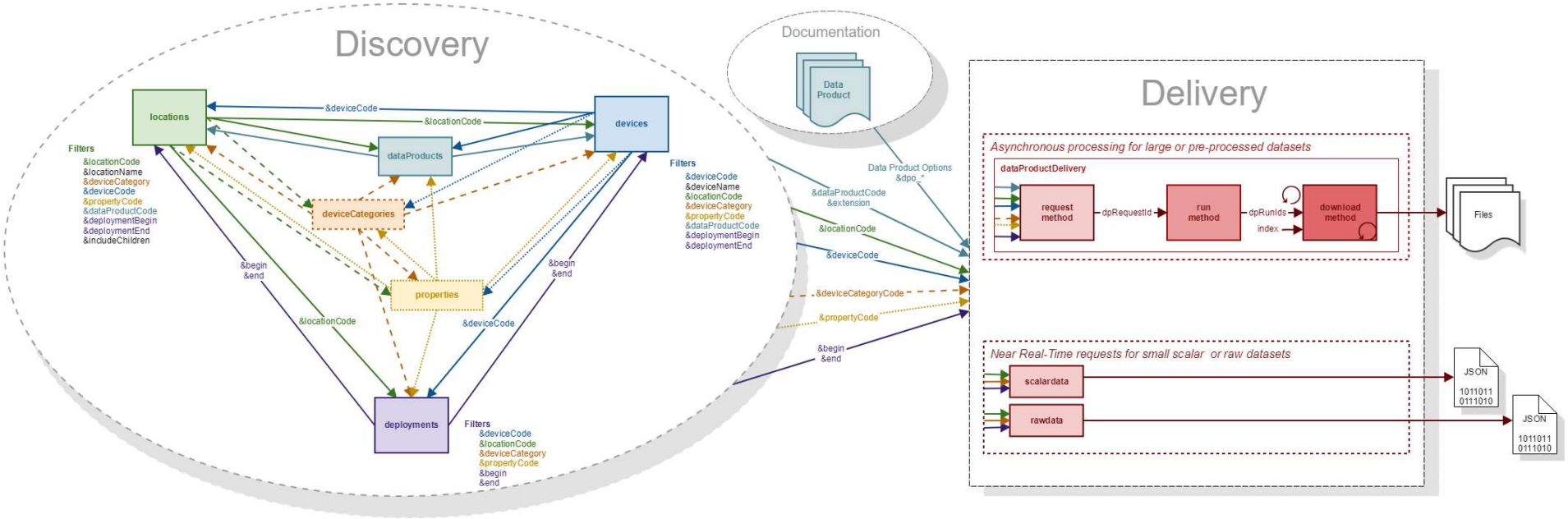
API Phase Overview

Results

- 12 user interviews and questionnaires
- 14 existing and future users
- 4 Use Cases
  - Bird Studies Canada
  - Ouranos
  - Academic
  - ONC Internal
- 3 new personas
  - Lead Researcher
  - Scientific Modeller
  - Scientific Developer
- 54 User Stories
- 4 functional groups
  - Discovery
  - Filtering
  - Scripting
  - Internal Needs

ONC API

Understanding the web services

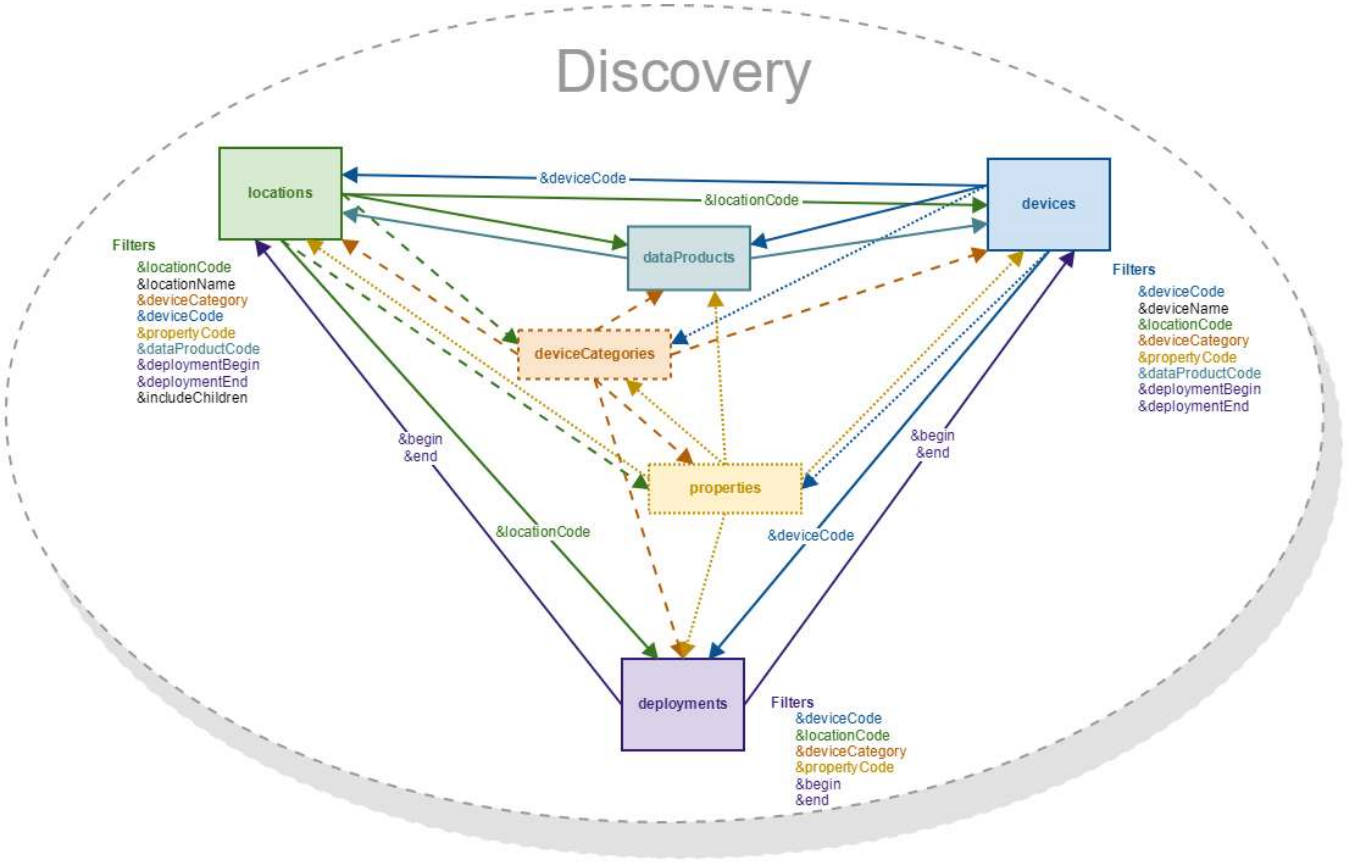


- Data Discovery
  - 6* web services to help discover what data can be downloaded
    - Where is data available
    - What devices are available
    - What device categories are available
    - What properties are available
    - What data products are available
    - When are devices deployed

      * deployments service in development
- Data Delivery (Download)
  - 1 web service to download data as customizable data product
    - Request a Data Product be created
    - Download the Data Product once complete
  - 2 web services for near real-time data access
    - Scalar data
    - Raw data

ONC API

Data Discovery Services



- 6 Services to Discover the codes needed for the Delivery services
- Common filtering and output allows for discovery of missing/unknown information

ONC API

Data Discovery Services

6 Service Endpoints

…/api/locations

- ONC search tree nodes (aka, stations) that data can be downloaded from
- Uniquely identified by locationCode.
  ie, BACME *(Barkley Canyon / MidEast)*

…/api/devices

- Instruments that have one or more sensors that observe a property or phenomenon with a goal of producing an estimate of the value of a property
- Uniquely identified by deviceCode
  ie, CAM-TEMPO-MINI-2 *(Tempo-Mini Colour Video AXIS Q1755 [Camera 2])*

…/api/deviceCategories

- Device Category grouping
- Uniquely identified by deviceCategoryCode
  ie, VIDEOCAM *(Video Camera)*
  DSC *(Still Camera)*

…/api/properties

- Observable phenomenon (aka, variables)
- Common name given to sensor types (ie, oxygen, pressure, temperature, etc)
- Uniquely identified by propertyCode

…/api/dataProducts

- Data Products available for download
- Uniquely identified by dataProductCode & extension
  Ie, 3DCIS *(3D Camera Image Stitching)* & tar
  MP4V *(MP4 Video)* & mp4

…/api/deployments *

- Instrument deployments
- Uniquely identified by siteDeviceId
- Provides a relationship between locations and devices
- Each Site Device has a location code, device code, lat/long/depth and deployment date range.
  Ie, ???

** currently under development*

ONC API

Data Discovery Services

- Reductive Filtering
  - Additional filters can be added to the url to reduce the selection

Example - return all locations that have a Video Camera deployed between two dates

http://dmas.uvic.ca/api/locations?method=get
&token=x6ecb030-1835-1dc5-29nc-q58l7wjpl607
&deviceCategory=VIDEOCAM
&deploymentBegin=2016-07-27T00:00:00.000Z
&deploymentEnd=2016-08-01T23:59:59.999Z

- Returned as a JSON payload


```
[{"dataProductCode":"AD",
  "dataProductName":"Audio Data",
  "extension":"wav",
  "helpDocument":"https://wiki.oceannetworks.ca/display/DP/7"}]
```

ONC API

Data Product Delivery

- Three step process using dataProductDelivery web service
  1. Request a data product using the **request** method
     - Include data product filters in URL
       - locationCode and deviceCategoryCode *or* locationCode, deviceCategoryCode and propertyCode *or* deviceCode *or* deviceCode and propertyCode
       - dataProductCode and extension
       - begin and end
       - Data Product Options vary with data product
     - Returns a RequestId
     - Returns run time and size estimates
       - Can be used for request verification
  2. Run the data product using the **run** method
     - Include RequestId in URL
     - Starts the data product generation process by adding it to the Task Queue
  3. Download the data product using the **download** method
     - Include RunId and index in URL
     - Informs on process status with messages in the payload HTTP status codes
     - Downloads file when process is complete
     - Requires same token as run request

ONC API

Using the web services

- Via browser
  - Make a simple http request using a browser link to return information

- Via code
  - Any language that supports HTTP requests including:
    - Python *
    - MatLab *
    - R *
    - JavaScript
    - C++
    - Java

      *Available ONC client libraries

- All requests require a user token
  - Create user at https://dmas.uvic.ca/Registration?service=http://dmas.uvic.ca/
  - Generate user token on 'Web Services API' tab at http://dmas.uvic.ca/Profile

  - Please use your own token
    - It allows us to better understand your data needs and inform you when changes or improvements are made to ONC web services

ONC API

Service requests in language of choice

Python

```python
import requests
response = requests.get('http://dmas.uvic.ca/api/locations',
                        params={'method':'get',
                                'token':'b6ede000-1865-4ac3-94ad-e87d8bdfd307',
                                'deviceCategoryCode':'VIDEOCAM',
                                'deploymentBegin':'2016-07-27T00:00:00.000Z',
                                'deploymentEnd':'2016-08-01T23:59:59.999Z'})
```

MATLAB

```matlab
response = webread('http://dmas.uvic.ca/api/locations',...
                   'method','get','token',...
                   'b6ede000-1865-4ac3-94ad-e87d8bdfd307',...
                   'deviceCategoryCode','VIDEOCAM',...
                   'deploymentBegin','2016-07-27T00:00:00.000Z',...
                   'deploymentEnd','2016-08-01T23:59:59.999Z');
```

R

```r
library(httr)
response <- GET("http://dmas.uvic.ca/api/locations",
                query = list(method="get",
                             token="b6ede000-1865-4ac3-94ad-e87d8bdfd307",
                             deviceCategoryCode="VIDEOCAM",
                             deploymentBegin="2016-07-27T00:00:00.000Z",
                             deploymentEnd="2016-08-01T23:59:59.999Z"))
```

ONC API

Client Libraries

Client libraries provide

- Quick, easy and consistent access to ONC data and resources
- In scientific programming language of choice
- With minimal lines of code

Currently available

- python 2.7+
- Matlab R2017a
- R 3.3+

https://wiki.oceannetworks.ca/display/O2A/Client+Libraries

ONC API

Using the Client Libraries

- Python
  - Install package using pip
  - Add to library to script using

```
from onc.onc import ONC
```

  - Create ONC object using

```
onc = ONC("YOUR_TOKEN")
```

or

```
onc = ONC("YOUR_TOKEN",True,False,"c:/ONC/Data")
```

- MATLAB
  - Download Add-On Toolbox from ONC Wiki and install
  - Create ONC object using

```
o = ONC("YOUR_TOKEN")
```

or

```
o = ONC("YOUR_TOKEN",true,false,"c:/ONC/Data")
```

- R
  - Download package from ONC Wiki and install
  - Add library to script using

```
library(onc)
```

  - Create ONC object using

```
onc = new("onc",token="YOUR_TOKEN")
```

or

```
onc = new("onc",token="YOUR_TOKEN",production=TRUE,showInfo=FALSE,outPath="c:/ONC/Data")
```

ONC API

Client Libraries – Single line of code

Download Time Series Scalar Data Product in CSV format for ADCP 2 MHZ at Barkley Canyon - Axis

- Python

```
results = onc.orderDataProduct({'locationCode':'BACAX',
                                'deviceCategoryCode':'ADCP2MHZ',
                                'dataProductCode':'TSSD',
                                'extension':'csv',
                                'begin':'2016-07-27T00:00:00.000Z',
                                'end':'2016-08-01T00:00:00.000Z',
                                'dpo_qualityControl':1,'dpo_resample':'none', 'dpo_dataGaps':0})
```

- MATLAB

```
results = o.orderDataProduct(struct('locationCode','BACAX',...
                                    'deviceCategoryCode','ADCP2MHZ',...
                                    'dataProductCode','TSSD',...
                                    'extension','csv',...
                                    'begin','2016-07-27T00:00:00.000Z',...
                                    'end','2016-08-01T00:00:00.000Z',...
                                    'dpo_qualityControl',1,'dpo_resample','none','dpo_dataGaps',0));
```

- R

```
results = onc.orderDataProduct(onc,list(locationCode="BACAX",
                                    deviceCategoryCode="ADCP2MHZ",
                                    dataProductCode="TSSD",
                                    extension="csv",
                                    begin="2016-07-27T00:00:00.000Z",
                                    end="2016-08-01T00:00:00.000Z",
                                    dpo_qualityControl=1, dpo_resample="none", dpo_dataGaps=0))
```

ONC API

Documentation

- Guide
  - Overview of the
  - Requesting Data Products using the API
- API Reference
- Sample Code
- Client Libraries
- Use Cases

https://wiki.oceannetworks.ca/display/O2A/Oceans+2.0+API+Home

ONC API

Demo

ONC API

Beta Test

- Track Testing Results at
  https://internal.oceannetworks.ca/display/ONCData/1.4.2.3+User+Testing+Results

ONC API

Road Ahead

- Internal Testing – July 2017
- External User Testing (Beta Program) – Aug 2017
  - Webinar for all participants
- Issue Resolution – Aug / Sept 2017
- Release – Oct 2017

OCEAN NETWORKS CANADA

THANK YOU!

Ocean Networks Canada is funded by the Canada Foundation for Innovation, Government of Canada, University of Victoria, Government of British Columbia, CANARIE, and IBM Canada.

@ocean_networks OceanNetworksCanada visit: **oceannetworks.ca**