Making Sense of Millions of Observations Using Open Standards

Air Sensors 2013
Big Data: Management and Analysis

Luis Bermudez, Ph.D.
EPA, North Carolina, March 19, 2013
Agenda

• BIG Data
• Why do we need standards for BIG Data?
• Why Open Geospatial Consortium (OGC) Standards?
• Sensor Web Enablement (SWE)
• Exemplar Projects
Big Data = 4Vs

[M. Stonebraker and IBM]
Volume

Twitter
90 Million tweets / day
8 terabytes / day

640 terabytes of operational data on just one Atlantic crossing

Velocity

3 GB per second
LOFAR: distributed sensor array farms for radio astronomy
Veracity

Level-2 Flags

<table>
<thead>
<tr>
<th>Bit</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>ATMFAIL</td>
<td>Atmospheric correction failure</td>
</tr>
<tr>
<td>02</td>
<td>LAND</td>
<td>Pixel is over land</td>
</tr>
<tr>
<td>03</td>
<td>PRODWARN</td>
<td>One or more product warnings</td>
</tr>
<tr>
<td>04</td>
<td>HIGLINT</td>
<td>High sun glint</td>
</tr>
<tr>
<td>05</td>
<td>HILT</td>
<td>Observed radiance very high or saturated</td>
</tr>
<tr>
<td>06</td>
<td>HISATZEN</td>
<td>High sensor view zenith angle</td>
</tr>
<tr>
<td>07</td>
<td>COASTZ</td>
<td>Pixel is in shallow water</td>
</tr>
</tbody>
</table>

How was this calculated?
Variety
Variety – Benefit Areas

- Disasters
- Health
- Energy
- Climate
- Water
- Weather
- Ecosystems
- Agriculture
- Biodiversity
Variety – Systems
Variety - Sensors
Variety - Models

Approximate temperature changes Canada (2050)/Global (2100):

- Blue: -1 – 0
- Green: 1– 2/4
- Yellow: 2/4 – 3/5
- Orange: 3/5 – 5/7

Short Term | Long Term
What’s in common?
Geospatial Location
Geospatial Integration
Air Quality Monitoring Station 03, August 15 2002

- Good
- Fair
- Poor

Population per square kilometre 2001, NB Health Regions

- 4.84000
- 4.84001 - 5.51000
- 5.51001 - 7.06000
- 7.06001 - 16.30000
- 16.30001 - 18.25000

Copyright © 2004 New Brunswick Lung Association / Association Pulmonaire du Nouveau Brunswick

Courtesy: Eddie Oldfield, Director, NB Climate Change Hub

Copyright © 2013 Open Geospatial Consortium
How?

Welcome to the Geospatial Platform

The Geospatial Platform provides shared and trusted geospatial data, services, and applications for use by government agencies, their partners and the public.

Join the Dialogue

Browse Popular Ideas

Specify Open Standards

The Roadmap for the Geospatial Platform states it “will be established as a service-oriented architecture based upon common, secure, interoperable and scalable open-standards based technologies.” The Roadmap also says that through the FGDC, standards that shall be evaluated for adoption by the Federal Government include: Geospatial data publication standards from the Open Geospatial Consortium (OGC), the International... more »
Specify Open Standards

The Roadmap for the Geospatial Platform states it “will be established as a service-oriented architecture based upon common, secure, interoperable and scalable open-standards based technologies.” The Roadmap also says that through the FGDC, standards that shall be evaluated for adoption by the Federal Government include: Geospatial data publication standards from the Open Geospatial Consortium (OGC), the International... more »

http://geoplatfom.ideascale.com
BIG DATA = ... VARIETY OF DATA

COMMONALITY = LOCATION

IMPORTANT FOR INTEGRATION
-> GEOSPATIAL PLATFORM

NEED OPEN STANDARDS
Open Geospatial Consortium

Only industry organization in the world focusing on location standards
Forum, Process and People
450 Member Organization
4000 Individuals
Collaboration and Coordination
Capturing Real Needs

Geospatial and location standards for:

- Aviation
- Built Environment & 3D
- Business Intelligence
- Defense & Intelligence
- Emergency Response & Disaster Management
- Geosciences & Environment
- Government & Spatial Data Infrastructure
- Mobile Internet & Location Services
- Sensor Webs
- University & Research
Interoperability Program

I have not failed, I’ve just found 10,000 ways that won’t work.

Thomas Edison
Compliance Program

700 implementing products
190 compliant products in the market
Sensor Web Enablement
Sensor Web Requirements

- Discover
- Describe
- Get Data
- Task
- Subscribe
SWE Solution
SWE Interface Standards

• Sensor Observation Service
• Sensor Planning Service
• Sensor Alert Service
• PUCK
SWE Encoding Standards

- Observations and Measurements (O&M)
- SensorML
- SWECommon
- PUCK
Sensor Observation Service

Client → SOS

1. GetCapabilities
2. Capabilities.xml
3. DescribeSensor
4. SensorML
5. GetObservation
6. O&M / SWE Common
SWE Harmonization

OASIS Emergency Management TC

Harmonization with SAS (PULSENet)

• Common Alert Protocol (CAP)
• EDXL Distribution Element
SWE Harmonization

IEEE P1451 IEEE 1451 Smart Transducer Interface Standard

“Ocean Science Interoperability Experiment” Report details use of SWE over 1451
SWE Harmonization

DoD Chemical, Biological, Radiological, and Nuclear (CBRN) sensors

PULSENet demonstrated the use of SensorML to describe these sensors
The emerging Internet of things:
-- indoor/outdoor location
-- sensor webs
-- building information models
-- location apps
-- location marketing
-- smart grid
Sensor Web for IoT Working Group

Develop OGC standard for access to sensors in an IoT/WoT environment

52° North Initiative for Geospatial Open Source Solutions
AIT Austrian Institute of Technology GmbH
Arizona State University (ASU)
Botts Innovative Research
Cosm Ltd
CSIRO
EADS ASTRIUM
EDINA, University of Edinburgh
Envitia Ltd.
Esri
Esri Canada
ETRI (Electronics & Telecommunications Research Institute)
European Centre for Medium-Range Weather Forecasts
Federal Agency for Cartography and Geodesy
Fraunhofer-Gesellschaft

GEOMATYS
Geonovum
Hankuk University of Science and Technology
Hitachi, Ltd., Center for Sustainable Social Infrastructure
Industrial Technology Research Institute
Institut National de Recherche pour l'Intelligence Artificielle (INRIA)
interactive instruments
Intergraph Corporation
Lockheed Martin
METEO-FRANCE
National Institute of Water and Atmospheric Research
NAVTEQ
Open Geospatial Consortium, Inc.
Open Site Plan
Oracle USA

Perey, Christine
Prodevelop, S.L.
Terradue Srl
The SI Organization, Inc.
Universitat Politècnica de Catalunya, UPC (SARTI research group)
University of Calgary
University of California, San Diego Supercomputer Center
University of Muenster - Institute for Geoinformatics
University of the Bundeswehr - ITIS
University of Tokyo CSIS
US National Geospatial-Intelligence Agency (NGA)
WISC Enterprises
Woolpert
Heterogeneous Sensors Become Homogeneous Things in Smart Cities

Maria Fazio, U. of Mezzina

The architecture implementation is based on Sensor Web Enablement standard specifications and makes use of the Contiki Operating System for accomplishing the Internet of Things.
Exemplar Projects
Volume Examples
Integrated Ocean Observing System 2500 Platforms
OBIS – 31 Million Observations

Uses OGC Services to publish maps and download data
Groundwater Interoperability

2 million Wells in Canada
Velocity
NASA Sensor Web

Data Processing Node:
- Web Coordinate Transformation Service (WCTS)
- Web Processing Service (WPS)
- Web Coverage Service (WCS)

SWE Node:
- Web Map Service (WMS)
- Web Feature Service (WFS)
- Sensor Planning Service (SPS)
- OGC Pub/Sub (OPSB)
- Sensor Observation Service (SOS)
- GeoTorrent Distribution Service

Sensor Data Products

Internet

RSS Feeds

Campaign Manager API

Workflows

Campaign Manager

Identity Management Service (OpenID 2.0)

Components outlined in red are part of NASA generic SensorWeb toolbox

floods, fires, volcanoes etc

EO-1 Satellite

In-situ Sensor Data Node

UAV Sensor Data Node

Sensor Data Products

Campaign Manager API

RSS Feeds

Workflows
Open GeoSMS Standard

http://maps.google.com/maps?q=38.9985,-77.030275&GeoSMS
I am here ..
How OGC’s Open GeoSMS serves for Disaster Management

[Guest blog post by Kuo-Yu slayer Chuang from Taiwan's Industrial Technology Research Institute, ITRI.]

I'm Kuo-Yu slayer Chuang from ITRI, a government funded research institute in Taiwan. We developed an open standard for exchanging location information via SMS among mobile devices called Open GeoSMS, which has currently been adapted officially by Open Geospatial Consortium (OGC). Several use cases of this standard are shown with the following video:
Variety
Air Quality Sensors and Modeling

PULSEnetTM
CAP Alerts via OGC Sensor Alert Service (SAS)
Unmanned Aircraft Systems

Models
Veracity Example
Sensor Fusion

Impact assessment
>> close beach? <<

Beach attendance / day

Microbial contamination / day

Bathing water quality model
Air Quality monitoring

Air quality management

- Vendor independence
- Cross-border monitoring and alerting
- QA automation
- SANY-compliant data acquisition systems

Other data sources

SANY infrastructure services

Added-value Generic Services

Fusion Services
- temporal
- spatial
- #kind of data

Modelling services
- diffusion
- transport

Visualisation Services
- Colour-coded maps
- Time series
**QUARTOD Quality Control Tests: Waves**

The participants in this effort (shown on the table tab) propose that the following required tests be performed on wave data to meet minimum IOOS quality control standards. In addition, there are a number of recommended tests.

Note: All tests listed are for open ocean waves.

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
<th>Order</th>
<th>Flag</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Gaps</td>
<td>Consecutive N missing data. Maximum number of missing data.</td>
<td>1</td>
<td>Soft</td>
<td>N is user defined. Include in % count.</td>
</tr>
<tr>
<td>Spikes</td>
<td>User defined Points &gt;= M*std with P iterations</td>
<td>2</td>
<td>Soft</td>
<td>Interpolate/extrapolate up to N points. N is user defined. M can be user defined. Recommended M=4. Include in % count.</td>
</tr>
<tr>
<td>Range test</td>
<td>Location, instrument defined.</td>
<td>2</td>
<td>1. Soft</td>
<td>Max/min user defined. 1. Interpolate/extrapolate up to n points. N is user defined. Include in % count. 2. Instrument spec exceeded, reject.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Hard</td>
<td></td>
</tr>
<tr>
<td>Mean shift (segments)</td>
<td>A mean shift &quot;D&quot; occurs in this time series.</td>
<td>3</td>
<td>Hard</td>
<td>Reject entire record. P is user defined.</td>
</tr>
<tr>
<td>Acceleration test</td>
<td>User defined (a&gt;M*'g')</td>
<td>3</td>
<td>Soft</td>
<td>Recommended M&lt;=1/2. Interpolate/extrapolate up to N contiguous points. N is user defined. Include in % count.</td>
</tr>
<tr>
<td>Mean test, variance test</td>
<td>User defined, location dependent</td>
<td>4</td>
<td>1. Soft</td>
<td>1. Flag unexpected values.</td>
</tr>
<tr>
<td>Percent points good</td>
<td>Check for M% good data (based on above 6 criteria)</td>
<td>5</td>
<td>Hard</td>
<td>Recommended M&gt;=90%</td>
</tr>
</tbody>
</table>

**SPECTRAL VALUES**

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
<th>Order</th>
<th>Flag</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>NON-DIRECTIONAL:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operational frequency range test</td>
<td>*defined by the environment and instrument</td>
<td>1</td>
<td>1. Soft</td>
<td>1. Max/min user defined. 2. Instrument spec exceeded, reject.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Hard</td>
<td></td>
</tr>
<tr>
<td>DIRECTIONAL:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident low frequency energy direction</td>
<td>Location defined</td>
<td>1</td>
<td>Soft</td>
<td>User defined</td>
</tr>
</tbody>
</table>
Observable Properties

Sensor/Deployment Files (SensorML)
- Original Equipment Manufacturer (OEM)
- Configuration/Ownership/Deployment (CONDEP)

Process Files (SensorML)
- QC Tests – with QC flags
- Processing Descriptions

Observed and Derived Properties and QC Flags
Conclusion - BIG Data

Velocity

Veracity

Variety

Volume
Conclusion - OGC
Conclusions - Sensor Web Enablement

- Discover
- Describe
- Get Data
- Task
- Subscribe
Conclusion - Successful Examples

Tsunami Warnings

Air Quality

Mobile

Sensor Data Fusion

Quality Marine Sensors
Questions?

Luis Bermudez, Ph.D.
Ibermudez@opengeospatial.org
@berdez on Twitter