

# **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

[http://www.information-management.com/issues/21\\_5/big-data-is-scaling-bi-and-analytics-10021093-1.html](http://www.information-management.com/issues/21_5/big-data-is-scaling-bi-and-analytics-10021093-1.html)



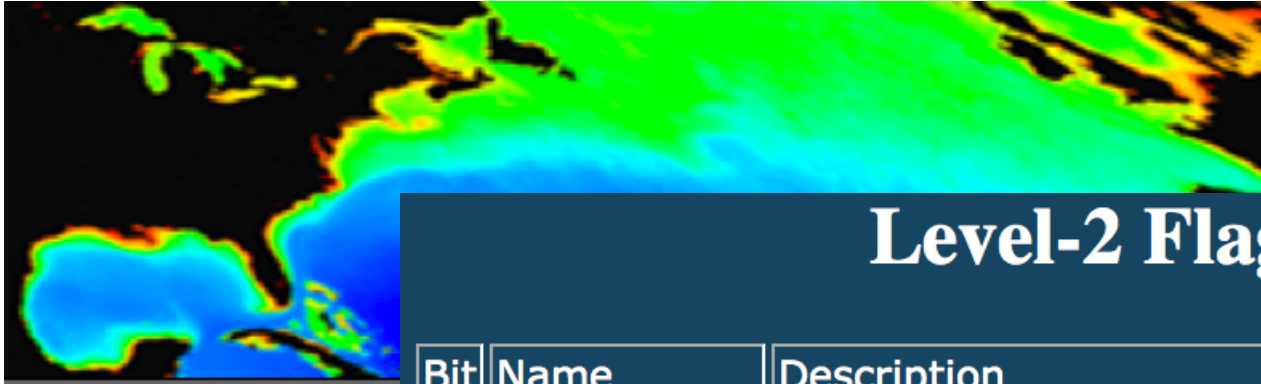
# Velocity



**3 GB per second**

**LOFAR: distributed sensor array farms for radio astronomy**

# Veracity



## Level-2 Flags

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

How was this calculated ?

# Variety

# Variety – Benefit Areas



**Disasters**



**Health**



**Energy**



**Climate**



**Water**



**Weather**



**Ecosystems**



**Agriculture**



**Biodiversity**

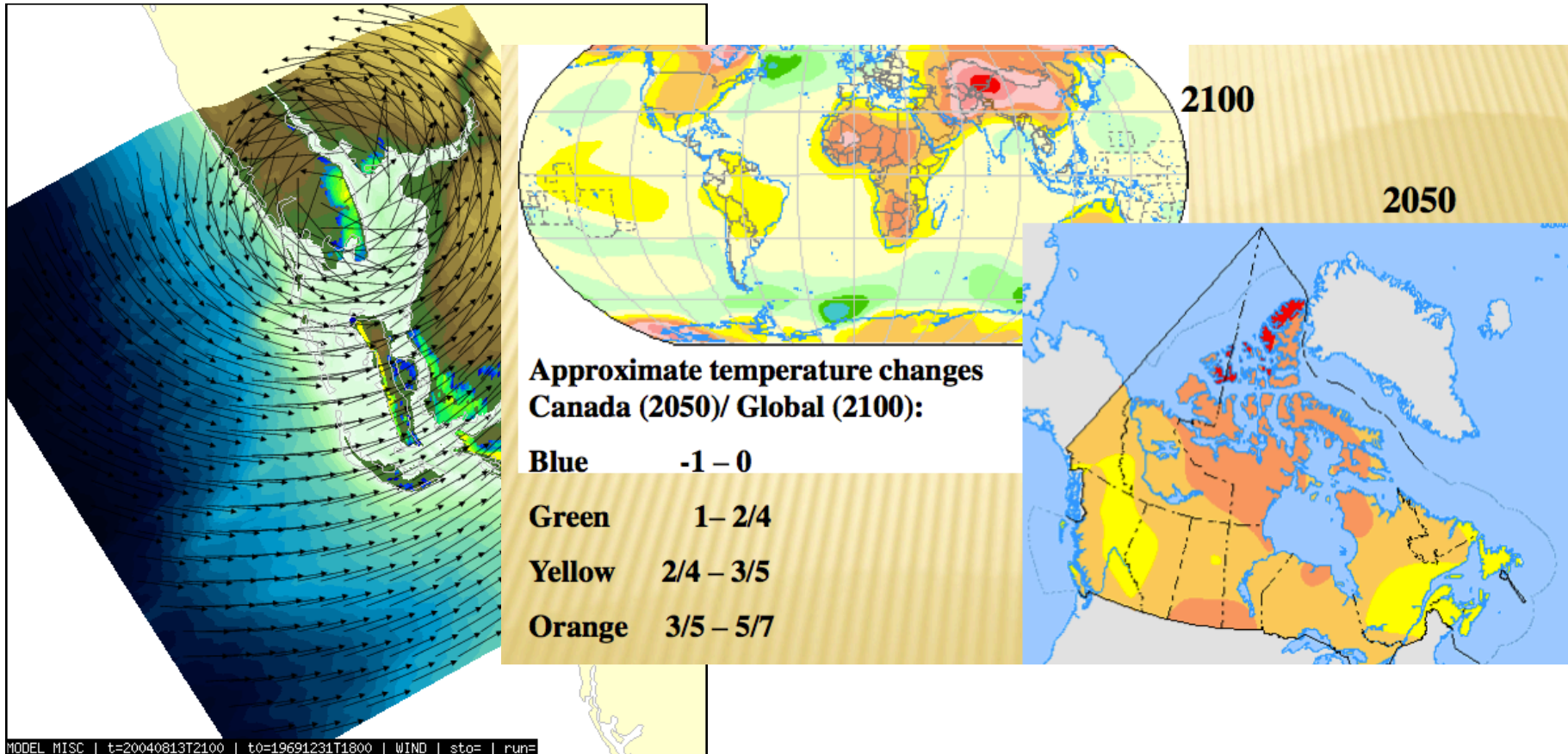
# Variety – Systems





[illegible]

# Variety - Models

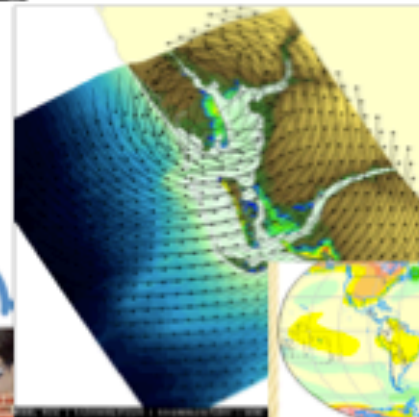


**Short Term**

**Long Term**



# What's in common?



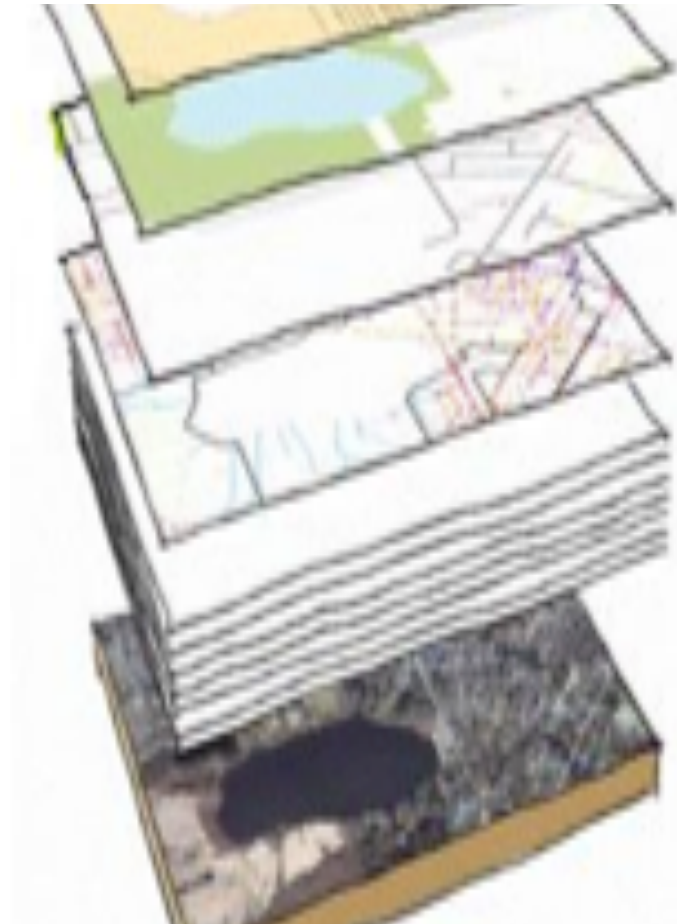


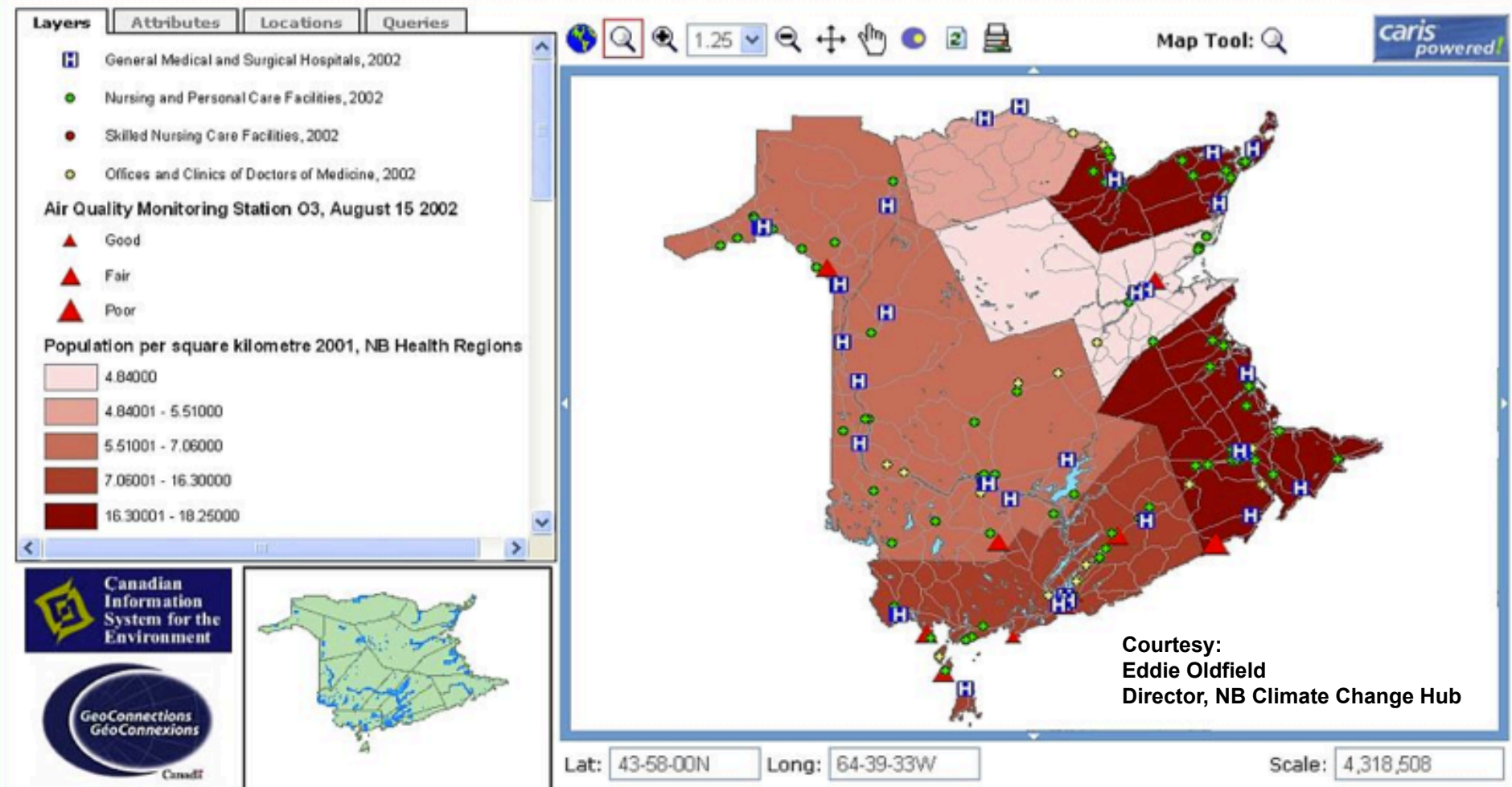


# Geospatial Location



# Geospatial Integration







# How?

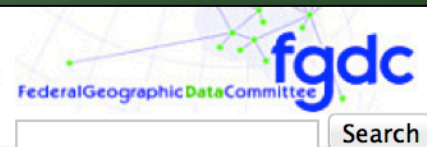
Welcome to the

## 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



new idea 

### Browse Popular Ideas

Recent (28) Popular (28) Hot (0)



I agree

35  
votes



I disagree

COMMON DATA, SERVICES AND APPLICATIONS »

### 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](#) »

# How ?



## Join the Dialogue

 Search

# Specify Open Standards

Campaign

All Ideas

Top Priorities

Additional Resources

Archive

Common Data, Services  
and Applications

I agree

35  
votes

I disagree

COMMON DATA, SERVICES AND APPLICATIONS »

## 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://geoplatform.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**

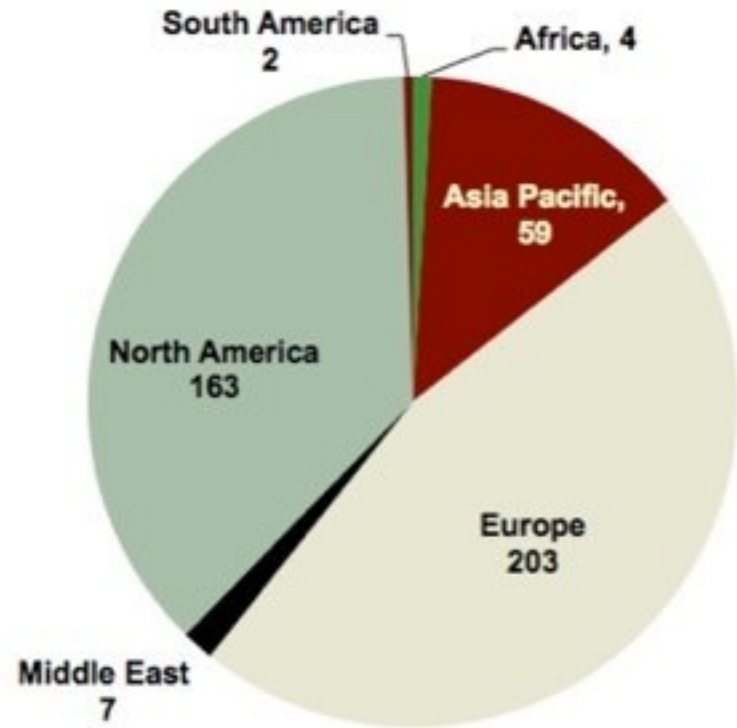
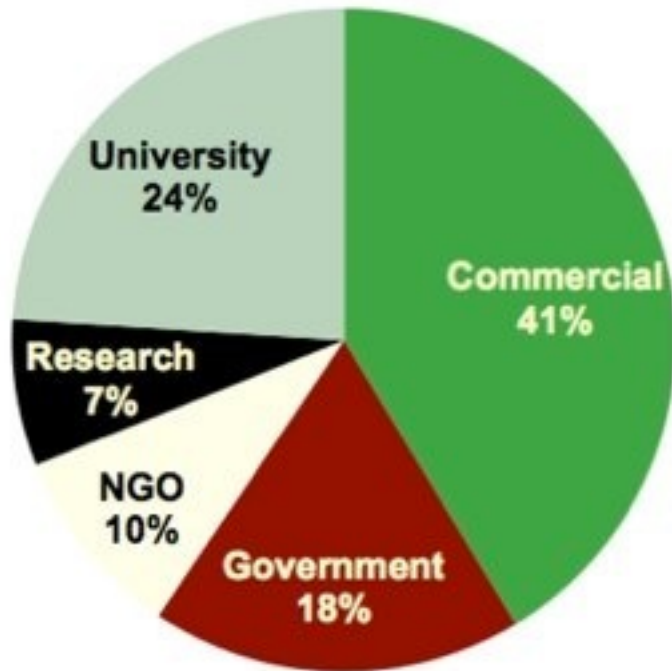


# Forum, Process and People



# 450 Member Organization

## 4000 Individuals



# Collaboration and Coordination



# Capturing Real Needs

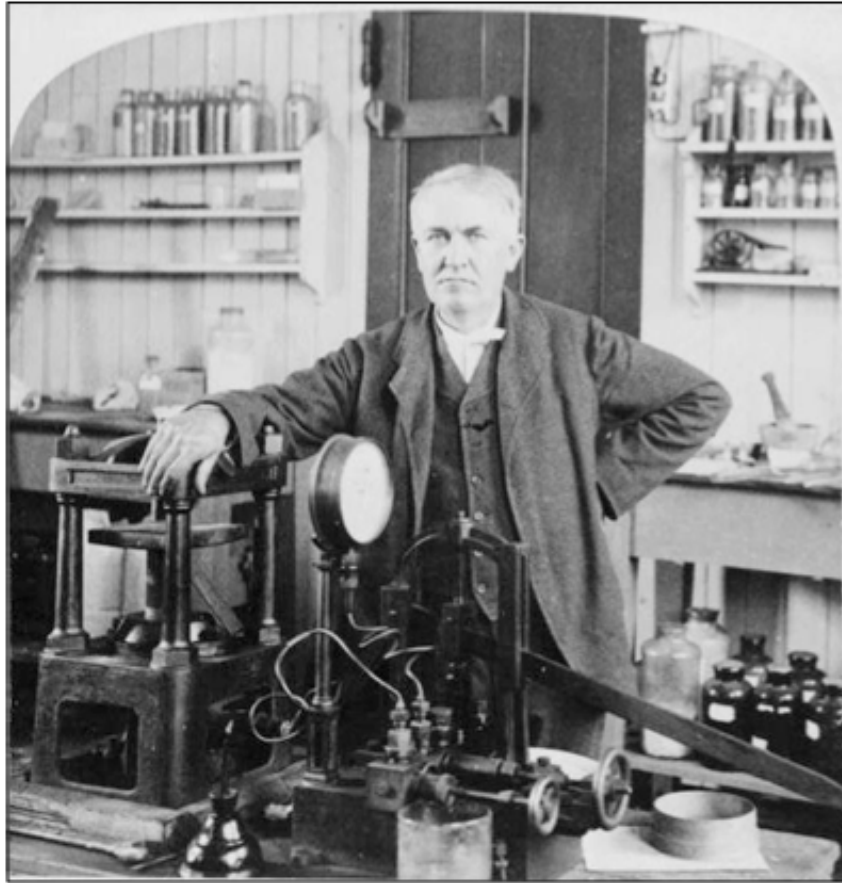
## Geospatial and location standards for:

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



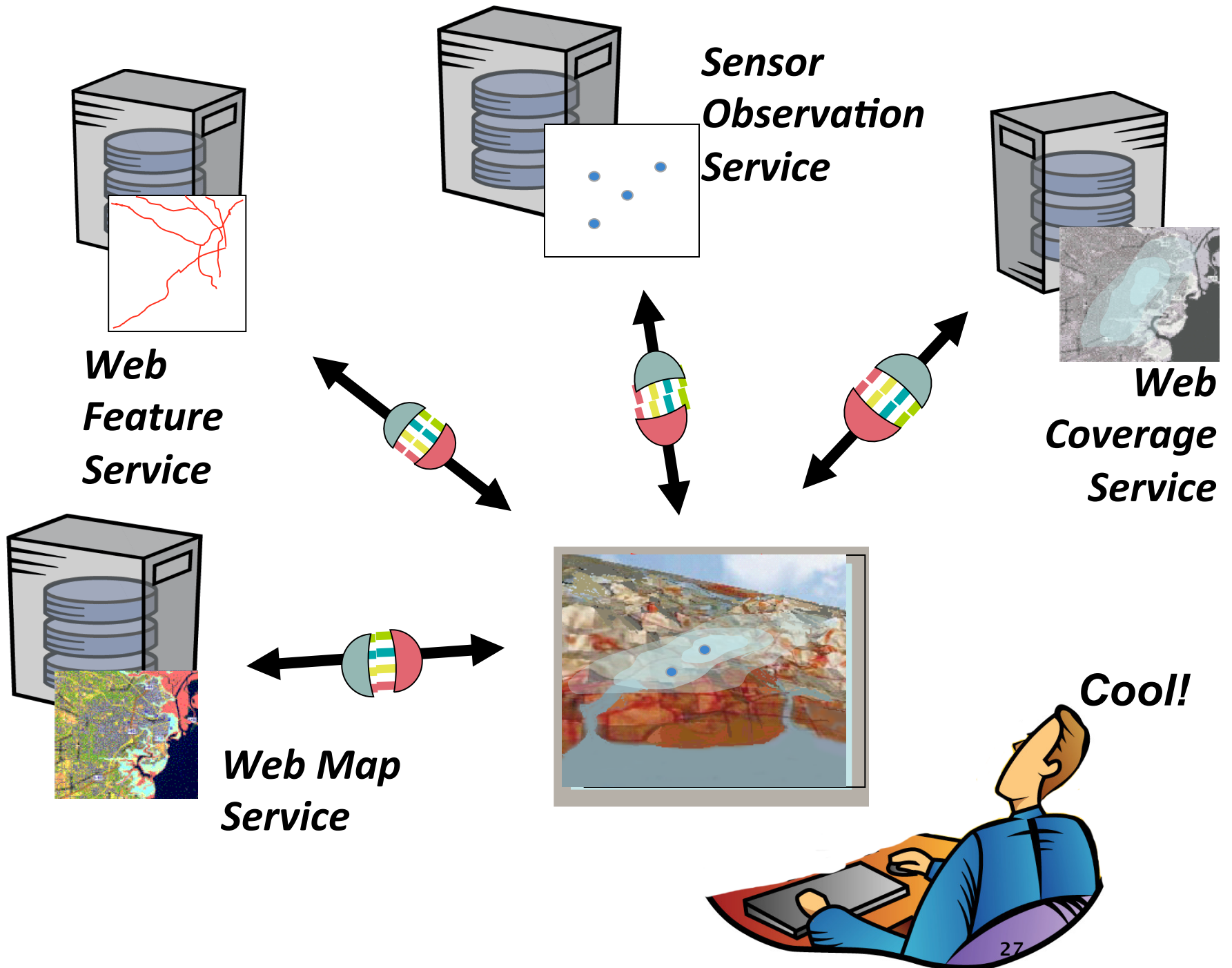
**I have not  
failed, I've  
just found  
10,000 ways  
that won't  
work.**

**Thomas Edison**

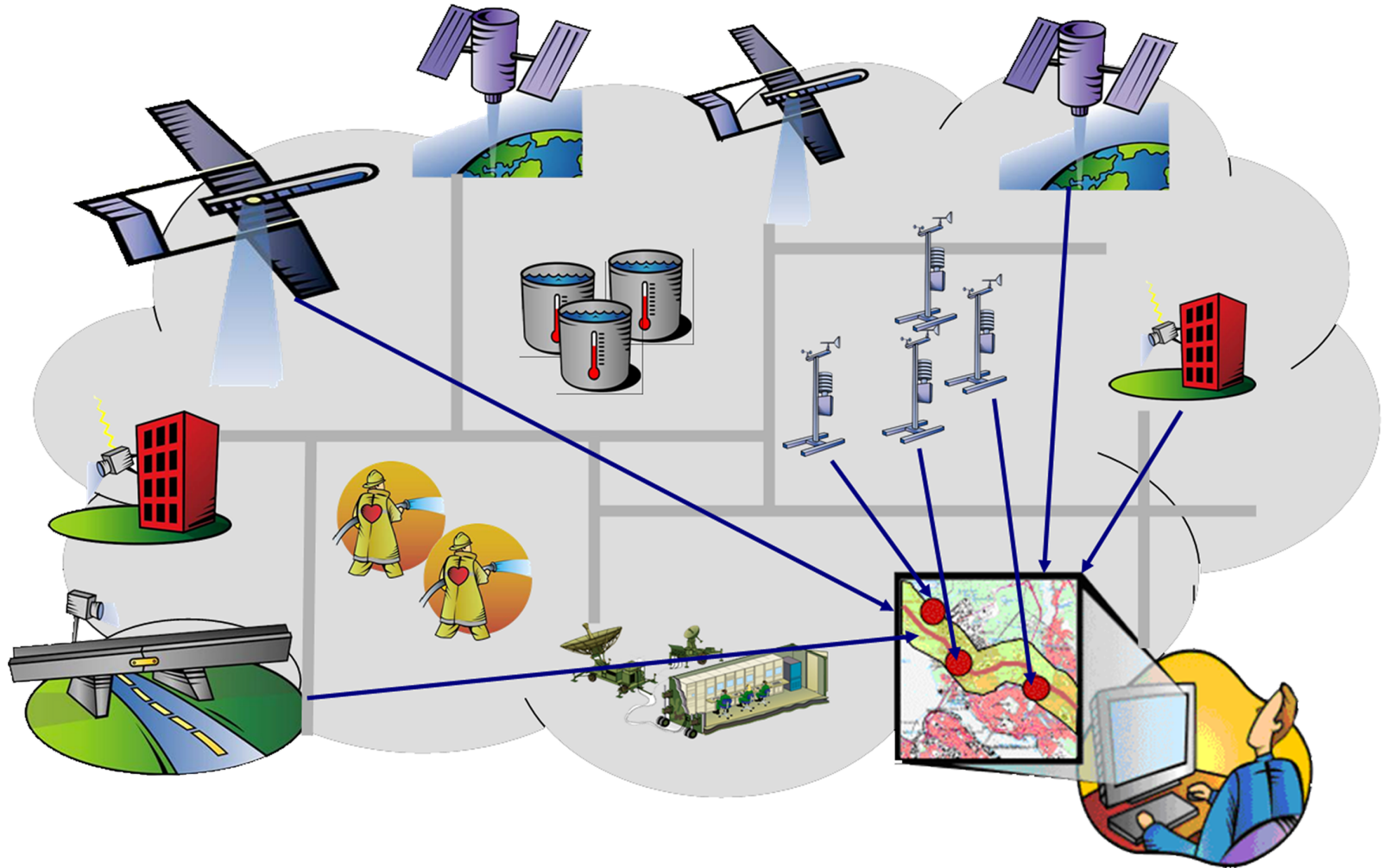
# Compliance Program

700 implementing products

190 compliant products in the  
market

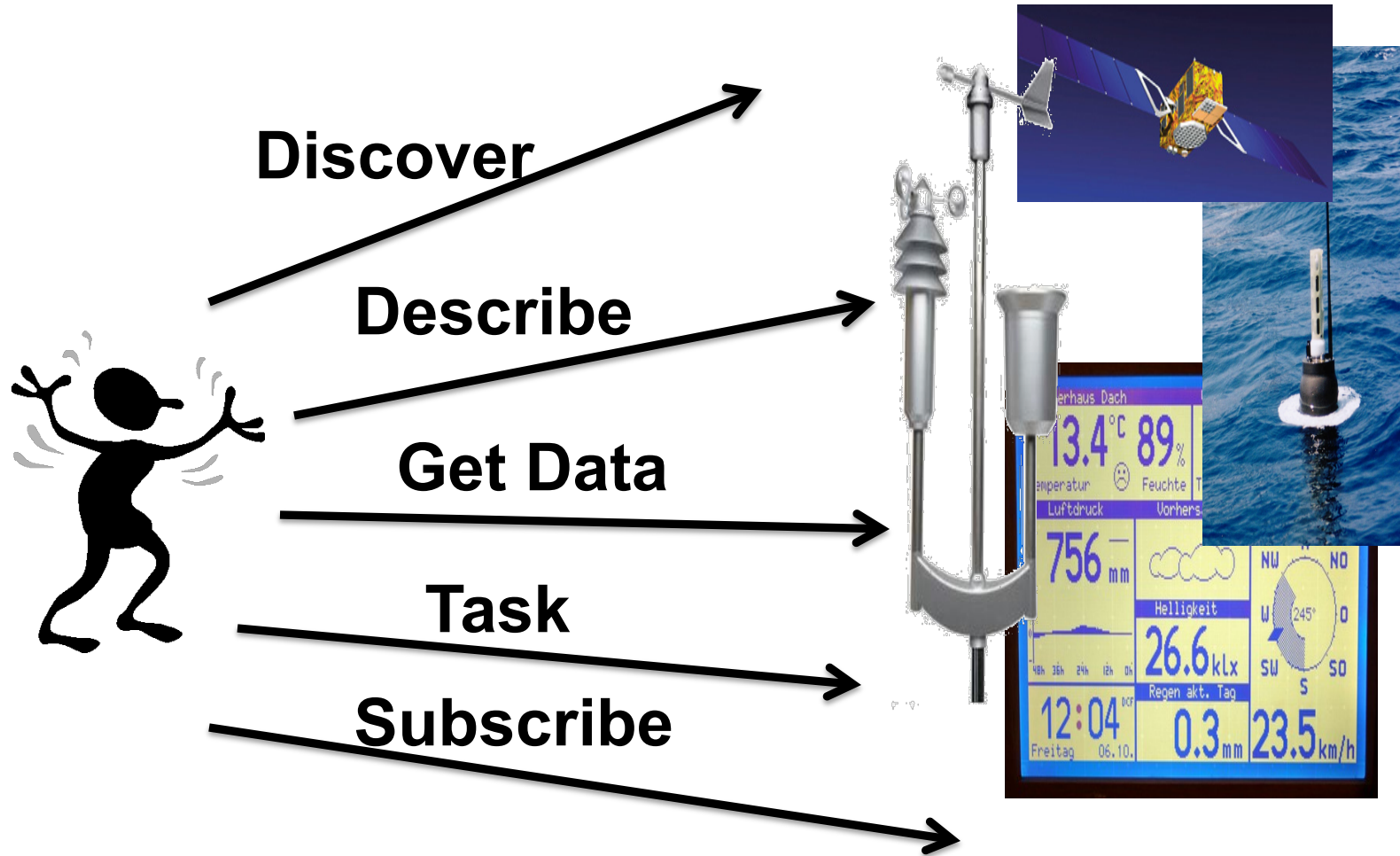


# Sensor Web Enablement

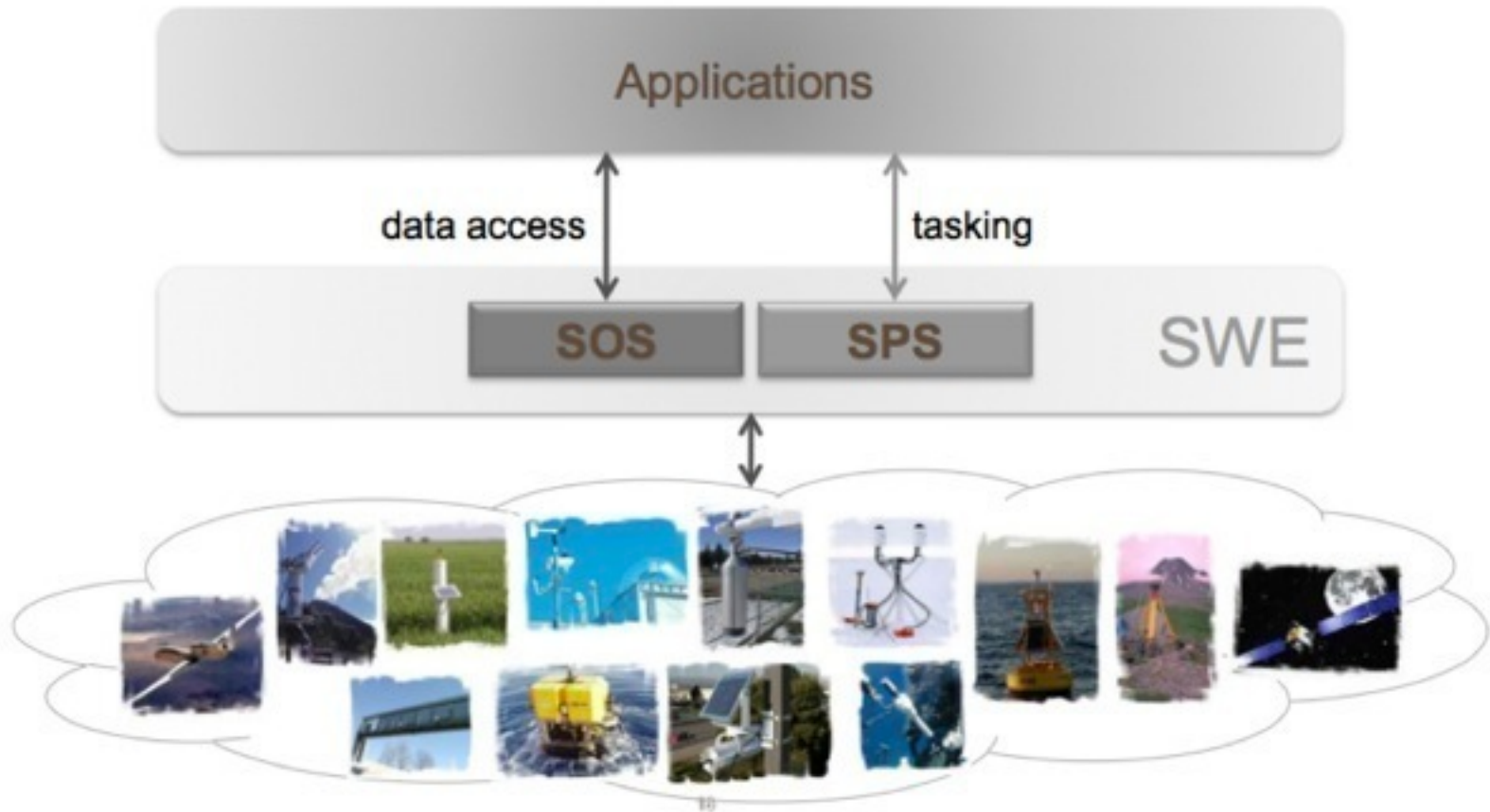




# Sensor Web Requirements



# SWE Solution



# SWE Interface Standards

- Sensor Observation Service
- Sensor Planning Service
- Sensor Alert Service
- PUCK

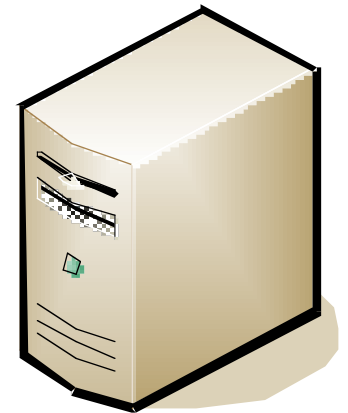
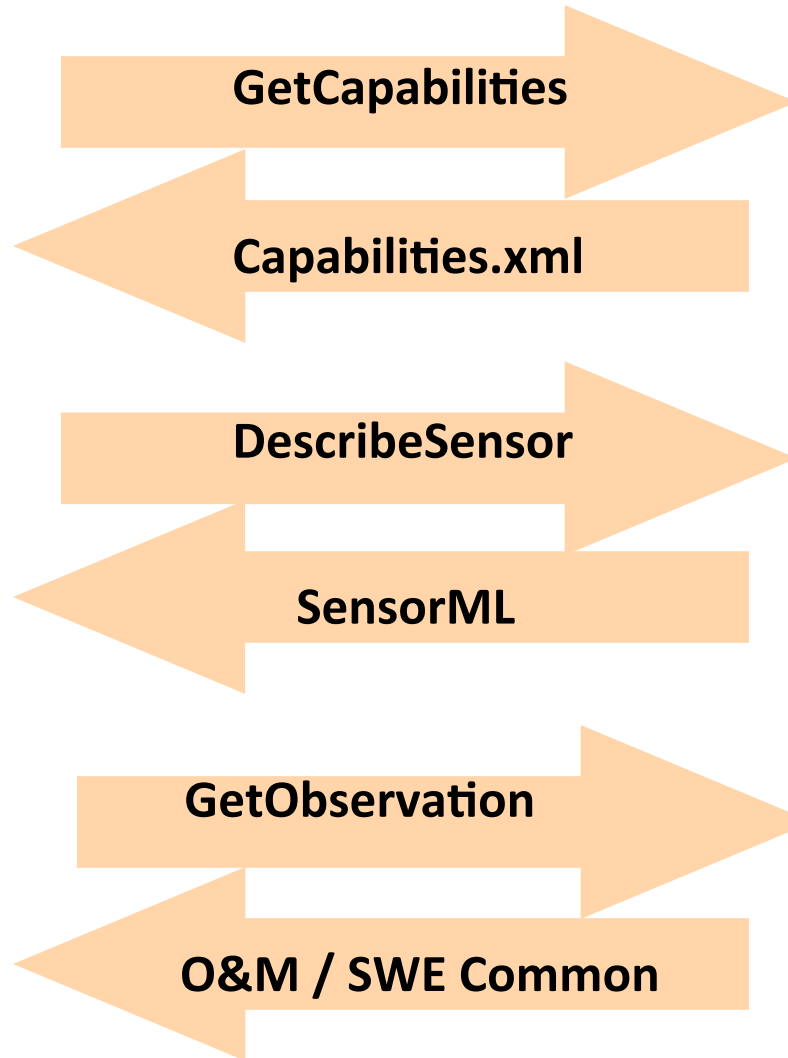
# SWE Encoding Standards

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

# Sensor Observation Service



**Client**



**SOS**

# 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

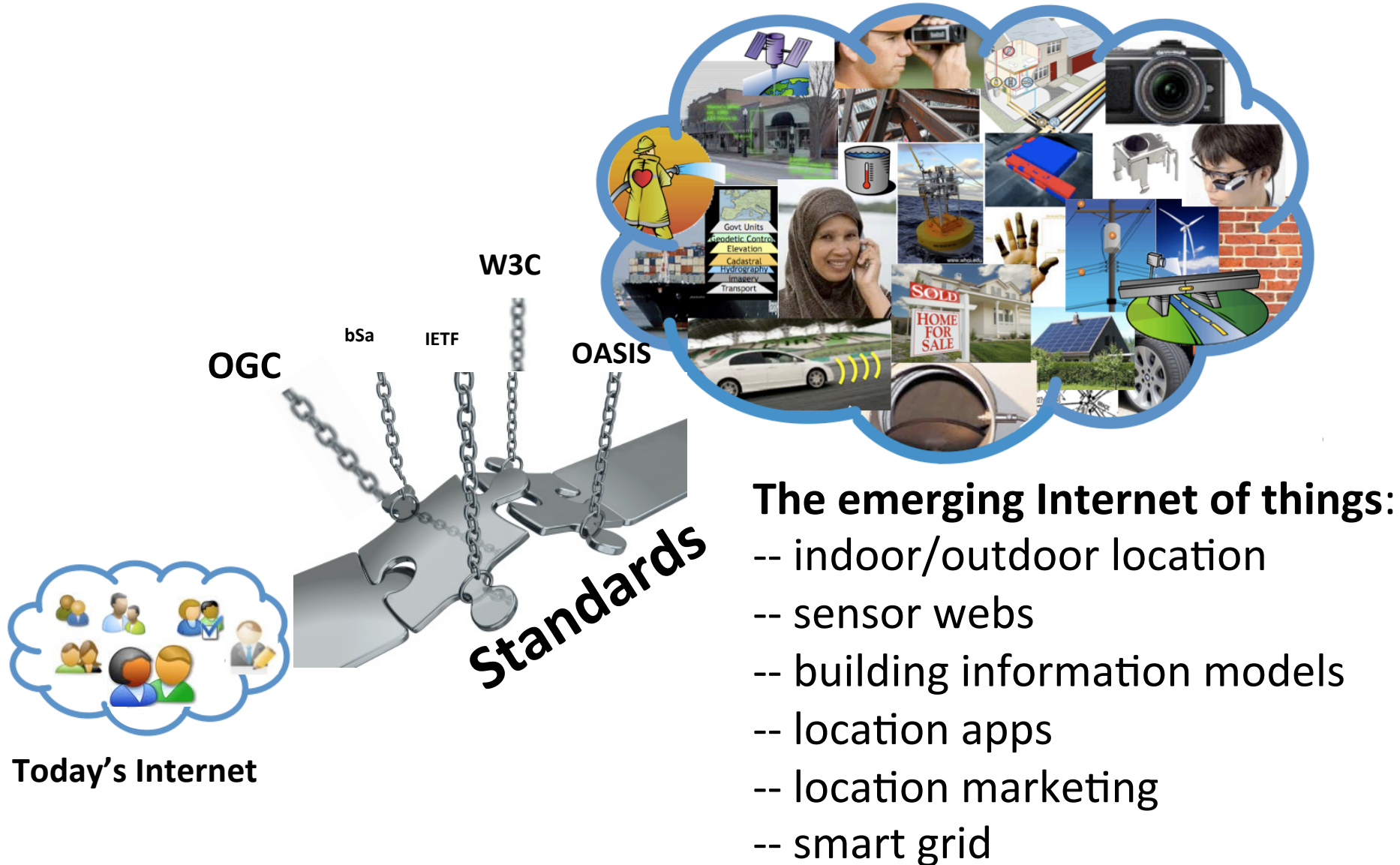
# SWE Harmonization

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

PULSENet demonstrated the use of SensorML to describe these sensors



# Internet of Things



# 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 S	Perey, Christine
AIT Austrian Institute of Technology GmbH	Prodevelop, S.L.
Arizona State University (ASU)	Terradue Srl
Botts Innovative Research	The SI Organization, Inc.
Cosm Ltd	Universitat Politècnica de Catalunya, UPC (SARTI research group)
CSIRO	University of Calgary
EADS ASTRIUM	University of California, San Diego Supercomputer Center
EDINA, University of Edinburgh	University of Muenster - Institute for Geoinformatics
Envitia Ltd.	University of the Bundeswehr - ITIS
Esri	University of Tokyo CSIS
Esri Canada	US National Geospatial-Intelligence Agency (NGA)
ETRI (Electronics & Telecommun	WiSC Enterprises
European Centre for Medium-Ra	Woolpert
Federal Agency for Cartography	
Fraunhofer-Gesellschaft	
GEOMATYS	
Geonovum	
Hankuk Universit	
Hitachi, Ltd., Cen	
Industrial Techno	
Institut National	
interactive instru	
Intergraph Corpor	
Lockheed Martin	
METEO-FRANCE	
National Institute of Water and Atmospheric Research	
NAVTEQ	
Open Geospatial Consortium, Inc.	
Open Site Plan	
Oracle USA	

# 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

1935 Platforms 100 Rectangles.

Variables:

☐ Cluster platforms

Click the dots for in-situ observations.  
Click the rectangles for gridded data.

☒ Recent observations w/in:

☐ No observations

Start:

End:

Time is UTC. Start: 00:00 End: 23:59

Regions:

☐ Search by bounding box mode.

(Click a gridded data rectangle to filter platforms.)

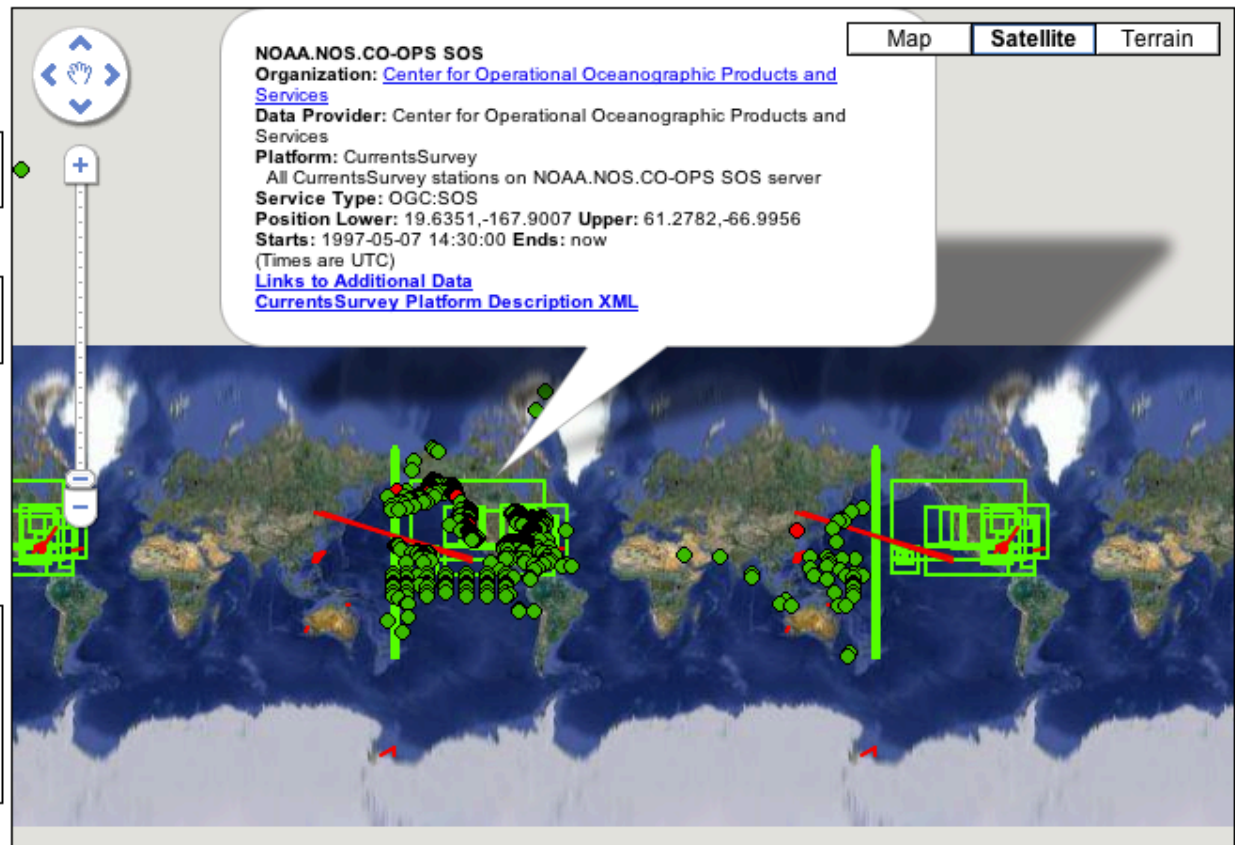
Service types:

Servers:

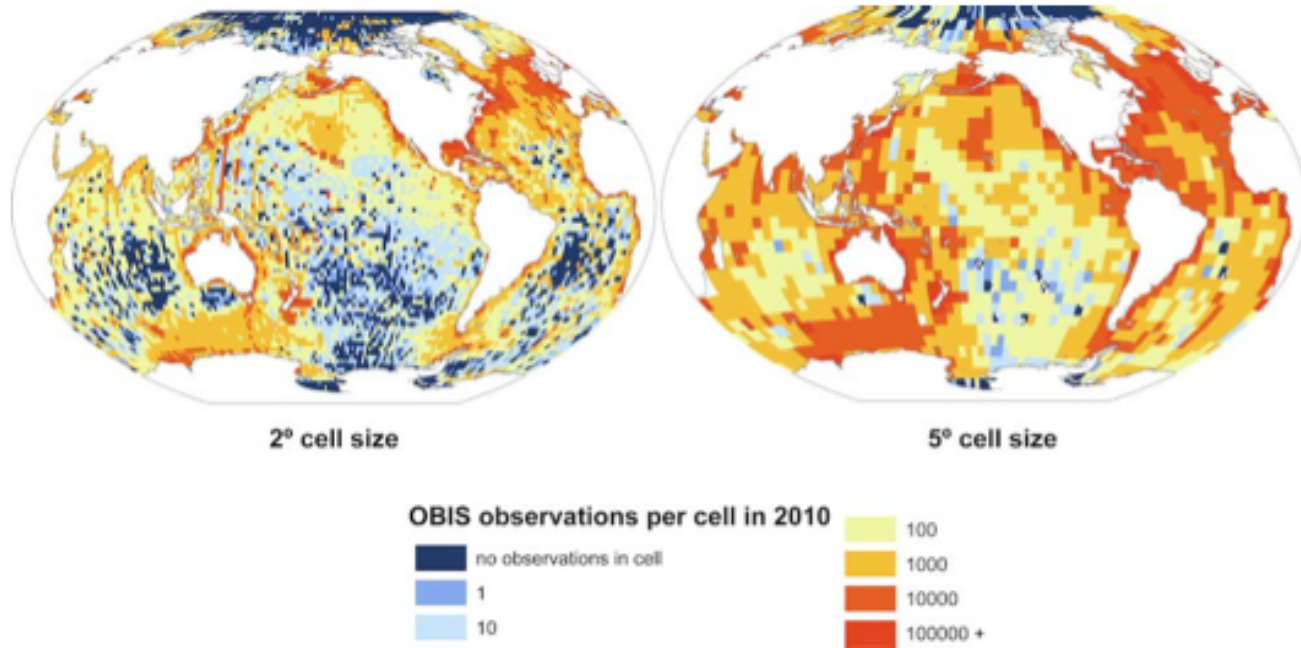
Data Providers on this Server

All

Data Providers:



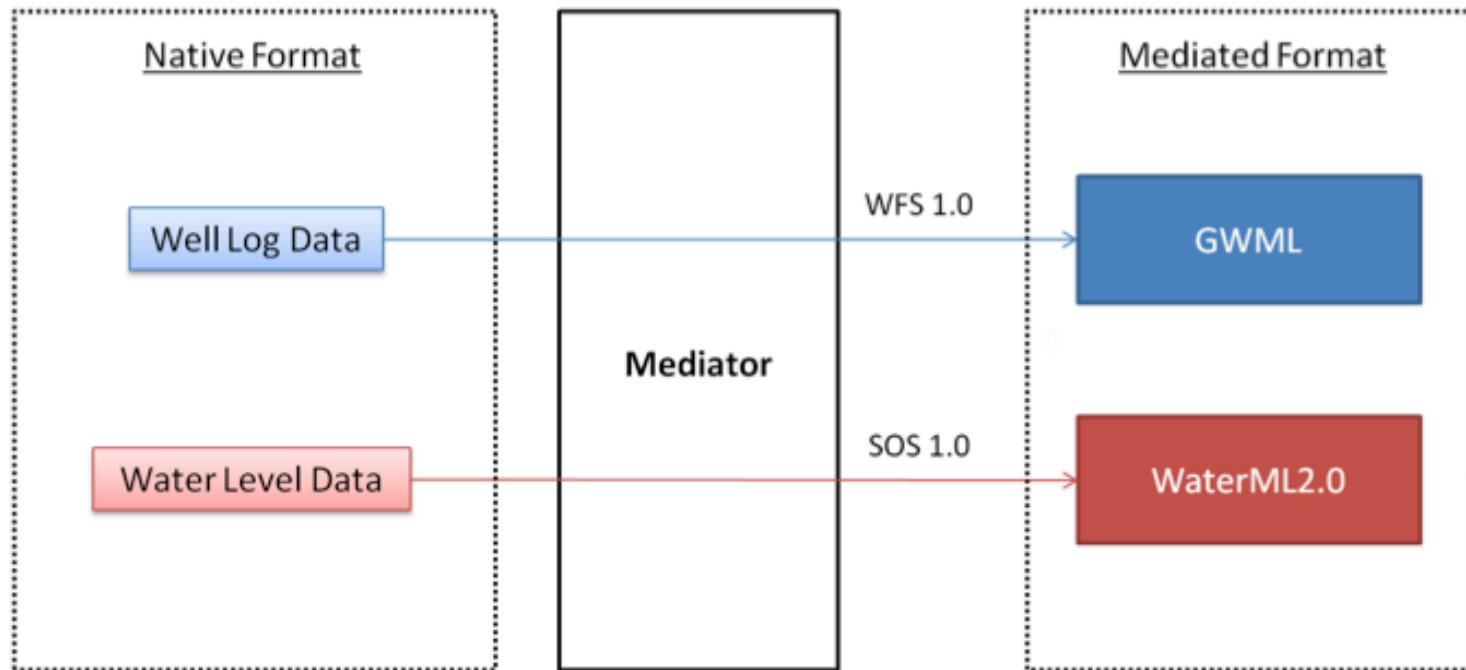
# OBIS – 31 Million Observations



Uses OGC Services to publish maps and download data



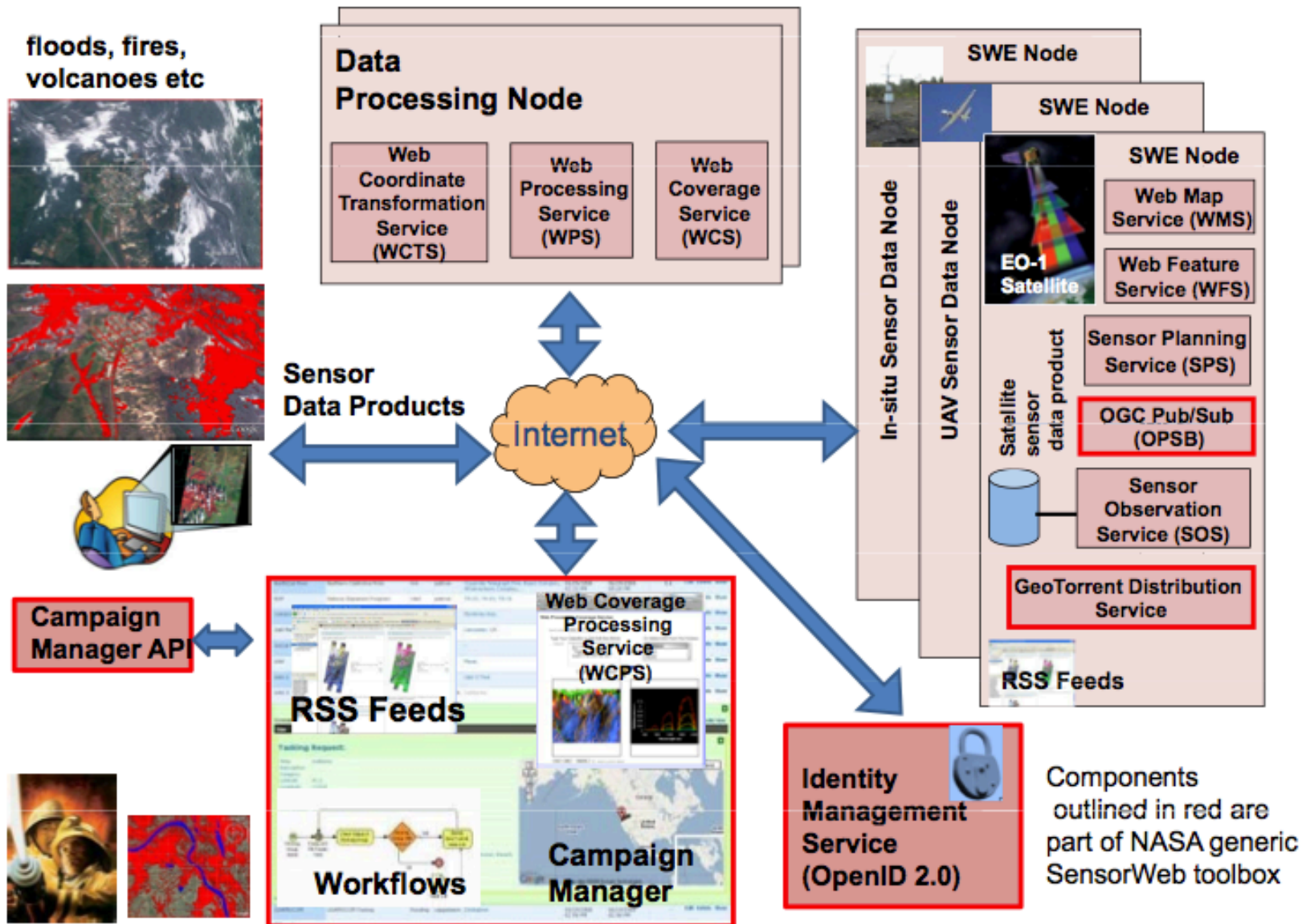
# Groundwater Interoperability



2 million Wells in Canada

# Velocity

# NASA Sensor Web





# GeoBliki EO1

***Sensor Web Enabled (SWE) Data Node***

## ➔ Data Distribution

### ➔ Atom 1.0

**Atom Enabled**

### ➔ GeoRss

**GeoRss with**

### ➔ Publish/Subscribe

**OGC Site**

### ➔ OGC Pubsub Service – Basic (OPS-B)

#### ➔ OGC PubSub page

#### ➔ OGC OPS-B Page

## OGC Web Feature Server – Basic (WFS-B)

### ➔ OGC Site

### ➔ Examples

## OGC Sensor Observation Server(SOS)

**GeoBliki SOS demo**

### ➔ Sensor Tasking

## OGC Sensor Planning Service Specification

**GeoBliki SPS Demo**

**GeoBliki SPS WSDL**

**GeoBliki SPS Invoke**

# Open GeoSMS Standard

[http://maps.google.com/maps?  
q=38.9985,-77.030275&GeoSMS](http://maps.google.com/maps?q=38.9985,-77.030275&GeoSMS)

I am here ..





# Ushahidi

[HOME](#) [CONTACT US](#) [ABOUT](#) [USHAHIDI](#) [SWIFTRIVER](#)

Categories: [2007 Elections](#) [Academic Awards](#) [Badges](#) [Checkins](#) [Code Releases](#) [Communi](#)  
[Crowdmap](#) [CrowdmapID](#) [Crowdsorcerers](#) [Crowdsourcing](#) [Data](#) [Deployment](#) [Design](#) [Devel](#)  
[Elections](#) [Evaluation](#) [Events](#) [How to Help](#) [Jobs](#) [Kenya](#) [Localization](#) [Mali](#) [Mapping Resou](#)  
[News](#) [Partnerships](#) [Peace efforts](#) [Plight of the Displaced](#) [Podcasts](#) [Product](#) [Random Thought](#)  
[Research](#) [Security](#) [Services](#) [Strategy](#) [SwiftRiver](#) [Team](#) [Testing](#) [Uchaguzi](#) [Uncategorized](#)  
[Video](#) [Violence](#)

## How OGC's Open GeoSMS serves for Disaster Management

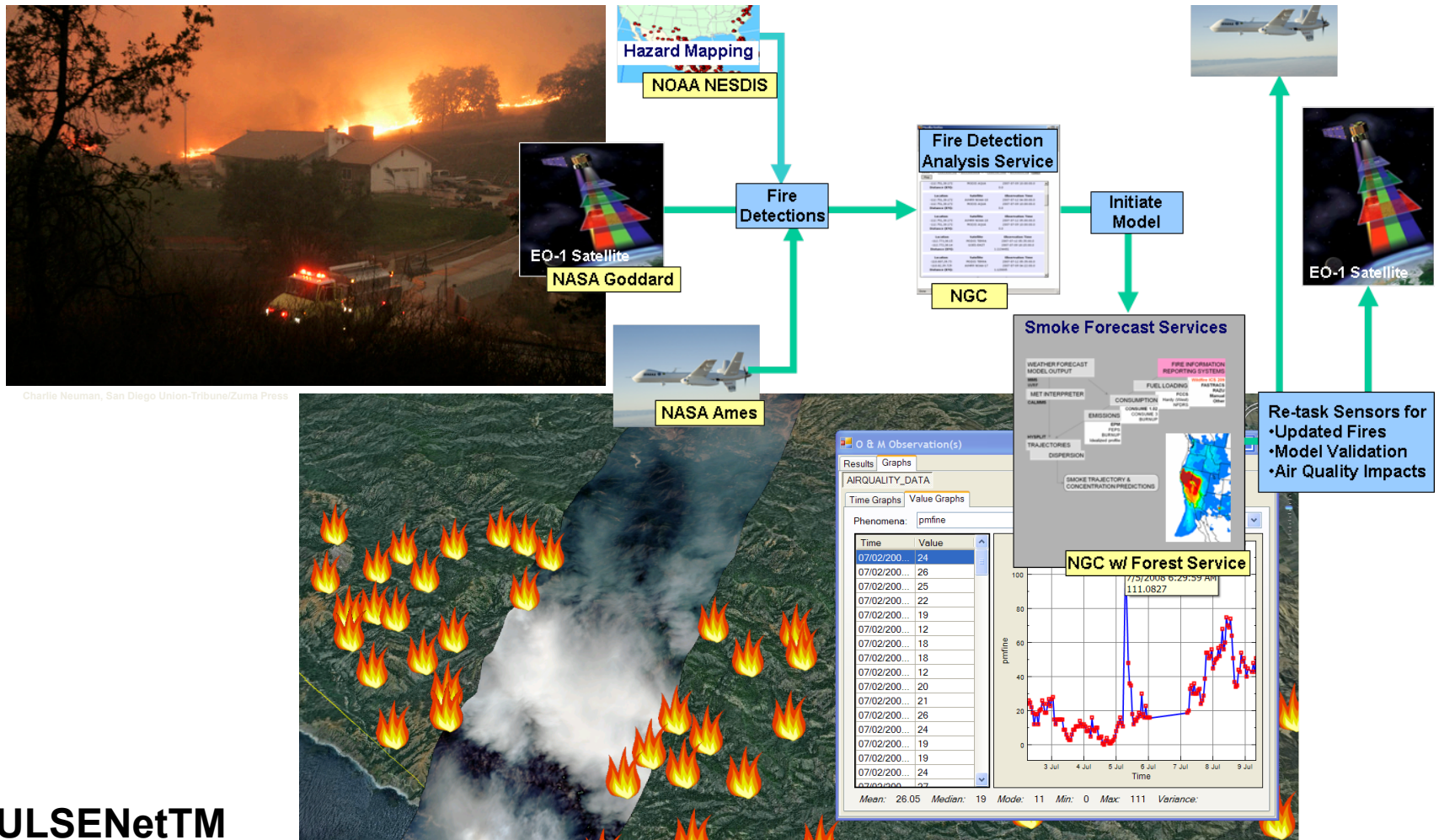
*[Guest blog post by Kuo-Yu slayer Chuang from Tawain's Industrial Technology Research Institute, ITRI.]*

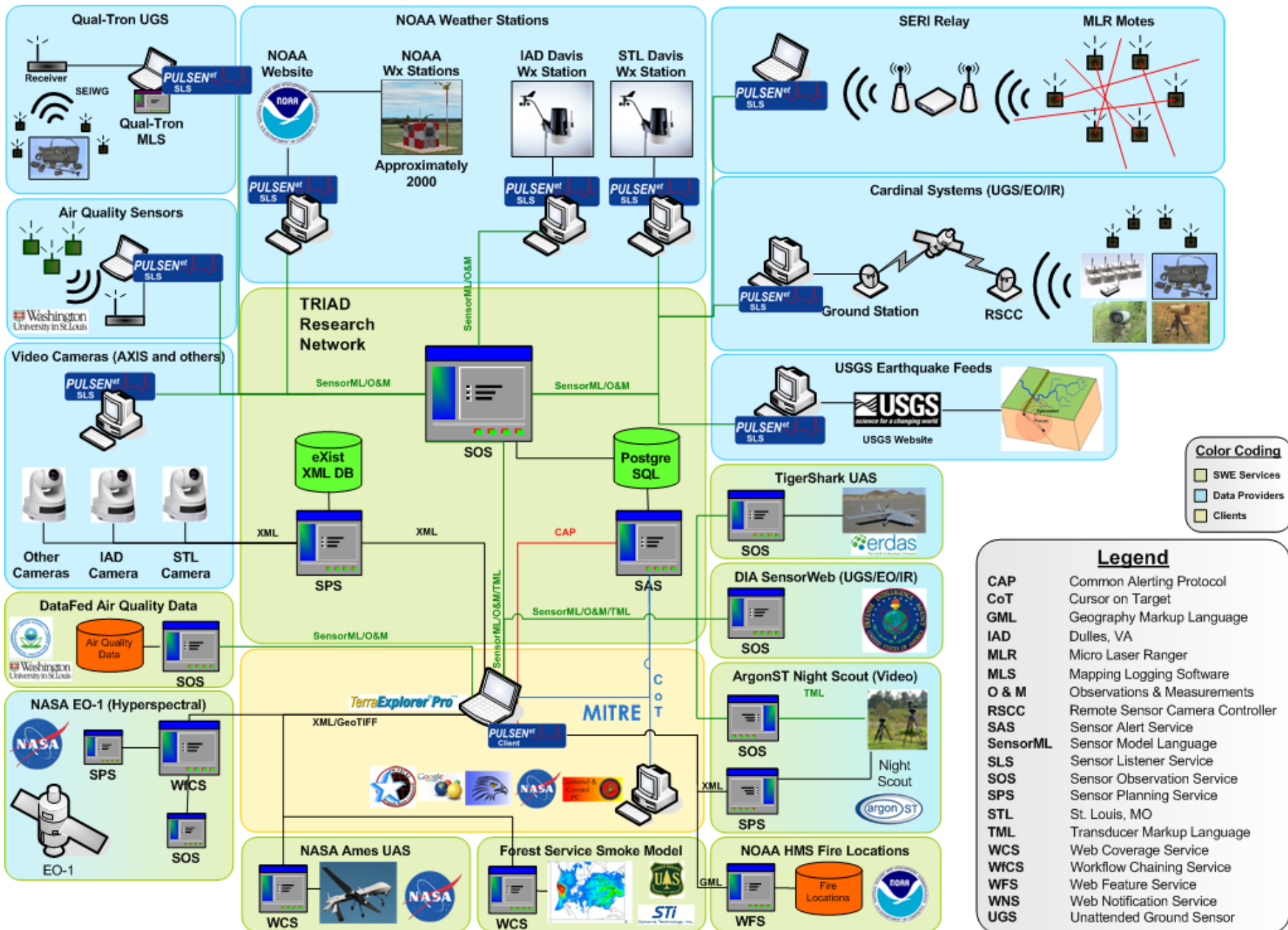
I'm Kuo-Yu slayer Chuang from ITRI, a government funded research institute in Taiwan. We developed an open standard for exchanging location information via SMS among mobile devices called Open GeoSMS, which has currently been adapted officially by Open Geospatial Consortium (OGC). Several use cases of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



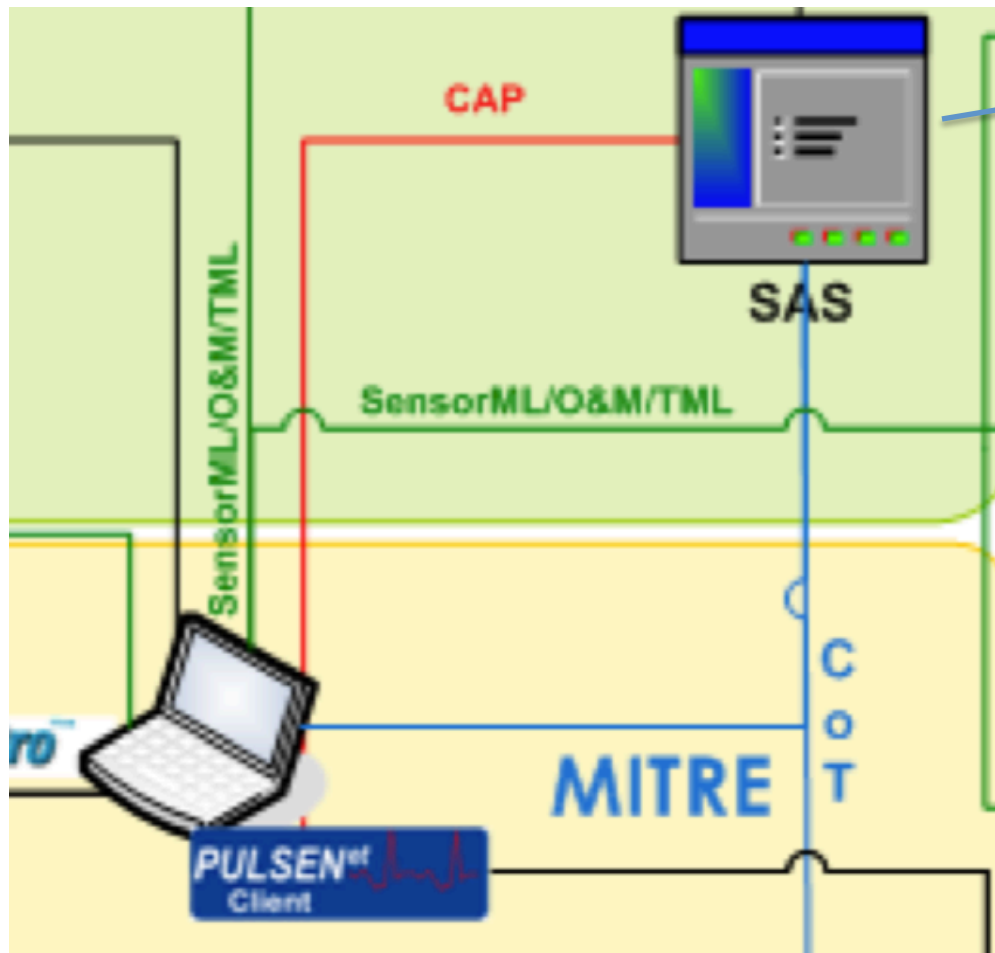
# Variety

# Air Quality Sensors and Modeling

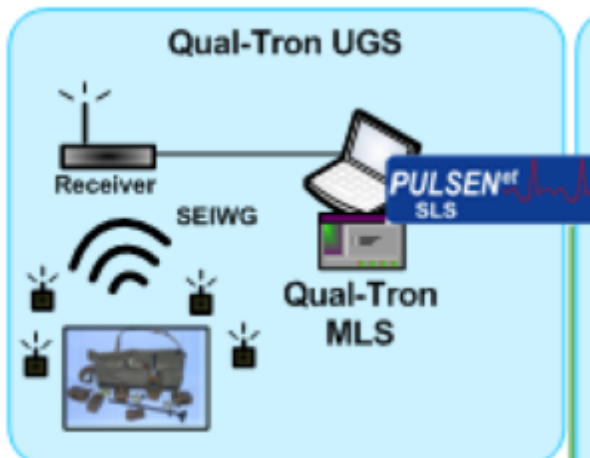




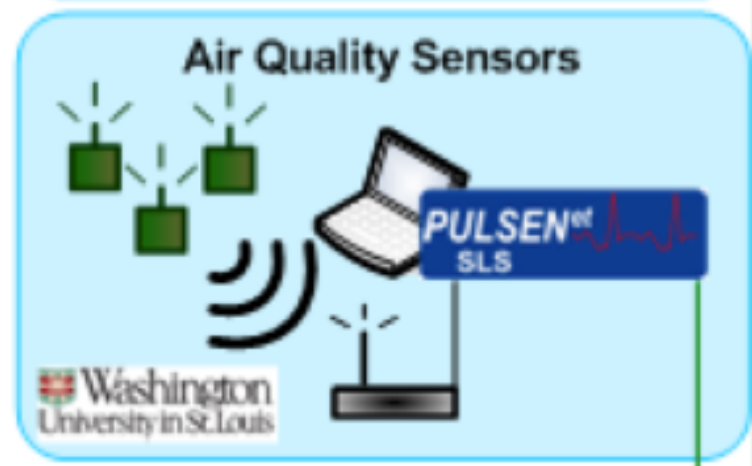
# CAP Alerts via OGC Sensor Alert Service (SAS)



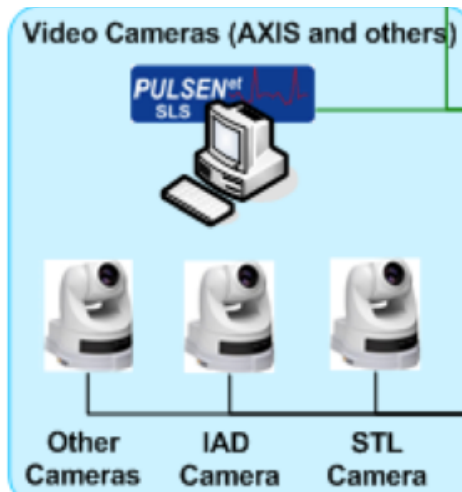
**OGC  
SWE  
Clients**



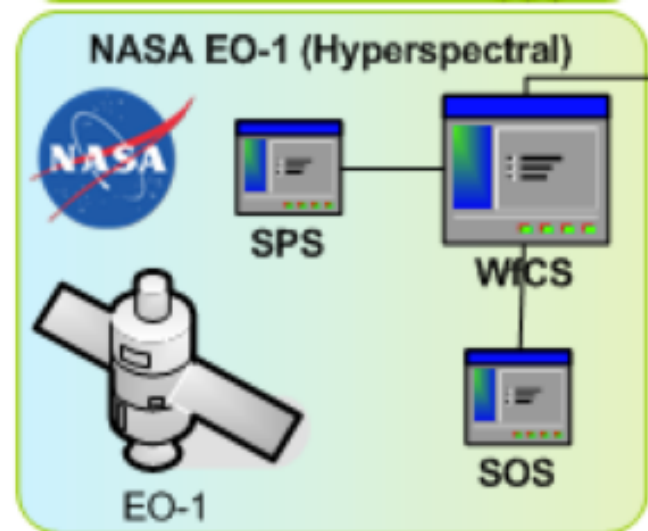
**Seismic**



**Air Quality**



**Videos**



**Satellite**



**Models**

**Unmanned  
Aircraft Systems**



# Veracity Example

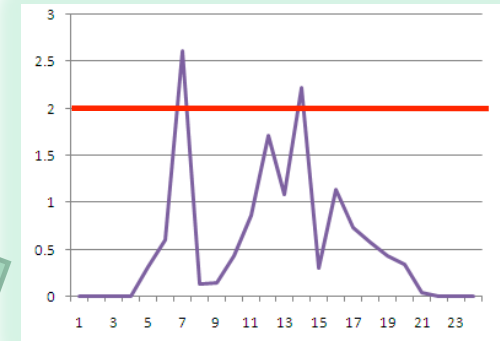
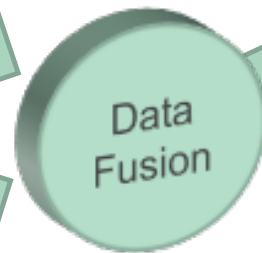
# Sensor Fusion



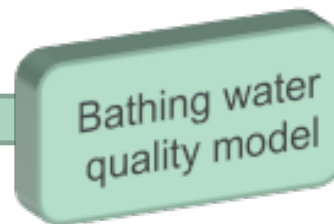
Beach attendance / day



Microbial contamination / day



Impact assessment  
>> close beach? <<

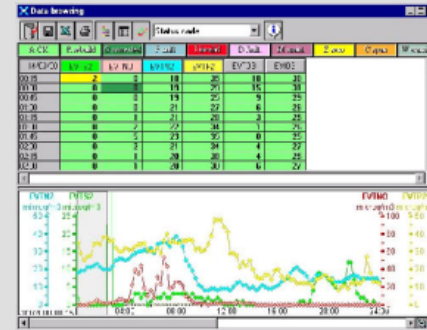


## Air Quality monitoring



## Air quality management

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



## 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

## QARTOD Quality Control Tests : Waves

The participants in this effort (shown on the table tabs) 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.

TIME SERIES (Raw Calibrated Data)				
Category	Criteria	Order	Flag	Action
Data Gaps	Consecutive N missing data. Maximum number of missing data.	1	Soft	N is user defined. Include in % count.
Spikes	User defined Points >= M*std with P iterations	2	Soft	Interpolate/extrapolate up to N points. N is user defined. M can be user defined, recommended M=4. Include in % count.
Range test	Location, instrument defined.	2	1. Soft 2. Hard	Max/min user defined. 1. Interpolate/extrapolate up to n points. N is user defined. Include in % count. 2. Instrument spec exceeded, reject.
Mean shift (segments)	A mean shift "P" occurs in this time series.	3	Hard	Reject entire record. P is user defined.
Acceleration test	User defined ( $a > M \cdot g$ )	3	Soft	Recommended $M \leq 1/2$ . Interpolate/extrapolate up to N contiguous points. N is user defined. Include in % count.
Mean test, variance test	User defined, location dependent	4	1. Soft 2. Hard	1. Flag unexpected values. 2. Reject unreasonable values.
Percent points good	Check for M% good data (based on above 6 criteria)	5	Hard	Recommended $M \geq 90\%$
SPECTRAL VALUES				
Category	Criteria	Order	Flag	Action
NON-DIRECTIONAL:				
Operational frequency range test	^defined by the environment and instrument	1	1. Soft 2. Hard	1. Max/min user defined. 2. Instrument spec exceeded, reject.
DIRECTIONAL:				
Incident low frequency energy direction	Location defined	1	Soft	User defined

# QUARTOD

## Quality Tests for Marine Sensors and DATA

Observable Properties

```
graph TD; A[Observable Properties] --> B; subgraph B [ ]; direction TB; B1[Sensor/Deployment Files (SensorML)] --> B2[Process Files (SensorML)]; end; B --> C[Observed and Derived Properties and QC Flags];
```

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

Bit	Name	Description
01	ATMPAIL	Atmospheric correction failure
02	LAND	Pixel is over land
03	BADANC	Reduced quality of ancillary data
04	HIGLINT	High sun glint
05	HIST	Observed radiance very high or saturated
06	HISATZEN	High sensor view zenith angle
07	COASTZ	Pixel is in shallow water
08	NEGLW	Negative water-leaving radiance retrieved
09	STRAYLIGHT	Straylight contamination is likely
10	CLOICE	Probable cloud or ice contamination
11	COCCOLITH	Coccolithophores detected
12	TURBIDW	Turbid water detected
13	HISOLZEN	High solar zenith
14	HITAU	High aerosol optical thickness
15	LOWLW	Very low water-leaving radiance (cloud shadow)

## Veracity



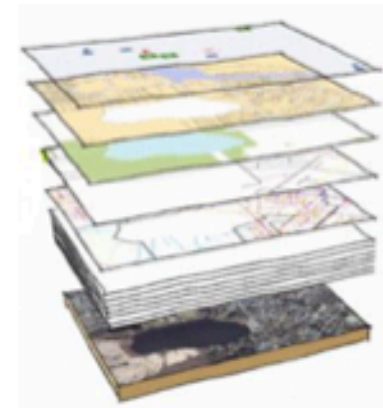
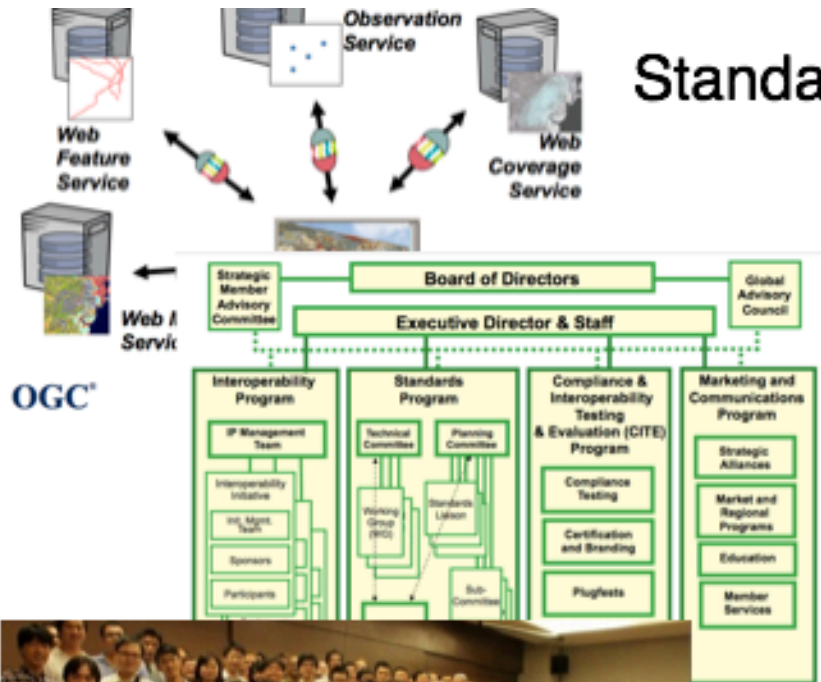
## Variety



## Volume



# Conclusion - OGC



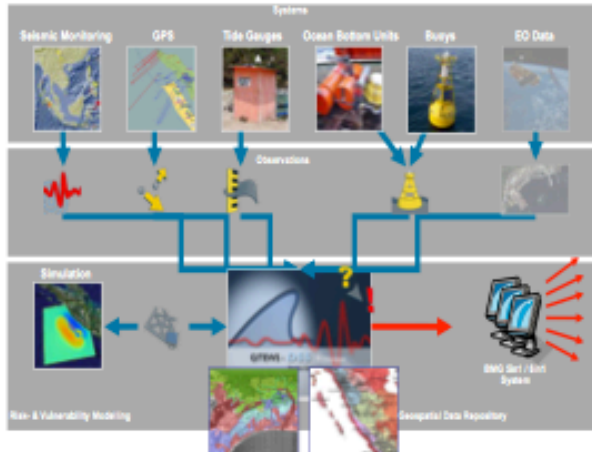
Process  
Infrastructure  
Community

Integrate  
Geospatial  
Data

# Conclusions - Sensor Web Enablement



# Conclusion - Successful Examples



Tsunami Warnings



Air Quality



Mobile



Quality Marine Sensors



Sensor Data Fusion

# Questions?

Luis Bermudez, Ph.D.

[lbermudez@opengeospatial.org](mailto:lbermudez@opengeospatial.org)

@berdez on Twitter