Connecting "Long-tail" Scientists with Big Data Centers using SaaS

AGU 2012 Session IN015: Data Interoperability and Inter-Use Solutions

George Percivall and Luis Bermudez
Open Geospatial Consortium (OGC)

December 5, 2012

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

“Long-tail” Scientists and Big Data Centers

- Big data centers and long tail scientists –

  two extremes in the geoscience research community

- Software-as-a-service (SaaS) gives access to big data and processing to this underserved community of scientists

- Good examples today as basis for building interoperability and inter-use through web services to big data centers

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

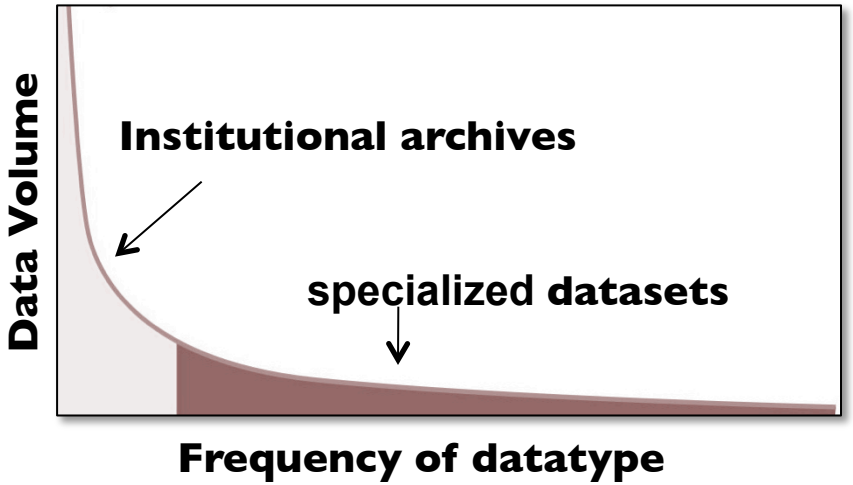
Large, institutional data centers

- Vital resources in geoscience community and will be for some time to come
- Provide “access to the data and a critical source of people who have experience in the use of the data." [NRC]
- …and will continue to grow [NASA ITIC Smarr]

	Big Data CI	10G→100G Network	GPU Clusters	Hybrid HPC
Non-Fed	Google/MS/ Amazon	GLIF/I2/ CENIC	Japan TSUBAME2 4224 GPUs 2.4 PF	China #2 Fastest 5 PF MC/GPU
NSF	Gordon	GENI Next Gen Internet	TAAC 512 GPUs	Blue Waters* MC/GPU 12 PF
DOE	Magellan	ANI ARRA 100Gb	ANL 256 GPUs	NG Jaguar* MC/GPU 20 PF
NASA	Nebula, Testbed	Goddard to Ames 10G	Ames 136 GPU 2x64 at Ames & GFSC	Pleiades MC 1PF

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

The “long-tail of science”



- Distribution of NSF Awards by Dollar Value [Heidorn]
- 2% of largest grants received 20% of total amount awarded.

- NSF EarthCube analysis of long-tail scientist use cases
  - Work with limited resources
  - Simple tasks become discovery and management nightmares
  - Downloading and transforming datasets takes weeks [Foster, et.al.]

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

Web Services Response to Long-Tail

- Comparison to long-tail in Web brings methods to bridge gap between large data centers and long-tail scientist
- Nearly a decade ago, OGC envisioned a greatly expanded geospatial information space would result from a platform of open web-based standards for geoprocessing
- SaaS for the geospatial web, e.g.
  - OGC Web Map Service [WMS]
  - OGC Web Coverage Service [WCS]
  - Data Access Protocol [DAP]
  - OGC Web Processing Service [WPS]

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

OGC Web Services Standards

The GeoWeb is enabled by OGC standards:

Web Map Service (WMS)

Web Feature Service (WFS)

Web Coverage Service (WCS)

Catalogue (CSW)

Geography Markup Language (GML)

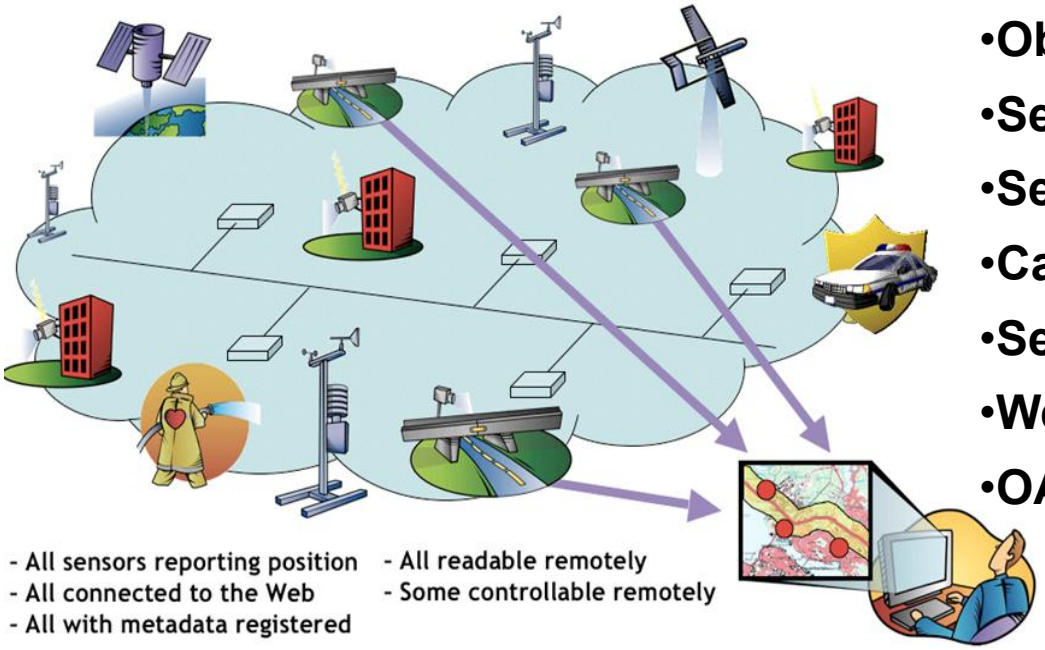
OGC KML

Others…

OGC Copyright 2012, Open Geospatial Consortium • *Making Location Count*

OGC Sensor Web Enablement

Discovery and tasking of sensors. Access, fusion and application of sensor observations for enhanced situational awareness



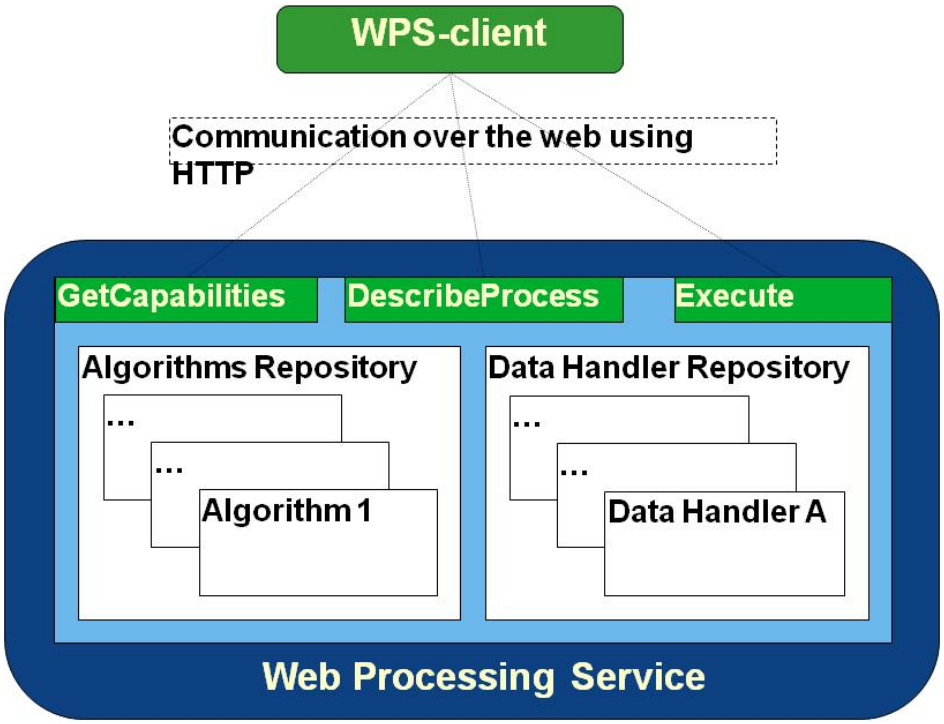
- **Sensor Model Language (SensorML)**
- **Transducer Markup Language (TML)**
- **Observations & Measurements (O&M)**
- **Sensor Planning Service (SPS)**
- **Sensor Observation Service (SOS)**
- **Catalogue Service**
- **Sensor Alert Service (SAS)**
- **Web Notification Service (WNS)**
- **OASIS (alert) standards**

OGC Copyright 2012, Open Geospatial Consortium • *Making Location Count*

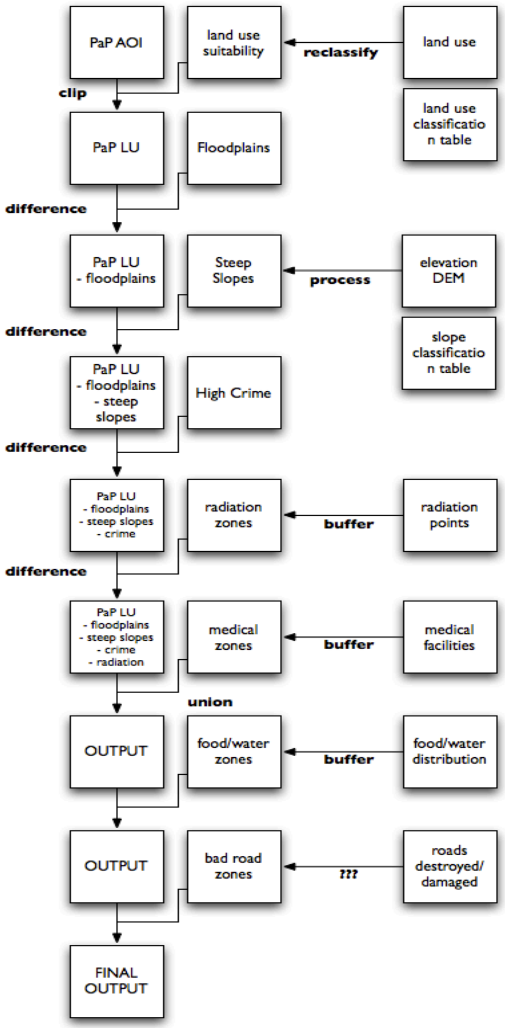
Geospatial Processing, Analysis, Workflow

Web Processing Service – WPS

- OGC Web Service access to algorithms
- Change detection, coordinate transformation, modeling and simulation…



Geoprocessing Workflow



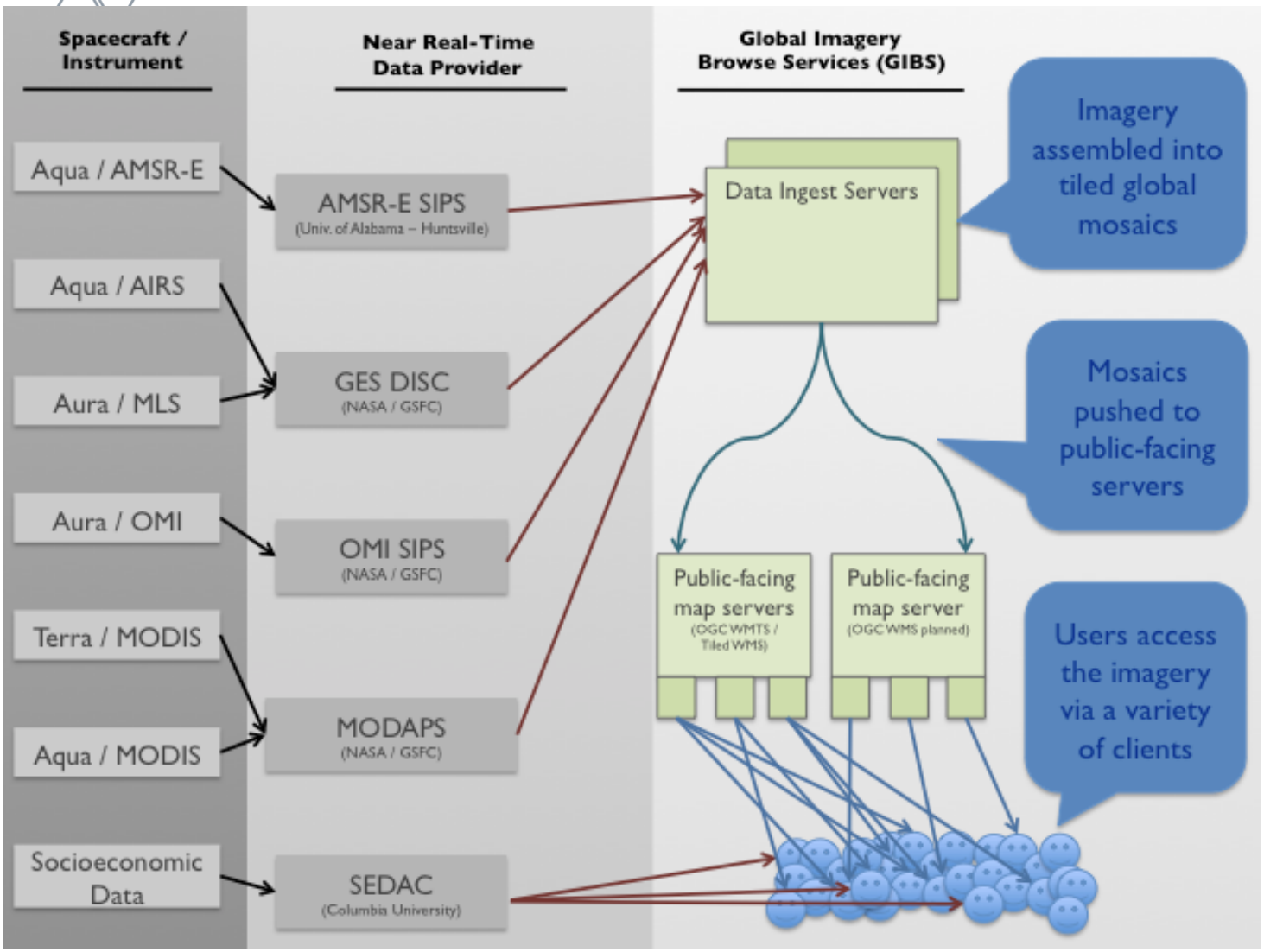
OGC Copyright 2012, Open Geospatial Consortium • *Making Location Count*

Connecting to Data Centers with SaaS

- “By adopting, adapting, and applying semantic web and SaaS technologies, we can make the use of geoscience data as easy and convenient as consumption of online media." [Foster, et.al.]
- Interoperability of web services being readily combined with the intuitional resources of large data centers.
- So how about some examples….

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

Example: NASA's EOSDIS GIBS



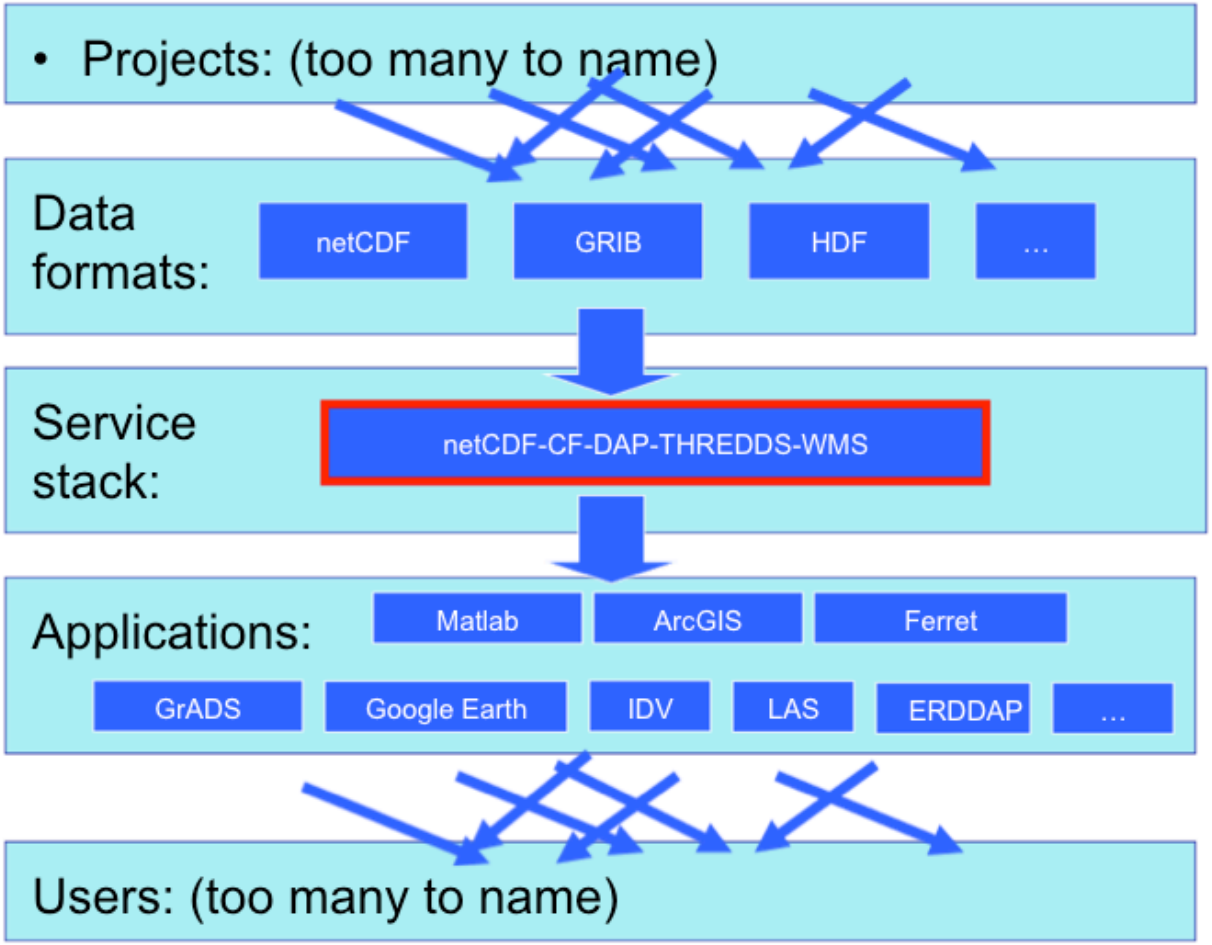
Access Methods
- **OGC WMTS**
- **Tiled WMS**
- **OGC KML**

Supported Clients
- **OpenLayers**
- **Google Earth**
- **NASA World Wind,**
- **Others**

Source: Ryan Boller, NASA
http://earthdata.nasa.gov/wiki/main/index.php/Global_Imagery_Browse_Services

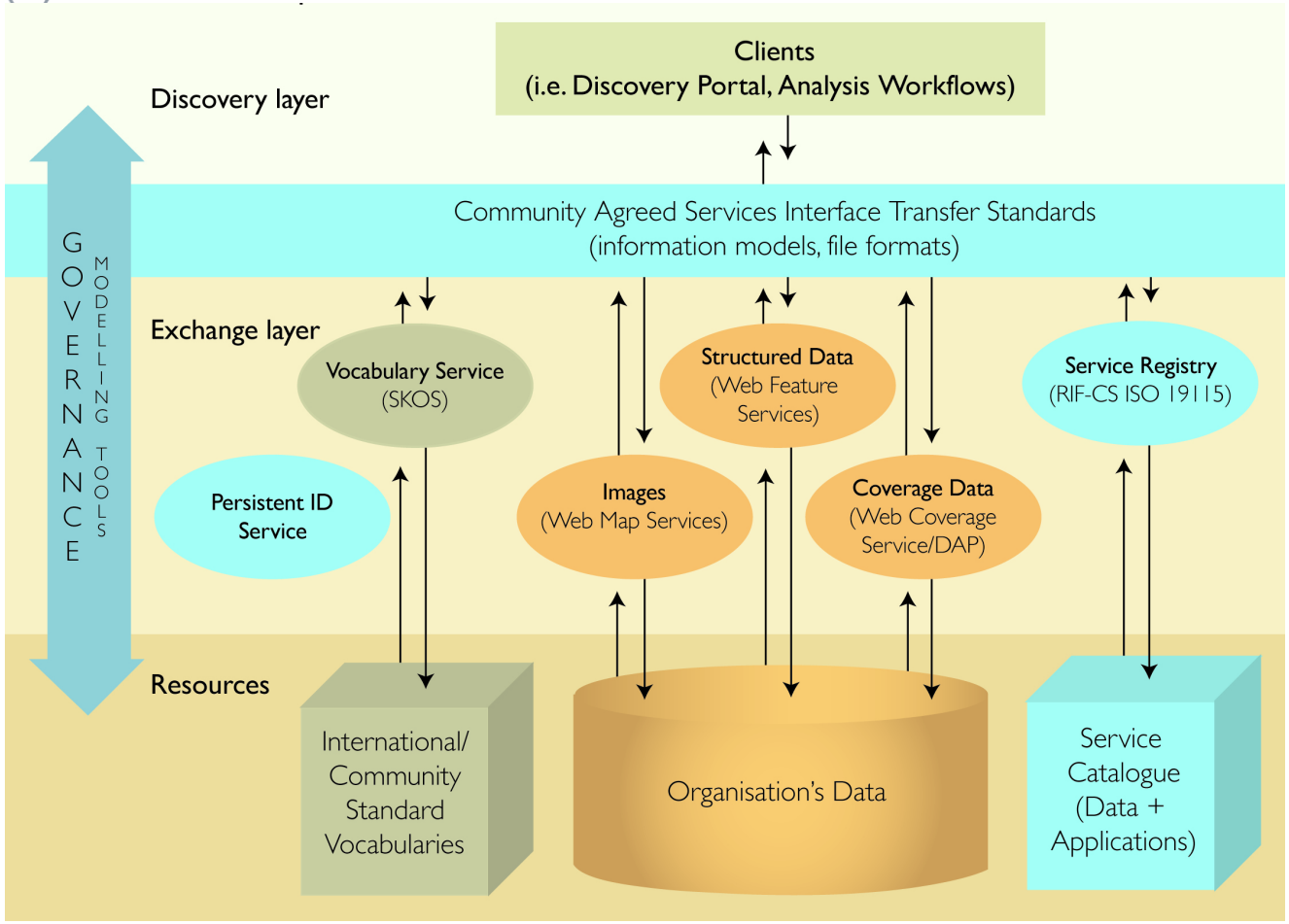
Example: NOAA UAF

- Unified Access Framework for Gridded Data (UAF Grid)
  - web service capability for direct access to a variety of datasets using netCDF, OPeNDAP, THREDDS, OGC WMS and WCS



OGC

AuScope Spatial Information Services Stack



AuScope –
infrastructure for understanding evolution and processes of Australian continent

SISS –
Suite of tools for spatial data interoperability using OGC standards and existing techologies

Cloud Services

DigitalGlobe's Cloud Services unlock the power of the world's largest online library of high-resolution Advanced Ortho imagery. With our extensive suite of OGCcompatible web services combined with our powerful hosting infrastructure, we can instantly deliver current, high-resolution imagery and geospatial information to desktops, portals, intranets and mobile devices around the world.

Delivering the World

With our Advanced Ortho Content Programs, our powerful

Easier Integration

Satellite imagery and metadata are accessible through a full

Resources

Supporting Documents

Cloud Services
Overview of DigitalGlobe's Cloud Services 3.0
> Download

Web Services Available

Web Feature Service (WFS)
- Query against metadata
- Real-time coverage display

Web Map Service (WMS)
- Natively interoperable with most GIS Software
- Images generated upon request per specification

Web Map Tile Service (WMTS)
- Image tiles delivered rapidly
- Renders rapidly for panning and zooming

Web Coverage Service (WCS)
- Easily download GeoTIFFs and other formats
- Perform multispectral analysis and visualization

Connecting "Long-tail" Scientists with Big Data Centers using SaaS

- Move important operations into the cloud providing long-tail scientists access with SaaS web services to make the operations invisible

- *"Civilization advances by extending the number of important operations which we can perform without thinking of them." - Alfred North Whitehead*

- Thanks – George Percivall, percivall@myogc.org

Copyright 2012, Open Geospatial Consortium • *Making Location Count*

References

- [NRC] National Research Council,, "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond," Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future, National Academies Press, 2007. http://www.nap.edu/catalog/11820.html
- [Heidorn] P.B. Heidorn, Shedding Light on the Dark Data in the Long Tail of Science. Library Trends, 57(2):280-299, 2008. https://www.ideals.illinois.edu/handle/2142/10672
- [Foster] Foster, Ian, EarthCube Charrette white paper 2011, "Wagging the long tail of earth science: Why we need an earth science data web, and how to build it"
- [ORM] The OGC Reference Model, V2.1, 2011. http://www.opengeospatial.org/standards/orm
- [WCS] OGC, Web Coverage Service 2.0, OGC Interface Standard, OGC Document 09-110r3 http://www.opengeospatial.org/standards/wcs
- [DAP] James Gallagher, Nathan Potter, Tom Sgouros, Steve Hankin, Glenn Flierl,
- "The Data Access Protocol — DAP 2.0." 2005-10-27. http://opendap.org/pdf/dap_2_data_model.pdf
- [CCI] OGC Cross-Community Interoperability.
- [Hype] http://en.wikipedia.org/wiki/Hype_cycle
- [ECW] Rose-marie Gonzalez, Kevin J Murphy, Matthew F Cechini, "NASA's Earth Data Coherent Web," AGU Fall Meeting 2011, Abstract ID: P43B-1674, December 2011. http://adsabs.harvard.edu/abs/2011AGUFM.P43B1674G
- [LANCE] D. Davies, et.al., "The Use of NASA LANCE Imagery and data for near real-time applications", IEEE IGARSS, Munich, July 2012. http://www.igarss2012.org/Papers/viewpapers.asp?papernum=4122
- [UAF] https://geo-ide.noaa.gov/wiki/index.php?title=Overview_and_benefits_of_the_GEO-IDE_UAF_Grid_Project
- [OGC] Implementations of OGC Standards. http://www.opengeospatial.org/resource
- [Fusion] G. Percivall, "OGC Fusion Standards Study, Phase 2 Engineering Report," OGC Document 10-184, 2010-12-13 http://portal.opengeospatial.org/files/?artifact_id=41573
- [OWS] http://www.opengeospatial.org/projects/initiatives/ows-9
- [AIP] http://www.ogcnetwork.net/AIP
- [NASA ITIC Smarr] March 8, 2012, Larry Smarr Chair, NASA Information Technology Infrastructure Committee (ITIC)

OGC Public Resources

- Adopted Standards:
  - http://www.opengeospatial.org/standards
- OGC Reference Model:
  - http://www.opengeospatial.org/standards/orm
- OGC Web Services 6 Testbed Demonstration Videos
  - http://www.opengeospatial.org/pub/www/ows6/index.html
- Compliance Testing and Certification
  - http://www.opengeospatial.org/compliance
- List of Registered Products using OGC Standards:
  - http://www.opengeospatial.org/resource
- OGC Network – member-contributed OGC "encyclopedia"
  - http://www.ogcnetwork.net
- OGC User – case studies of OGC implementations in the global community
  - http://www.opengeospatial.org, click on "Press Room"

Copyright 2012, Open Geospatial Consortium • *Making Location Count*