GAGE National Science Foundation's Geodetic Facility for the Advancement of Geoscience

UNAVCO

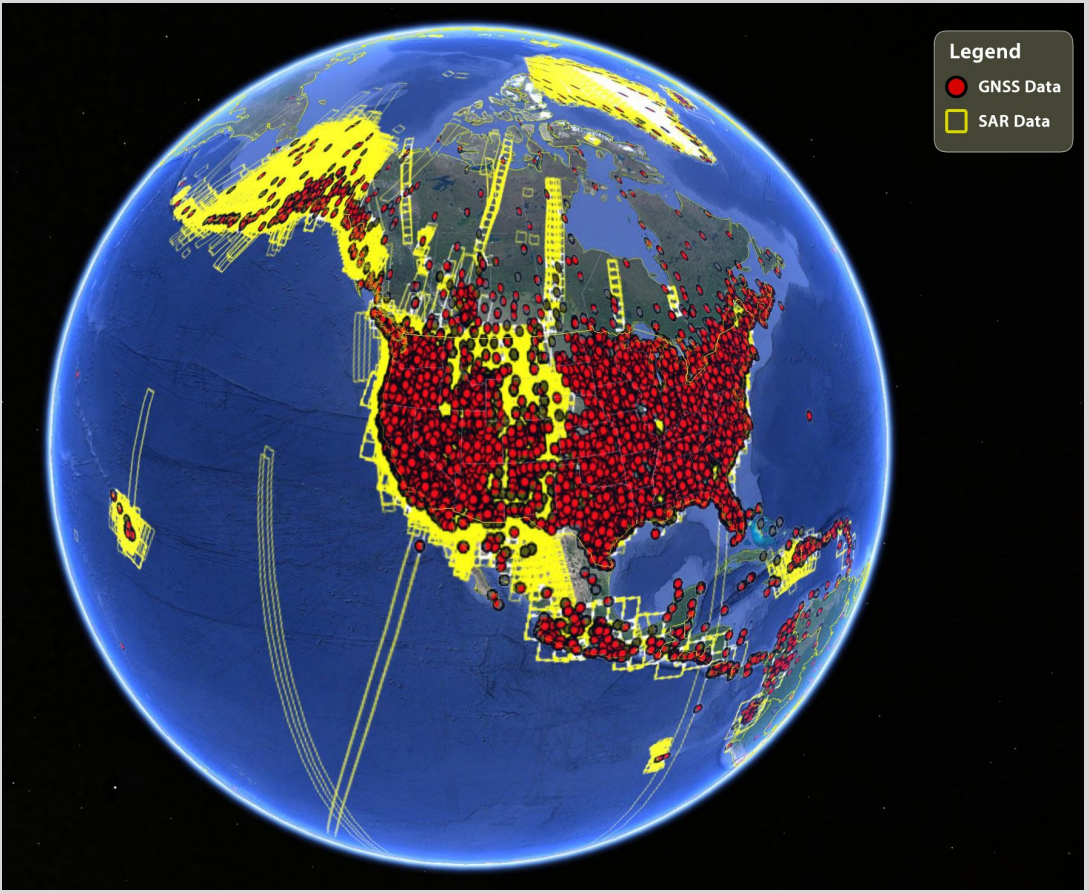


Image: Google Earth Data SIO, NOAA, U.S. Navy, NGA, GEBCO

GAGE Facility Geodetic Data Archive: Discoverability, Accessibility, Interoperability & Attribution

J. Riley, C. Meertens, D. Mencin, K. Hodgkinson, D. Ertz, D. Maggert, D. Reiner, C. Crosby, S. Baker (UNAVCO)

EGU2020-12417

Topics

- Data archiving multiple sensors and data product levels.
- Overview of GAGE Facility data archives, products and data access.
- Data access via FTP, web applications, and web services.
- Community contributed datasets.
- Persistent identifiers for Data attribution.
- Open standards & collaborations to develop and promote infrastructure, metadata and interoperability for the community.

GAGE Data Products

UNAVCO archives and provides search and access for data that our community of geodetic scientists uses for quantifying the motions of rock, ice and water that are monitored by a variety of sensor types at or near the Earth's surface.

The data types include:

- GPS/GNSS
- Imaging (point clouds) from SAR and TLS
- Borehole strain and seismic
- Meteorological

GPS/GNSS data sets, TLS data sets, and InSAR products are assigned digital object identifiers.

Level 0	Raw data products, typically in proprietary or complex formats.
Level 1	Pre-processed data products such as QC'd time series suitable for subsequent analysis.
Level 2	Post-processed data products derived using unambiguous/non-controversial methodologies.
Level 3	Scientific research products generated by the community and archived/distributed by GAGE.

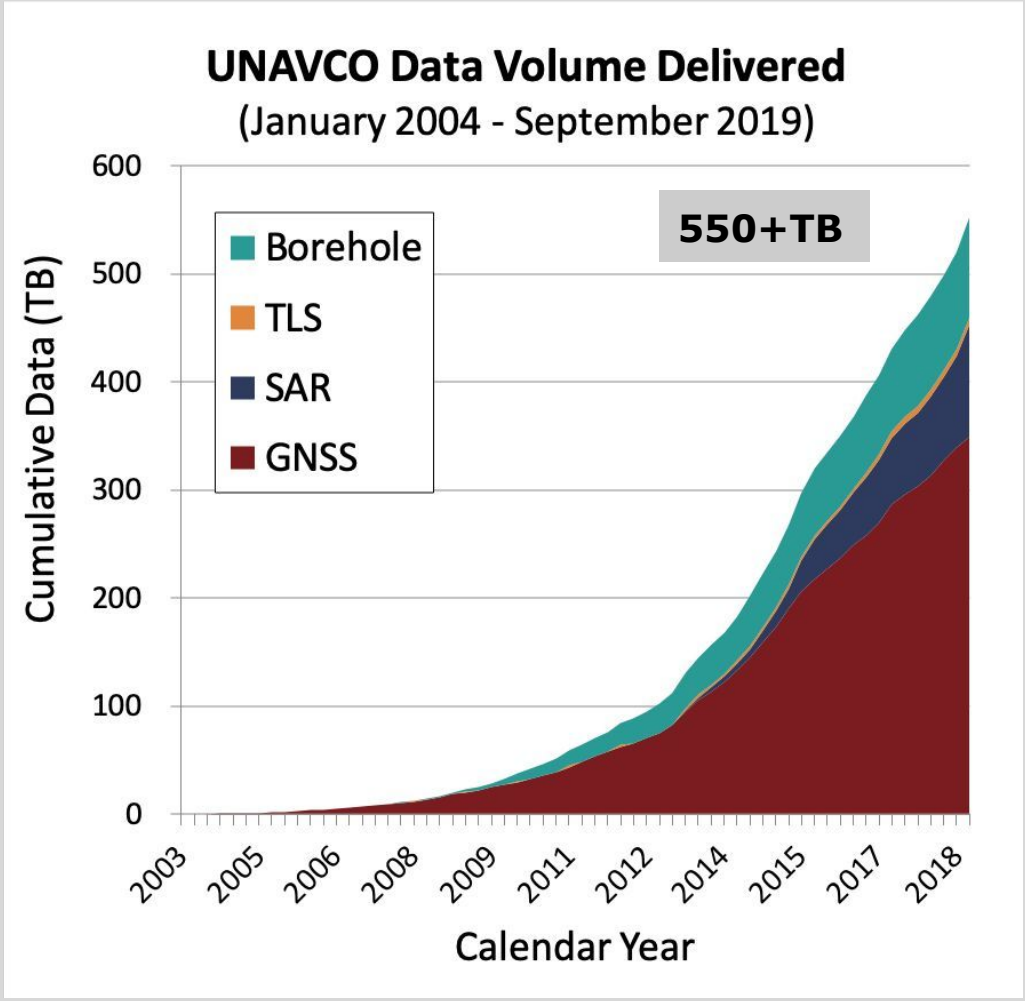
Data Center

On Premises: ~1.1 PB SAN Storage, ~40 host servers, 250 VMs, Internet 2, 50kw, offsite colocation of critical services, cloud backups

Data Volume: 350+ TB Archived, 550+ TB Delivered

Cloud services: (XSEDE, AWS, Google)

Data available by FTP, Web Applications, and Web Services



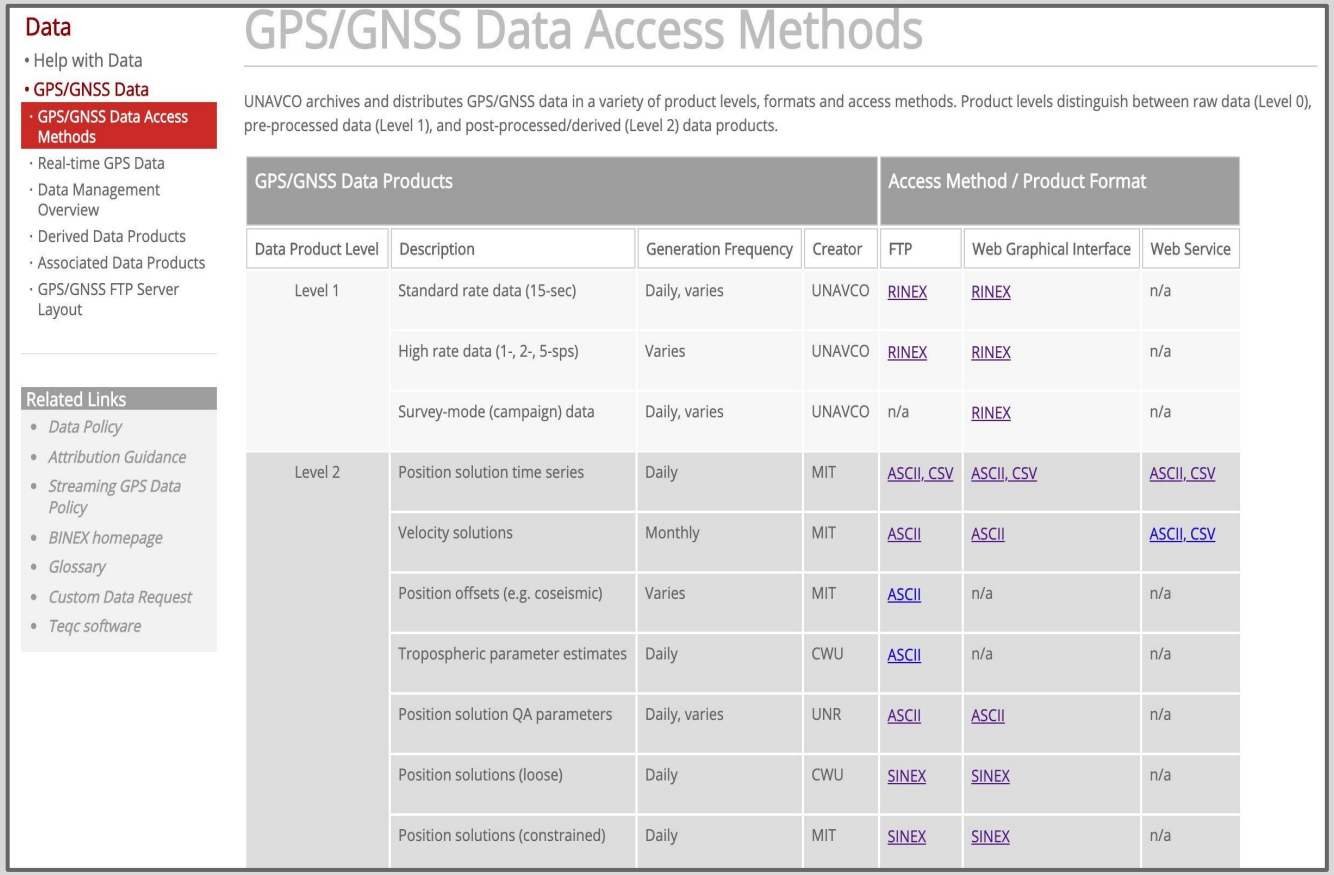
Data Search & Access

Data Findable and Accessible from UNAVCO by FTP, GUI and Web Services

Stations archived (permanent and campaign): 13,580

Permanent stations archived: 2,617

Permanent stations processed: 2,648



Data

- Help with Data
- GPS/GNSS Data
  - GPS/GNSS Data Access Methods
  - Real-time GPS Data
  - Data Management Overview
  - Derived Data Products
  - Associated Data Products
  - GPS/GNSS FTP Server Layout

Related Links

- *Data Policy*
- *Attribution Guidance*
- *Streaming GPS Data Policy*
- *BINEX homepage*
- *Glossary*
- *Custom Data Request*
- *Teqc software*

GPS/GNSS Data Access Methods

UNAVCO archives and distributes GPS/GNSS data in a variety of product levels, formats and access methods. Product levels distinguish between raw data (Level 0), pre-processed data (Level 1), and post-processed/derived (Level 2) data products.

GPS/GNSS Data Products				Access Method / Product Format		
Data Product Level	Description	Generation Frequency	Creator	FTP	Web Graphical Interface	Web Service
Level 1	Standard rate data (15-sec)	Daily, varies	UNAVCO	RINEX	RINEX	n/a
	High rate data (1-, 2-, 5-sps)	Varies	UNAVCO	RINEX	RINEX	n/a
	Survey-mode (campaign) data	Daily, varies	UNAVCO	n/a	RINEX	n/a
Level 2	Position solution time series	Daily	MIT	ASCII, CSV	ASCII, CSV	ASCII, CSV
	Velocity solutions	Monthly	MIT	ASCII	ASCII	ASCII, CSV
	Position offsets (e.g. coseismic)	Varies	MIT	ASCII	n/a	n/a
	Tropospheric parameter estimates	Daily	CWU	ASCII	n/a	n/a
	Position solution QA parameters	Daily, varies	UNR	ASCII	ASCII	n/a
	Position solutions (loose)	Daily	CWU	SINEX	SINEX	n/a
	Position solutions (constrained)	Daily	MIT	SINEX	SINEX	n/a

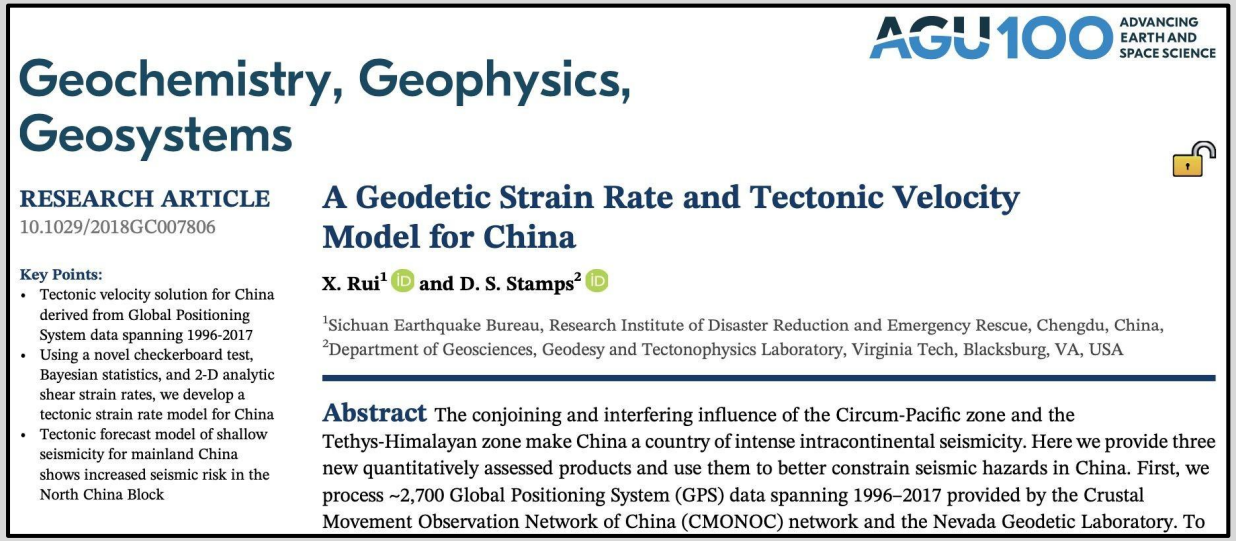
Community Contributed Datasets

In addition to Level 2 products generated by the GAGE GNSS Analysis Center (CWU) and Coordinator (MIT), community PI data products can be archived/distributed from UNAVCO Data Center.

- Rui and Stamps (2019) velocity solution file available from UNAVCO by FTP, and eventually by web services. Stamps et al. (2018) available soon.

Investigating RDM platform for Community/PI data product submission in a more streamlined fashion for publication-ready hosting and referencing using:

- ORCID for account management
- DOI minting
- Archiving
- Dataset landing pages



AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geochemistry, Geophysics, Geosystems

RESEARCH ARTICLE
10.1029/2018GC007806

Key Points:
- Tectonic velocity solution for China derived from Global Positioning System data spanning 1996-2017
- Using a novel checkerboard test, Bayesian statistics, and 2-D analytic shear strain rates, we develop a tectonic strain rate model for China
- Tectonic forecast model of shallow seismicity for mainland China shows increased seismic risk in the North China Block

A Geodetic Strain Rate and Tectonic Velocity Model for China

X. Rui[1] and D. S. Stamps[2]

[1]Sichuan Earthquake Bureau, Research Institute of Disaster Reduction and Emergency Rescue, Chengdu, China, [2]Department of Geosciences, Geodesy and Tectonophysics Laboratory, Virginia Tech, Blacksburg, VA, USA

Abstract The conjoining and interfering influence of the Circum-Pacific zone and the Tethys-Himalayan zone make China a country of intense intracontinental seismicity. Here we provide three new quantitatively assessed products and use them to better constrain seismic hazards in China. First, we process ~2,700 Global Positioning System (GPS) data spanning 1996–2017 provided by the Crustal Movement Observation Network of China (CMONOC) network and the Nevada Geodetic Laboratory. To

Persistent Identifiers

Data Products

Level 1: Pre-processed (RINEX files)

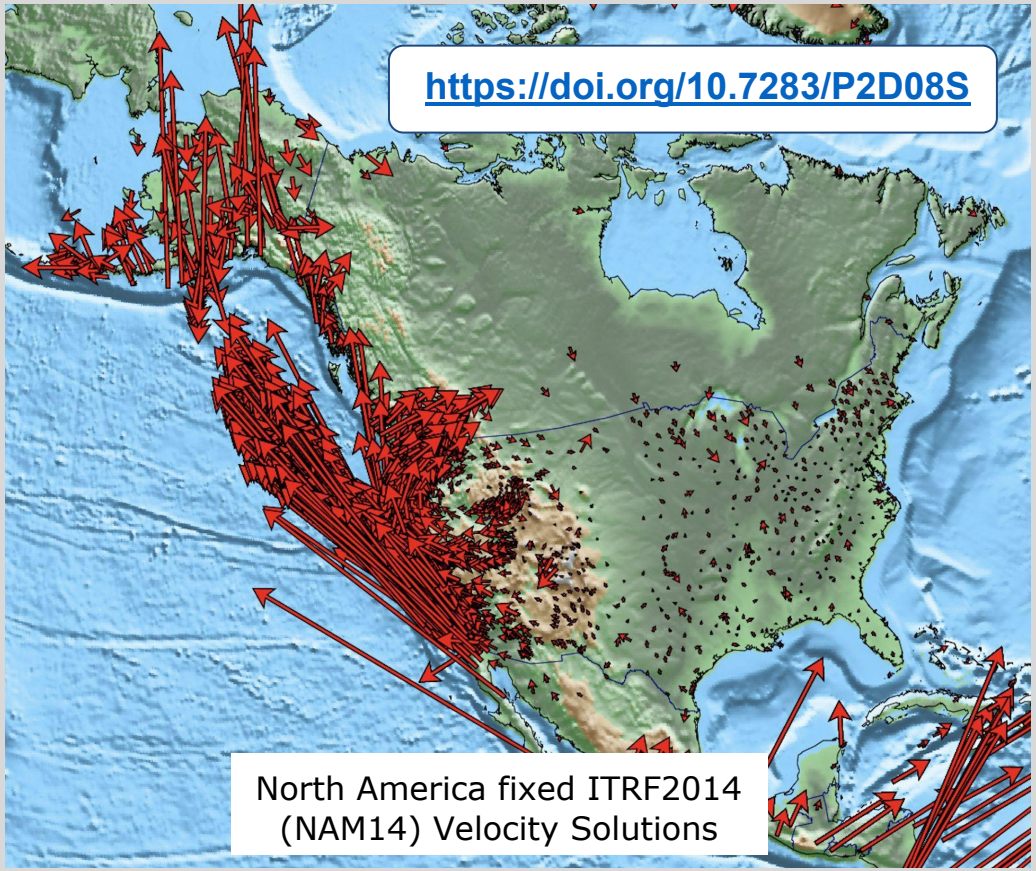
- DOI's issued for each continuous station
  - "Composite" DOI's possible
  - "Aggregate" DOI's possible
- DOI's issued for campaigns

L**evel 2**: Post-processed solution products (Position time series, velocities)

- New DOI issued for each annual "final" release

Investigating

- DOIs for Level 3 products
- RORs for organization identifiers
- ORCIDs for user identifiers

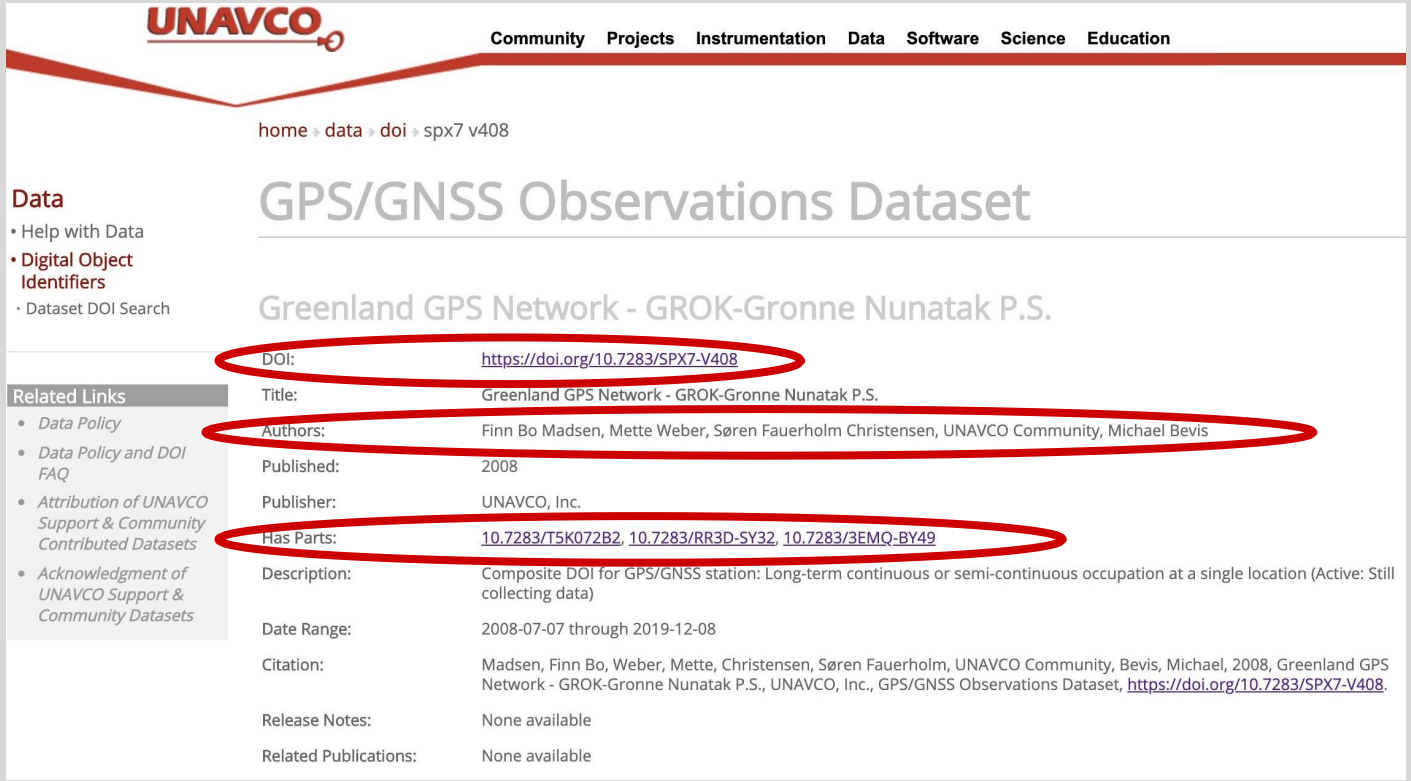


Persistent Identifiers & Attribution

Major Attribution changes, such as PI and funding source changes, are facilitated by "composite" DOI's for continuous GNSS station Level 1 data.

Investigating DOI best practices for datasets and versioning.

- Member of GGOS DOI Working Group w/ 22 members, convened 1st time 2019-12-09, chaired by Dr. Kirsten Elger (GFZ)
- Discussing data versioning preference over composite DOIs for changing metadata over time.



Open Standards

For interoperability, Geodetic Data Services re-architecting core data systems to better support a variety of geospatial data and real-time data streams and enable international standard approaches to geographical information search and access.

We are seeking standards for greater interoperability of our metadata within our data center, with partners, and with the world at large.

Implemented Schema.org structured data on DOI dataset landing pages in partnership w/Google Dataset Search and EarthCube Project 418/419.

GeodesyML development and adoptions led by **GeoScience Australia** with international collaboration.

Implementing OGC web services:

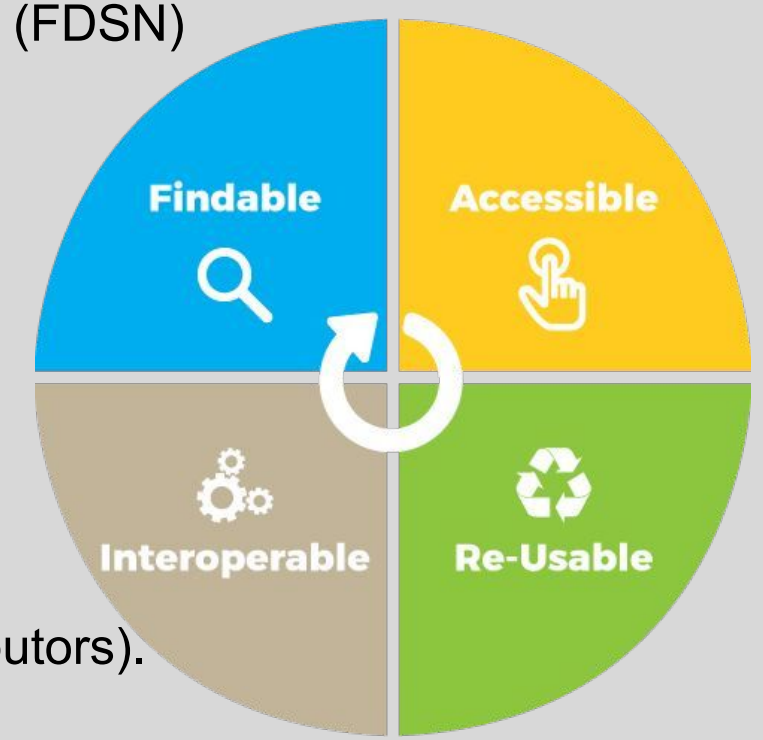
- WMS, WFS, WMTS

Investigating OGC web services:

- Sensor Web Enablement, SensorML
- ISO19115 from CSW

What: Geodetic data explosion require a modernization of how we share metadata, data, products

How: Standardize interfaces, use existing standards (RINEX3, GeodesyML, OGC) Learn from seismological community (IRIS): web services (FDSN)
Simple is powerful
Allow, share, document!



Why: Let scientists do the science
Metadata maintenance is painful but crucial
Validation and interoperability
Human readable formats
Provenance (DOI)
Open source community approach (champions, enablers, contributors).

Help us, tell us what you need! Geodetic Data Product survey, use cases.

Summary Slide from Community SIG at 2019 GAGE-SAGE Workshop convened by Elisabetta D'Anastasio (GNS Science), David Phillips (UNAVCO), Chad Trabant (IRIS), Mike Floyd (MIT)