## Open Data, Data Services, and Cross-Disciplinary Collaboration in Geophysics

Jeff Freymueller University of Alaska Fairbanks

## Outline

- Data: past present and future
- Open data: EarthScope defined the new normal
- Future integrated IRIS/UNAVCO data flows
- Seismogeodesy: An example of integrating cross-disciplinary types of data



## Now: Online, Download

- Today, all data are online, but transfer of large data sets remains much slower than local access
  - Most people download once, use locally repeatedly (local archive)
- Repeated download beginning to replace local archive
- A challenge: How to deal with meta-data updates?
  "Oops, the sensor was not installed pointing north", "Sorry, we reported the wrong type of antenna."

## Future: Will we access on demand?

- Increasing network speeds follow a power law
- How will you work if networking speeds are ~100x faster?



Image: Jacob Pedersen

## Move to Web Services

- IRIS and UNAVCO are moving away from older transfer protocols to web services
- Efficiency, scalability
- Send a URL string, get a data set/product back

#### Try it out! Hide Response

#### Curl

#### Request URL

http://web-services.unevcs.org/gps/deta/position/P328/v3TenelysisCenter-pboGreferenceFrame-iguM86 startilan-2000\_40-0110051A004A044A004A041ine-2000-03-0110043A0043A0043A004Feport=iang6 refCoordDpilan=from\_analysis\_center

#### Response Body

# dataset: GeoCSV 2.0

# field\_unit: 100 0601 datetime UTC, meters, meters, meters, meters, meters, number, number, number, degrees north, degrees east, meters, meters, meters, meters, meters, meters, number, number, number, UTF-0

- # field\_type: string, flost, f
- # attribution: http://www.unavco.org/community/policies\_forms/attribution/attribution.html
- # Request UR1: http://web-services.unavco.org/gps/data/position/9378/v37analysisCenter-pbcs

referenceFrame=igs06%starttime=2008-81-01700%3A00%3A00%sendtime=2008-83-01700%3A00%3A00%3A00%sepert=iong% refCoordOption=from\_analysix\_renter

- # Source File: 2010.pbc.ips00.pos Date: 2016-11-08 05:33:33
- # XTE Reference Coordinate (ige08) %: -2475699.75517 %: -3822330.2947 %: 6450718.33167
- # Applied Offset: H: 0 T: 0 E: 0
- Datelime, K. Y. H. Std Dev H. Std Dev Y. Std Dev H. Corr HY. Corr HH. Corr YH. H latitude. H longitude, Height, delta H. delta E. delta U. Std Dev H. Std Dev H. Std Dev U. Corr HH. Corr HU. Corr HU. Solution

2008-01-01200:00:00.-2475699.72594.-3822330.29432.4450718.32889.0.00234.0.00342.0.00356.0.80298.-8.75100.-0.81490.44.53498.237.04912.83.12234.0.00557.0.01593.-0.00963.0.00121.0.00157.0.00510.-8.62029.0.02750.-0.17109.repro

2008-01-02700-00:00,-2475699.73854,-3822330.29835,4450718.33338,0.00243,0.00354,0.00354,0.00367,0.88200,-0.75200,-0.81500,44.53458,237.06912,83.12893,0.00541,0.01594,-0.00306,0.00125,0.00142,0.00527,-0.62534,0.03542,-0.17589,repro

### EarthScope and Open Data

AAAAAAAAA

Drilling into the San Andreas Fault GPS Stations

Borehole Strainmeters

long-baseline Laser Strainmeters

A Transportable Seismic Stations

APermanent Seismic Stations

### EarthScope Data

# The flood of open data from EarthScope changed the way we do science

Data Shipments by Request Type

1202.95 terabytes projected on June 30, 2017







#### **IRIS DMC Archive**

#### as of 1 July 2017

#### 447.0 terabytes

#### reflects removal of replicated data







Cumulative Data Archived through Mar 2018 ~306 Tb

UNAVCO

- Cumulative Data and Derived Data Products Delivered through Mar 2018 ~474 Tb
- Total EarthScope PBO Data Volume Archived (all products) = 145 Tb
- Total EarthScope PBO Data Volume Delivered (all products) = 349 Tb

## Impacts of Open Data

- The vast holdings of open data have accelerated the pace of our science
  - Analyses of data that used to be research are not done automatically
  - The research frontier is now more in synthesizing and integrating
- Cultural shifts are in progress...

### The Future: IRIS and UNAVCO Together

• IRIS and UNAVCO have proposed to work together more closely on the future NGEO facility.

NSF is still working on the NGEO decision...

- Seismologists and geodesists commonly work on different things, but we can make use of each others' products
- Seismogeodesy holds great promise for integrating the fields, at least for large earthquakes

### **Proposed NGEO Data Access Point**



- Building on EarthCube GeoWS project, IRIS and UNAVCO proposed a unified NGEO data access point
  - Cross-disciplinary users are probably most likely to access products rather than raw data
  - Will work for raw data, too!
  - Future: GPS waveform data
- Easier to implement with modern protocols

### Seismogeodesy



### Seismic Wavefield from 2011 Great East Japan Earthquake



## Seismogeodesy

- Looser definition:
  - Use of high rate GPS positions to study dynamic ground motions
  - Using geodetic data the way a seismologist would
- Tighter definition:
  - Combination of co-located geodetic and seismic instruments to produce a single time series of ground motions

## Advantages and Applications

- Kinematic GPS position time series combine information about static and dynamic offsets
- GPS position/velocity records can be filtered to be used as seismograms
  - It is even possible to do tomography with GPS seismograms!
    - Why? For some places and times, there are more GPS than seismometers
- No instrumental saturation with GPS
  - No problems with clipping, ground tilts, etc

### Ground Motion/Magnitude Scaling holds to M~9



doi: 10.1093/jpi/jpg/78

Geophys. J. Int. (2014) 296, 461–472 Advance Access publication 2013 October 17

#### Determination of earthquake magnitude using GPS displacement waveforms from real-time precise point positioning

Rongxin Fang, Chuang Shi, Weiwei Song, Guangxing Wang and Jingnan Liu 2020 Insurit Contr. Water University, Pater 10079, Clina E and Milliofe educe.



### Back to Gutenberg (1945)!

GEOPHYSICAL RESEARCH LETTERS, VOL. 40, 6089-6094, doi:10.1002/2013GL058391, 2013

#### Earthquake magnitude scaling using seismogeodetic data

Brendan W. Crowell,<sup>1,2</sup> Diego Melgar,<sup>2</sup> Yehuda Bock,<sup>2</sup> Jennifer S. Haase,<sup>2</sup> and Jianghui Geng<sup>2</sup>

Received 21 October 2013; revised 26 November 2013; accepted 27 November 2013; published 10 December 2013.

$$\log(P_d) = -0.893 + 0.562M_w - 1.731\log(R)$$
(2)

### **GPS** Displacements/Velocities



## Conclusions

- Open data and fast network speeds have changed/are changing the way we work
- Future scientific frontiers will increasingly depend on synthesis and integration of large data sets
- IRIS and UNAVCO are poised to help all of us exploit these opportunities
- There are exciting scientific opportunities in seismogeodesy, the fusion of seismology and geodesy