

# CDI Monthly Meeting 20170614

## CDI Monthly Meeting - 20170614

The Community for Data Integration (CDI) meetings are held the 2nd Wednesday of each month from 11:00 a.m. to 12:30 p.m. Eastern Time.

### WebEx:

<https://usgs.webex.com/> - Under the Meeting Center tabs, search for meeting name: "Community for Data Integration."

### Audio:

**USGS/DOI Dial In Number:** (703) 648-4848 (for USGS and DOI offices)

**Toll Free Dial In Number:** (855) 547-8255 (for other offices and telecommute locations)

**Conference Code:** 47919# (same for both numbers)

## Webex Recording

Webex recordings are available to CDI Members approximately 24 hours after the completion of the meeting. Please login to view the recording. If you would like to become a member of CDI, please email [cdi@usgs.gov](mailto:cdi@usgs.gov).

## Agenda (in Eastern time)

11:00a [Scientist's Challenge: Social media, Breakfast with Bill, and a multi-beam community](#)

11:10a Welcome - Tim Quinn - Office of Enterprise Information Chief [\[PDF\]](#)

Tim Quinn: Internet of Things presentation at Amazon Cloud Workshop - Presenter took USGS ideas and applied them to Indonesia and their problems with monsoon and high water issues inundating the city. Took earthquake early warning people have done. Sensor based system. Implemented it via Amazon Cloud. Harvesting tweets that we use for earthquakes except harvesting them for flooding. Got ideas from USGS but implemented them in the Cloud.

11:15a Working Group Announcements [\[PDF\]](#)

Additional Announcements

Model archiving discussion

- Happy coincidence that modeling was part of the Workshop theme
- Discussion was pretty active, if inconclusive, during the Workshop
- good start tho!
- Cross Working Group topic:
  - definitely DMWG
  - possibly also SWWG
  - Very interested in finding others
- TSWG will host a series of talks that will hopefully result in some consensus building.
  - WMA has "first generation" protocols, with differing levels of sophistication across their disciplines. Groundwater is in the lead and will hopefully present on their motivations, concepts, and tech choices
  - CSS's Biogeographical Characterization Branch will (hopefully) present on their experiments using the DataOne Prov for describing data provenance
  - Semantic Meta-modeling follow-on
  - Number of related academic groups, we're looking at which is the best starting point for us
    - Earth System Documentation for models, based on the Common Information Model
      - <https://earthsystemcog.org/projects/es-doc-models/>
    - CSDMS (Community Surface Dynamics Modeling System)
    - Others
  - Develop lists or use cases around CDI member needs, Agency policy (FSP, publications, etc.)
- Definitely do not want to reinvent anything. Big, complicated, common problem. There are already some folks working on or around this issue w/in USGS, as well as beyond. Purpose is to:
  - educate ourselves a bit
  - Develop thinking about our special needs
  - Think about increasing impact of modeling-based science produced by CDI members
  - help guide policy & protocol development around this type of scientific information content.

Earth-Science Themes Working Group

- Co-sponsored an additional session of talks on National Geospatial Data Development
  - In addition to bringing national data set producers together
  - Proposed to develop more active method for making COPs aware of ongoing work (avoiding redundancy!)
  - Sharing methods is a big deal (accumulation of land surface characteristics downslope is a big example)
- ETWG: although not present for the workshop, advertised the
  - Bioinformatics Focus Group, which is really active
  - ICEMM (pronunciation): originally (and still) a Federal interagency group.
    - ~15 years old. Moving into CDI's wiki space.
    - Has own governance and internal structure of WGS
    - Got lots of friendly feedback from them in preparing, would like to develop relationships on this. Really more than ETWG scope here.
    - Hoping they might have opinions on model archiving
- ETWG had a series of topical talks focusing on the "river corridor"
  - Powell Center project chief
  - Ecosystems MA presented on Ecological Flows initiative, as well as Drought
  - Leveraging 3DEP lidar for extraction of features and landscape characterization
    - Examples include floodplain mapping, near-stream canopy height and biomass estimation
  - NAWQA presenter: soils data and riparian zone mapping from USFS
  - Been a fair bit of email chatter amongst speakers, trying to push things into the wiki for wider audience. Should try to line up talks, but haven't had time just yet.
- Really tried to drive home that adapting style of work, culture to be more open is key. Not necessarily hard, but requires a somewhat conscious effort to change long-established patterns.

11:25a **2017 CDI Workshop Debrief - Charting a Course toward Integrated Science** - Leslie Hsu, USGS

**Presentation:** Slides are available to CDI Members. Please login to download the slides. If you would like to become a member of CDI, please email [cdi@usgs.gov](mailto:cdi@usgs.gov).

#### Abstract:

The Community for Data Integration held an in-person meeting from May 16-19, 2017 with the theme of "Enabling Integrated Science." During the meeting, many presentations and discussions were focused on the activity of building a roadmap to integrated science, that is, assessing the current state, gaps, and steps that could be taken to increasing our ability to answer societal needs with cross-Mission Area data, methods, and knowledge. The results of those discussions will be passed on to USGS Leadership and also used to guide CDI activities in the coming year. We'd like to get as broad input as we can from the community, but because not everyone was able to attend the workshop, we'll review the outcomes of these discussions for further input at the June CDI Monthly Meeting.

Over 175 participants attended presentations, member-run topical breakouts, the DataBlast demo and poster session, and discussions about steps the USGS can take to advance integrated science in the Bureau. Keynote speakers included Bill Werkheiser, Acting Director of the USGS, Kevin Gallagher, Associate Director of Core Science Systems, Tim Quinn, Chief of the Office of Enterprise Information, Marty Goldhaber, USGS Senior Scientist, and Bruce Caron, community architect at Earth Science Information Partners.

[See the workshop webpage for more information.](#)

#### **Links for this month's presentation:**

- [Scientist's Challenge: Social media, Breakfast with Bill, and a multi-beam community: https://my.usgs.gov/confluence/x/JaMRlg](https://my.usgs.gov/confluence/x/JaMRlg)
- Member stories: <https://my.usgs.gov/confluence/display/cdi/CDI+Member+Stories>
- Working groups: <https://my.usgs.gov/confluence/display/cdi/CDI+Working+Groups>
- June Meeting page: <https://my.usgs.gov/confluence/display/cdi/CDI+Monthly+Meeting+20170614>
- [Slido.com](#), event code #W888
- 2017 CDI Workshop page: <https://my.usgs.gov/confluence/display/cdi/2017+CDI+Workshop>
- Workshop Presentations: <https://my.usgs.gov/confluence/display/cdi/2017+CDI+Workshop+Presentation+Page>
- Big Science has a buzzword problem (Scudellari et al. 2017), <http://doi.org/10.1038/541450a>
- Bruce Caron's blog: <https://cybersocialstructure.org/>
- DataBlast posters: <https://my.usgs.gov/confluence/display/cdi/2017+DataBlast+Posters>
- Roadmap recommendations voting results: <https://my.usgs.gov/confluence/display/cdi/Sli.do+polls+from+the+workshop>
- Advanced Computing Cooperative email for questions or help: [gs-css\\_csas\\_hpc\\_help@usgs.gov](mailto:gs-css_csas_hpc_help@usgs.gov)
- [Sli.do](#) questions from the workshop: <https://my.usgs.gov/confluence/display/cdi/Questions+and+comments+from+Sli.do>

12:30p Adjourn

## Presentation Q/A

Slido Questions and Polls [xlsx]

Viv: A couple questions on slido about exceeding the capacity of SB (large datasets). We are working on that as the SB team. We are partnering with ARC about thinking about how to get data produced on Yeti released. Also, partnering with EROS to see if we can use some of their servers for larger datasets. Has anyone used any creative solutions to bridge our current gap?

Stephanie Galvan: CDI Funded Data At Risk project has a featured story on Bathymetographs: <https://www.usgs.gov/news/usgs-bridging-generations-wwii-technology>.

Marcia: It's a really interesting story!

## Attendees

A WebEx Participant Report is available to CDI Members. Please login to download the report. If you would like to become a member of CDI, please email [cdi@usgs.gov](mailto:cdi@usgs.gov).