

Participants

Tanya Haddad, Scott Toews, Aaron Macgregor, Steve Steinberg, Anna Zivian, Rachel Rodriguez, Jan Newton, Alan Allwardt, Emilio Mayorga, Chris Romsos, Beth Timmons, Andy Lanier, Todd Hallenbeck, Tim Doherty, Paul Olin, Chris Cohen, Peter Lawson, Jim Power, Carrie Bretz, Evan Paul, Robbie Wilson, Stephanie Rozek, Tim Welch, Matt Armsby, Jon Bonkoski, Ryan Hodges

Agenda/Presentations

- [Regional Data Framework background presentation - Andy Lanier](#)
- [Report from Survey of the ACTs - Jon Bonkoski](#)
- [RDF Data Registry what will it look like - Tim Welch](#)

Discussion

Steve: For Tim: You've looked at multiple CSW catalogs - is there any value to using a single platform across groups that contribute to the registry

Tim: the Registry is a lightweight abstraction layer that talks to the catalog, via CSW, so in theory the catalog can be swapped in the background

Steve: question for Jon's presentation - how in the weeds did the questions about specific data sets get (eg. data about specific species).

Jon: pretty specific, but the spreadsheet is a summary

Evan Paul: wondering if the viewer has been part of what has been evaluated in Phase I

Tim: The registry is not a planning tool or viewer - any viewing component is minimal, it is primarily for discovery. Any planning tool or robust viewer is envisioned in a future phase of work.

Tim in Oakland: Is the schema proposed in the diagram in use in the Northeast on MARCO?

Tim from Ecotrust: It is more advanced - MARCO is similar to the NOAA CMSP registry (spreadsheet in the background), Any changes to services URLs require manual updating, as opposed to a more automated system that can read metadata directly from the host collection

Todd: does this system require extremely robust Metadata, or is metadata lite an option?

Tim: a basic minimal set of discovery fields will be required, but it is not a high bar - CSW

has a minimum that needs to be met, and even if a data set does not have those, we can provide an online form so they can be created on the fly when an entry is submitted

Robby: Recommend that we address the backend metadata challenges upfront. As you look at the data and themes ask the key questions upfront: where is the data?, where are the services?, do you have metadata? Recommend the front end people are talking to the back end people that are actually hosting and serving the data (in home institution). In a metadata driven system, the minimal stuff must be provided, or else there needs to be a “Man in the Middle” that fills in the gaps.

Tim: We recognize that the various capacities out there mean that it will be a gradual process for data providers to connect to the system. This design allows for providers to manage their own data and submit their records when they are ready (with their minimal metadata)

Emilio: General comment that that the architecture diagram really reflects the design direction that the technical discussions have taken, so that is encouraging. The specific issue related to “where is the data / metadata?” is realistic in that we may have willing providers, who do not have the resources to host data. Points out that the human network is hoping to connect such under-resourced providers with willing intermediaries who can do the lifting in terms of hosting of services. There are likely many such intermediaries (eg OOS institutions, state agencies, ESRI cloud community). Question for Tim: timeline on the next steps?

Emilio: maybe one next task is to begin to identify willing intermediaries from within our human network.

Tim: the requirement documents will be worked on in the next couple of weeks and finalized towards the end of September - there are several milestones on the horizon that might help evolve/influence some of our recommendations (emergence of the NE/MARCO/SE portal designs), other related efforts that might provide interesting examples to examine

Tanya: The ICAN v3 prototype will be going live at the end of September and there should be an early October presentation about it's back-end architecture that might be interesting to view to see if any interesting similarities emerge (it is a CSW-based back end as well).

Andy: the working groups now have a few tasks that can be tackled:

- the data group could look at Jon's worksheet and begin to think about gaps.
- the IT group could provide feedback on the architecture diagram, and
- the outreach group could begin the task of identifying intermediaries.

Chris R: Likes the idea that the Data working group can use the worksheet as a starting point to begin to work on the data gaps issue - and maybe the capacity issue (in terms of hosting resources / problems that people have)

Tanya: is there enough information in Jon's survey to begin to identify the potential hosts, vs. folks who need help hosting?

Jon: Yes - we should have enough beginning information already to give a good start to the data group

Emilio: we should make sure the solutions document is reviewed by the whole IT working group (not just the ACT) so that the feedback is as productive as possible.

Scott: Comment on the heavy lifting needs - remind people that California is going through this process right now, and so they are identifying data that is already ready vs data that will need work. So rather than reinvent the wheel we should use the results of that work (for California)

Todd: Timeline?

Steve: within a month or two.

Jon: the nature of the loose organization of ACTs is a bit of a roadblock to data publishing sometimes, so the matching of data holders with willing hosts (like the California Geoportal) is probably a very good idea.

Todd: Yes - the nature of the ACTs is that they do not have their own resources for this - they rely on their members to be the active contributors.

Robby: comment on NOAA products like ENC's these aren't easy products for outsiders to work with so if we can communicate with NOAA what exactly we need, then some of the Fed partners can do the internal heavy lifting to produce something appropriate for the RDF community.

Andy: That is a great offer/idea from Robby - exactly why we need a federal liaison on the ACT, so that we connect with the correct feds as we encounter such issues.

Chris: To try to do the determination if the products available are actually what people need is something that the Data group can incorporate into their gap analysis (if the format is wrong it is a form of gap).

Chris: Maybe we should communicate the function and reality behind the registry on the front end of the system so that users understand the big picture, and have realistic expectations that this registry is aggregated around the needs of the ACTs

Robby: focusing on steps 1,2,3 and 7 of the diagram help you solve the users problem early and solving the steps 5,6 are the longer term technological road map for all the data providers to aspire to

Andy: we also want the front end to make it easy for data producers/users to get involved - it is still very early in the timeline to know what the front end will look like, those discussions have not been had yet, but we will make sure that they do occur and involve everyone that wants to be involved

Matt: It will be interesting to see how the Excomm responds to what has been done to date, and we will be reporting back to folks about how things go

Next Steps

Data working group to meet and review data spreadsheet, discuss metadata standards/requirements.

IT group will review technical options when that matrix is finalized.

Adjourn 12 pm, Thanks to all for attending!