

NOAA Environmental Information Services (NEIS)

--

A framework of layered services designed to help NOAA's mission by facilitating the discovery, access, integration, and understanding of all NOAA data (past, present, and future).



EDM Conference
May 16th, 2012

Jebb Q Stewart
NOAA/OAR/ESRL/GSD



ESRL Global Systems Division

NEIS Background

GSD Director Discretionary Funds (DDF) Proposal

NEIS Requirement

- Fast access to any data, any time, right now, on any platform.

NEIS Group Objectives

- Build upon NEIS work in the past
- Assess GSD capabilities pertinent to the concept of NEIS.
- Evaluate existing, developmental technology, and solutions relevant to concept.
- Evaluate envisioned needs of primary users.
- Provide recommendation of how to best use GSD resources to contribute to a National capability.



Lots of Great Technology Exists.

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



THREDDS
Google Earth
MERMAid
Grads Data Server

LAS
NASA World Wind
REVERB
OGC Services

ERDDAP
GeoPortal
iRODS

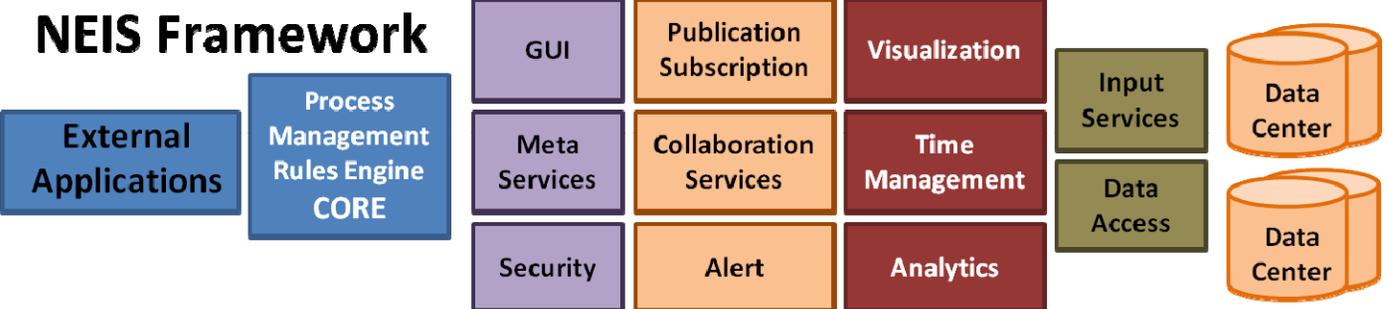
More...



ESRL Global Systems Division

NEIS Concept

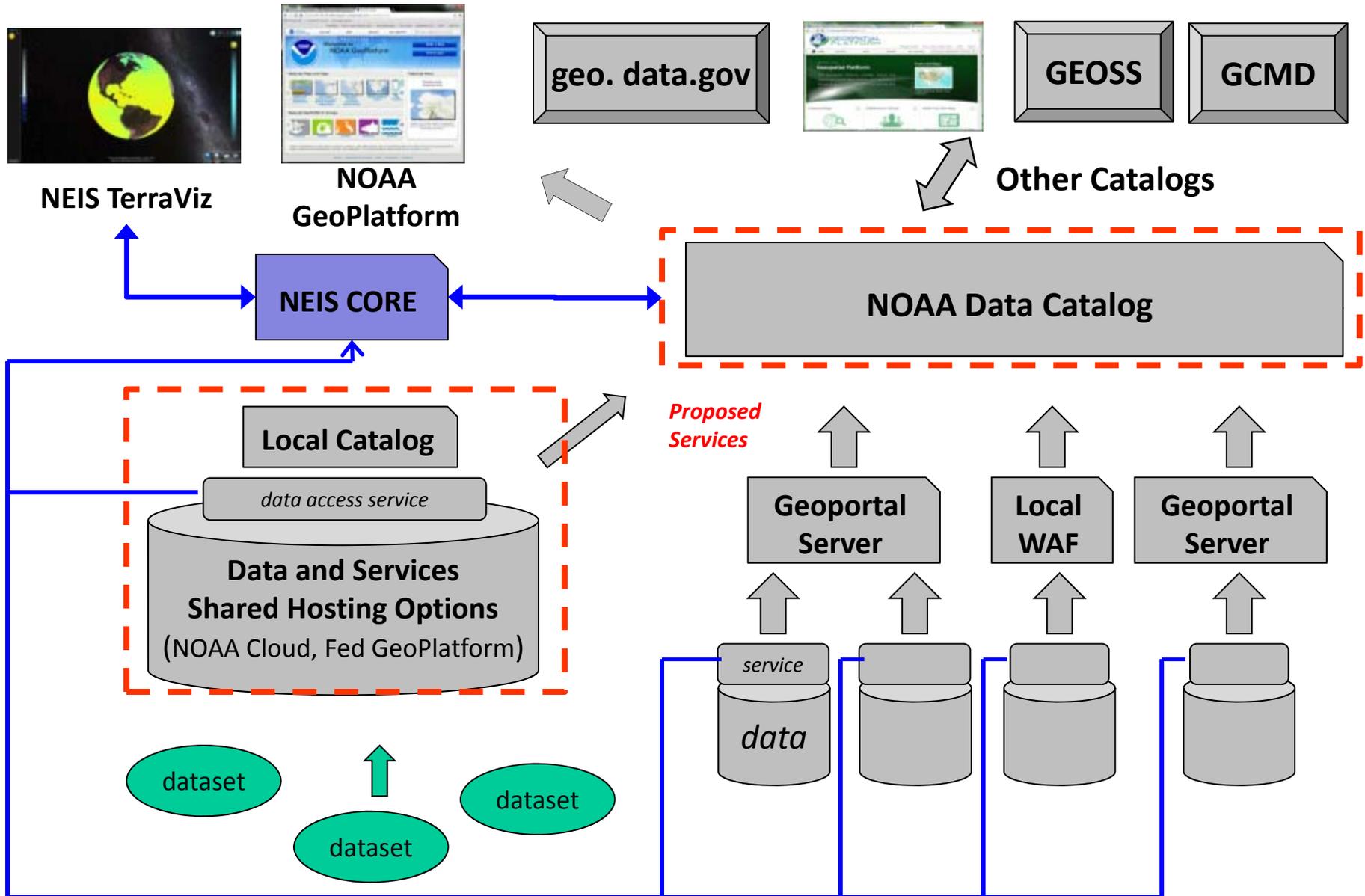
Take advantage of existing technology where we can to quickly find and access to data 100 of years in the past to 100 years in the future, from the top of the atmosphere to the bottom of the ocean in a single application.



Framework provides capability to answer questions that require data from different data sources.



NEIS in NOAA



What is TerraViz?

What is TerraViz?

- A system developed from the ground up by GSD based on Unity3D.
- There are no licensing restrictions in distributing the TerraViz executable.
- Unity3D is a well supported commercial game and data visualization engine with 750,000 active users
- Unity/C# is designed for 3D development, making it easy to create multiple views, create 3D meshes representing environmental objects, etc.
- Develop once, then will run under Windows, Mac, web browsers, iPhones, iPads, Android devices, and even game systems (Wii, Xbox 360, etc.)
- We can modify the user interface and add any feature we need.
- TerraViz Handles millions of polygons and performs quickly

TerraViz is one possible visualization tool built to develop, test, and showcase NEIS framework.



Why not Google Earth?

Google Earth, works great for certain applications and static products, however:

- Handles ~10,000 polygons before slowing to a crawl.
- Closed system (Google Earth source code not available) so we can't add needed features.
- Google Earth plug-in works on Windows and Mac 32-bit browsers only (won't run on iPhones, iPads, or Android devices.)
- Developing GE plug-in user interfaces in browser-dependent Javascript is tedious.
- Very Limited integrated data discovery or browsing, most uses require a separate application.
- Limited animation/data interaction support.
- Limited support for non-static data (ie Time Series)



NEIS - Finding and Discovering Data

NEIS Need: A way to programmatically discover and find data.

- NEIS is using Apache Solr to index ISO metadata records from metadata catalogs (currently geoportal) and related service endpoints.
- Solr provides temporal and geospatial filtering.
- Using facets on keywords, service name to fine tune searches.



The screenshot displays the NEIS search interface. At the top, there are search filters: a text input for 'Keywords', a 'Lat' field set to '50', a 'Lon' field set to '-70', and a 'Radius Km' field set to '2000'. To the right of these fields is a globe icon with a circular path around it, and a 'Search' button. Below the search filters, there are 'Min Date' and 'Max Date' fields, and a 'Clear' button. The main content area is divided into two columns. The left column, titled 'Available Tags', lists various tags with their counts: sea (326), oceans (323), ocean (293), water (284), seawater (279), monthly (220), mean (207), degree (197), temperature (195), chemistry (190), and deviation (189). Below this list are 'Include' options for 'local data' and 'online data'. The right column, titled '1 - 50 of 375 Matching Data Sets', shows a list of data sets. The first three items are: 'NOAA Marine Obs on 4-24-2012 (KML)' from NOAA National Data Buoy Center, 'NOAA Ship Locations on 4-6-2012 (KML)' from NOAA, and 'NOAA Fisheries Locations' from NOAA Fisheries, with a time range of 1980-01-01 to 2020-12-31. Below the list are buttons for 'Load Dataset', 'More Information', and 'Force Reload?'. The interface also includes a 'Tags->' section with a 'noaa' tag selected, and 'Prev' and 'Next' navigation buttons.



Finding and Discovering Data Challenges

Challenges:

- Having information readily available on data and services for searching.
- Harvesting versus Aggregated searching
- Metadata ISO records are not the only source of information. Some information is needed from service endpoints.
- Inability to access data programmatically due to lack of service information in Metadata ISO record.

Recommendations:

- Metadata records must contain service endpoints of where to programmatically access data.
- Helpful metadata on time information includes dataset start time, periodicity (daily, monthly, yearly), a list of outages, and an end time if dataset is no longer produced.



Metadata Challenges – Tags and Keywords

Challenges:

- Metadata keywords inconsistent
- Lack of meaningful tags (eastward, longitude).
- Keyword repetition – same keywords are used on several data sets.
- Semantics, Hierarchy, Taxonomy, Ontology, Relationships – All difficult to infer.

Examples:

“Sea Surface Temperature” “SST” “Sea” “Surface” “Temperature”

“NOAA” vs “National Oceanic Atmospheric Administration”, “Wind” and “Winds”

Ocean -> Ocean Circulation -> Ocean Currents

Recommendations:

GCMD Keywords: http://gcmd.nasa.gov/Resources/valids/archives/keyword_list.html

NASA JPL SWEET Ontology <http://sweet.jpl.nasa.gov/>

Allow user defined keywords (Crowdsourcing)



Data Access and Display Challenges

Challenges:

- Lack of adherence to specifications and consistency behind standards.
- Generation of graphical representations of data. How do I stylize point information? What color palette is the best for this data?
- Server uptime and data availability. Indexed ISO records say data is available for time but server may be down during access.

Examples:

- ncWMS, the WMS behind THREDDS, uses a default data range when applying a color palette to create an image, often resulting in an empty image.

Recommendations:

- Metadata uniformity helps improve user experience – need best practices, examples, and/or automated ways to produce information.
- Dashboard system to give feedback to data providers.
- Improving WMS experience, color palette is dynamically applied to entire range of data requested or use predefined color table that fits entire data.



Other Considerations

How do we encouraging good data stewardship?

Metadata is vital for applications such as this.

Recommendations:

- Showcase of what is possible with good metadata.
- Provide Best Practices/Examples of what is good.
- Provide tools to help automate process of creating metadata.
- Provide feedback to data providers on metadata (rubrics, other information).
- Show what metadata was provided by similar data sets.

Helpful information in metadata record:

- Links to data services
- Short names
- Timespan/Periodicity/Spatial Information
- Preview/Thumbnail



Dashboard Concept

The image displays two browser window mockups of a Geoportal dashboard. The left mockup shows a grid of status indicators for General and Compatibility categories. The right mockup shows a detailed view of the Compatibility section with specific messages for Visualization, Messaging, Analytics, and Security.

| General | | | |
|---------------|-------------|-----------|-------------|
| Availability | Consistency | Metadata | Information |
| Compatibility | | | |
| Visualization | Messaging | Analytics | Security |

Visualization
Compatible - Based on metadata for this data set, you are compatible with visualization services

Messaging
Compatible - Based on metadata for this data set, you are compatible with messaging services

Analytics
Not Compatible - Please update metadata record to include information on Units of Measure for variables to be compatible with Analytic Services

Security
Compatible - Based on metadata for this data set, you are compatible with security services

- Suggest metadata enhancements to improve data integration.
- Monitor for problems discussed above.
- Provide information on standards adherence (ex. OGC).
- Monitor data for changes (changes of fields, service endpoints, units, uptime, outages.)



Next Steps

Improving search and discovery of data:

Better understanding of semantics and relationships.
Crowdsourcing tagging.
Suggested data sets, like Netflix or Amazon.

Scenario based searching:

Fisherman / Hurricane / Education/ Climate.
Learn what is used in these profiles and score higher for future searches.

Improving data access and integration:

Using additional data access methods (ERDDAP, more OGC.)
Cross service Server Side Processing / Analytics.



Looking to the Future

NEIS is looking towards the future for data visualization and interaction through integrating and leveraging new and emerging technologies to meet NEIS goal 'any data, any location, any platform, now'

- Perform processing within cloud environment and with high speed connectivity to data sources, taking advantage of large processing power within clouds.
- Send graphics and server side processed/rendered/streamed data to GUI, improving bandwidth utilization.
- Take advantage of fast networking to make remote requests and processing appear like local application.
- Similar to how the new Amazon Silk Browser works.



Questions?

Jebb.Q.Stewart@noaa.gov

<http://www.esrl.noaa.gov/neis>



ESRL Global Systems Division

Extra Slides



ESRL Global Systems Division

