Data Science for NOAA Big Data

Story

Slides

Slide 1 Data Science for NOAA Chief Data Officer and Big Data Analytics Meetup
Slide 2 Agenda
Slide 3 Calendar
Slide 4 FDA Data Innovation Lab and Predictive Analytics Meetup
Slide 5 Conference for NSF Data Scientists, Data Infrastructure, and Data Publication
Slide 6 Semantic Insights Followup
Slide 7 NOAA Embraces the Business of Big Data
Slide 8 NOAA Big Data Partnership Model
Slide 9 NOAA Big Data Industry Day
Slide 10 Start: Government Data Hubs
Slide 11 Click To: Three NOAA Data Hubs
Slide 12 Click To: Department of Commerce
Slide 13 Click To: Data.gov Department of Commerce
Slide 14 Click To: Department of Commerce APIs
Slide 15 Also Filter To: Data.gov NOAA
Slide 16 Also See Prototype: data.noaa.gov
Slide 17 Filter to Excel in Prototype: data.noaa.gov
Slide 18 Also See: Environmental Research Division's Data Access Program RESTful Web Services
Slide 19 Data Science for NOAA Big Data: Build Knowledge Base
Slide 20 Data Science for NOAA Big Data: Knowledge Base Contents
Slide 21 Data Science for NOAA Big Data: Spreadsheet Knowledge Base Index
Slide 22 Data Science for NOAA Big Data: Spreadsheet Data.gov DoC Index
Slide 23 Data Science for NOAA Big Data: Spreadsheet Data.gov NOAA Index
Slide 24 Data Science for NOAA Big Data: Spreadsheet data.NOAA.gov Index
Slide 25 Data Science for NOAA Big Data: Spotfire Cover Page
Slide 26 Data Science for NOAA Big Data: Spotfire Data Ecosystem
Slide 27 Data Science for DHS: Global Terrorism Database Knowledge Base
Slide 28 Data Science for DHS: Big Data Symposia Knowledge Base
NOAA embraces the business of big data

Government Data Hubs

**Department of Commerce Data Sets and Information for Developers**

Complete catalogue of publicly-available Commerce data sets

- Topics
- Topic Categories
- Dataset Type
- Tags
- Formats
- Organization Types
- Organizations
- Publisher

**Department of Commerce Developer Application Programming Interfaces (APIs)**

- Welcome
- Bureau of Economic Analysis
- Census Bureau
- International Trade Administration
- National Institute of Standards and Technology
- National Oceanic and Atmospheric Administration
- National Telecommunications and Information Administration
- Patent and Trademark Office

**NOAA Data Catalog**

- Dataset Type
- Tags
- Formats
- Communities
- Organization Types
- Organizations
- Community Categories

**Big Data Industry Day Announcement Email**

**Big Data Industry Day Announcement**

- Synopsis
- Big Data Industry Day Information

**Big_Data_Industry_Day_Cover_Memo.pdf (228.78 Kb)**
NOAA announces RFI to unleash power of 'big data'

Big Data RFI from NOAA

Synopsis
Background
Purpose of this Request for Information (RFI)
Information Requested
1. Value Proposition
2. Data Infrastructure
3. Data Services
4. Data Management
5. Data Security
6. Partnership Methods
7. Capabilities
Instructions
Contracting Office Address
Primary Point of Contact
Secondary Point of Contact

Story

Data Science for NOAA Big Data

The NOAA [Big Data Industry Day Announcement Email](mailto:Data_Science_for_NOAA_Big_Data) and [Big Data Industry Day Announcement](mailto:Data_Science_for_NOAA_Big_Data) says:

This is the second RFI (Request For Information done for market research) released by NOAA’s Office of Acquisition and Grants (AGO) with regards to the NOAA Big Data Initiative. RFI number ST-1330-14-RP-0039 was posted to the Federal Business Opportunities Website on February 21, 2014. This RFI is focused on the draft Statement of Objectives (SOO). At the Big Data Industry Day on October 17, NOAA will review the draft SOO and offer industry an opportunity to ask questions. All comments and suggestions to the SOO will be collected in the excel comment and suggestion sheet.
included in the Federal Business Opportunities posting and will be due on October 24, 2014 by 3:00 PM EDT. Any presentations and notes provided will be posted to the Federal Business Opportunities website by October 21, 2014.

The Big Data Industry Day Cover Memo says:

The Big Data Initiative is an exciting project that will require industry to be creative and innovative. We hope that this initiative will create jobs and spur economic growth while allowing NOAA to be as transparent as possible.

1. Is a no-cost contract a viable business arrangement for this SOO? **My Answer: Yes**

   a. What risks or issues do respondents anticipate encountering in building this business model? **My Answer: Lack of full cooperation from and participation by NOAA.**
   
   b. Do the requirements specified in this SOO enable respondents to build a business model to recover costs? **My Answer:** Their "costs" are covered by business development, overhead, and community service.
   
   c. If you propose to recover your cost, is your cost efficiency in providing access an appropriate criterion for government selection? **My Answer: We do not propose to recover your cost.**

2. The Government anticipates multiple partners, selected based upon a “best value” evaluation.

   a. What may limit potential solutions in a “best value” evaluation, specifically in the areas of:

      i. Inciting creative uses and innovative approaches to spur economic growth; **My Answer: An open community of data scientists**
      
      ii. Ensuring equal access on equal terms; **My Answer: An open community based upon sound open community principles.**
      
      iii. Supporting NOAA's compliance with the US Open Data Policy (Executive Order 13642 and OMB Memorandum M-13-13); **My Answer: Understanding of and experience with implementing the Open Data Policy with Open Government Data.**
      
      iv. Providing a self-sustaining business model; and **My Answer: A well-run Community of Practice with ongoing cooperation and participation by NOAA.**
      
      v. Data-handling and data curation policies for extracted data. **My Answer: A Data FAIRport/Data Commons like the Semantic Community Sandbox.**

   b. What recommendations do interested parties have for managing this communication? **My Answer: Massive Open Online Communities (MOOCs) like the Federal Big Data Working Group Meetup. See for example ESIP Earth Sciences Data Analytics.**

The original RFI drew 70 responses from individuals, academia and industry organizations before closing March 31. But NOAA’s market research is ongoing (see NOAA embraces the business of big data);

"It gives you a good idea of what they see as a potential for value," said David McClure, lead analyst for open government services at NOAA and the man behind the agency's big data partnership business model, which built on the longtime efforts of McClure and former NOAA CIO Joe Klimavicz.
Besides providing a useful service, another motivation behind the effort is so NOAA can recover the costs it incurs when opening up its data sets. When President Barack Obama signed the open data executive order in May 2013, no additional funds were provided to support the effort. Money is being spent to open up government data, but nothing was being done to cover that cost.

In the NOAA big data partnership business model that McClure conceptualized, there would be a small public fee for accessing the data. The cost would be minimal -- in most cases zero. People would not be paying for the data, but for accessing the data -- a fine line, McClure acknowledged.

He said the process will give NOAA an opportunity to spend more time enriching its data, providing more value to the public and, theoretically, boosting NOAA’s return on investment.

The data mining process, conclusions, and recommendations are (see slides below):

- **Data Mining**:
  - Start with Government Data Hubs and DoC/NOAA Data Catalogs (see Slide 10 Start: Government Data Hubs);
  - Build a Knowledge Base (see Slide 19 Data Science for NOAA Big Data: Build Knowledge Base);
  - Build a Spreadsheet Index of the Knowledge Base (see Slide 21 Data Science for NOAA Big Data: Spreadsheet Knowledge Base Index); and
  - Import the Spreadsheets into Spotfire and Build a Data Science Publication (see Slide 25 Data Science for NOAA Big Data: Spotfire Cover Page).

- **Conclusions**:
  - The Government Data Hubs shows the Complete catalogue of publicly-available Commerce data sets has 22,365 data sets, of which 3,560 are NOAA, and the Department of Commerce Developer Application Programming Interfaces (APIs) have 7 agencies with APIs, with 11 from National Oceanic and Atmospheric Administration;
  - Specifically: Environmental Research Division's Data Access Program is a data server that gives you a simple, consistent way to download subsets of scientific datasets in common file formats and make graphs and maps;
  - This NOAA Data Catalog is a prototype under active development. Availability and completeness are not guaranteed. 55,602 data sets, but no SHP and only one CSV and 222 Excel!
  - The data catalogs show considerable diversity in data formats making data integration very challenging;
  - The NSF/NIH Data FAIRport/Commons criteria of F (findable), A (accessible), I (Interoperable), and R (reusable) are F(yes), A (yes), I (somewhat), and R (challenging) for the NOAA data ecosystem looked at in this pilot;
  - The three NOAA Data Hubs are a good start for building a Data FAIRport of data science publications.

- **Recommendations**:
  - NOAA should have a chief data officer (like the Department of Commerce) that leads a team of data scientists that build an NSF/NIH-like DataFAIRport/Commons; and
  - The Federal Big Data Working Group Meetup is a Community of Data Science Practice that can help a NOAA Chief Data Officer and a Team of Data Scientists make the data more interoperable and package it in data science publications like this pilot example for our November 3rd Meetup.
<table>
<thead>
<tr>
<th>Reference Document (SOO, Eval Criteria, General)</th>
<th>Paragraph Number</th>
<th>Question or Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data#Story">http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data#Story</a></td>
<td></td>
<td>NOAA should have a chief data officer (like the Department of Commerce) that leads a team of data scientists that build DataFAIRport/Commons</td>
</tr>
<tr>
<td><a href="http://www.meetup.com/Federal-Big-Data-Working-Group/events/213175262/">http://www.meetup.com/Federal-Big-Data-Working-Group/events/213175262/</a></td>
<td></td>
<td>The Federal Big Data Working Group is a Community of Data Science Practice that can help a NOAA Chief Data Officer and Team of Data Scientists make the data more interoperable and package it in data science publications like this pilot November 3rd Meetup</td>
</tr>
</tbody>
</table>

---

**Slides**

**Slide 1 Data Science for NOAA Chief Data Officer and Big Data Analytics Meetup**

[http://semanticommunity.info/](http://semanticommunity.info/)

---

**Slide 2 Agenda**

[http://www.meetup.com/Federal-Big-Da...nts/213175262/](http://www.meetup.com/Federal-Big-Da...nts/213175262/)
Agenda

- 6:30 p.m. Welcome and Introduction (New Tutorial and Mentoring) Slides
  - Background: Data Science for NOAA Big Data and DHS Global Terrorism Database. See Big Data Symposium Google Find. Department of Homeland Security and see Recent Presentation below
  - Data Science for the NOAA Chief Data Officer Story and Slides
- 7:00 p.m. Brief Member Introductions
- 7:10 p.m. Treeminer.com Video, Mark Silverman and Biplab Pal
  - Background: Data Science for Vertical Data Mining
  - Video Demo 1: HEC classification: https://docs.google.com/file/d/0B8SY...89p0rT0/edit and Demo 2: How Treeminer works for document classification reference to a patent invalidation: https://docs.google.com/file/d/0B8SY...3M8aTYM3/edit
- 7:30 p.m. Predictive Analytics in the Era of Big Data, Dave Vennersgrind, Director, Data Analytics Center of Excellence, SalientFed Slides
- 8:30 p.m. Open Discussion
- 8:45 p.m. Networking
- 9:00 p.m. Depart

Slide 3 Calendar

Calendar

- October 23, Big Data Symposium at the National Research Council, National Academy of Sciences Keck Building, Room 100, 500 Fifth Street, NW, Washington, DC 20001. 8:45 a.m.
  - Contact the Board director, Paul Uhli at puhi@email.com or at 1 202 334 1531, to register in advance.
  - https://mitprofessionalx.org/courses/MITProfessionalX/6_BDX/2T2014AJbUp
- November 4 - “Diverse Data Analytics Applications” a joint George Mason University and IBM ASC Symposium.
  - Register: http://ibm.com/ascd

Slide 4 FDA Data Innovation Lab and Predictive Analytics Meetup

http://www.meetup.com/Federal-Big-Data-Working-Group/events/209068792/
FDA Data Innovation Lab and Predictive Analytics Meetup

- Great discussion and very informative.
- The meeting was extraordinary - an onion-peeling exercise in which discussions stripped away layers of issues and led to what appears to be basics of predictive analytics and data science.
- Good discussion on FDA and open data. Lots of potential value, and much remaining to be done. Also enjoyed the predictive analytics discussions.
- The formatted data is truly helpful.

http://www.meetup.com/Federal-Big-Data-Working-Group/events/201107182/

Slide 5 Conference for NSF Data Scientists, Data Infrastructure, and Data Publication

http://www.meetup.com/Federal-Big-Data-Working-Group/events/201107182/

Conference for NSF Data Scientists, Data Infrastructure, and Data Publication

- Excellent content, very informative!
- A hearty thank you on behalf of the Digital Government Institute event team for your presentation and participation in the 2014 Government Big Data Conference. We very much appreciate your willingness to share with the government audience today – I am sure the attendees thoroughly enjoyed your presentation and the feedback will be extremely positive. DGI is very fortunate to have had the opportunity to work with the FBDWG team – your participation is appreciated.
- Thank you all very much! I had so many great conversations and contacts at the event.

http://www.meetup.com/Federal-Big-Data-Working-Group/events/201107182/

Slide 6 Semantic Insights Followup

http://www.semanticinsights.com/

Applying High-speed Pattern Recognition to Generate Queryable Semantic from Big Data
Semantic Insights Followup

- Looking for interested individuals who wish to participate in our Natural Language Understanding and Reasoning research. We welcome educational institutions and individual researchers interested in working collaboratively with us. Accounts are available for beta test:
- Applying High-speed Pattern Recognition to Generate Queryable Semantics from Big Data - Big Data is filtered and reduced in real-time for event and pattern discovery:
  - [Applying High-speed Pattern Recognition to Generate Queryable Semantics from Big Data](http://fcw.com/Articles/2014/06/24/NOAA-embraces-big-data.aspx?p=1)

---

Slide 7 NOAA Embraces the Business of Big Data


NOAA Embraces the Business of Big Data

- NOAA February 2014 RFI seeking solutions for how to make its 20 terabytes of daily data available quickly and at scale.
  - The request for information drew 70 responses from individuals, academia and industry organizations.
- Now sharing only about 10 percent of that information, NOAA wanted to hear about ways to get more information into the hands of users – and maybe make a little money on the side.
  - "It gives you a good idea of what they see as a potential for value," said David McClure, lead analyst for open government services at NOAA and the man behind the agency's big data partnership business model.

Slide 8 NOAA Big Data Partnership Model

NOAA Big Data Partnership Model

Slide 9 NOAA Big Data Industry Day

NOAA Big Data Industry Day

- At the Big Data Industry Day on October 17, NOAA will review the draft Statement of Objectives (SOO) and offer industry an opportunity to ask questions.
- All comments and suggestions to the SOO will be due on October 24, 2014.
- My response (see Story) is that we can do this RFI in a Community of Data Science Practice like the Federal Big Data Working Group Meetup.

Slide 10 Start: Government Data Hubs

http://project-open-data.github.io/data-hubs/
Slide 11 Click To: Three NOAA Data Hubs

Click To: Three NOAA Data Hubs

- The CHAMP Program gathers near real-time data from instrumented arrays and satellites covering important coral reef areas from all over the globe. Please choose an area of interest from the Data drop-down menu.
- A primary focus of U.S. IOOS is integration of, and expedited access to, ocean observation data for improved decision making. The Data Management and Communication (DMAC) subsystem of U.S. IOOS serves as a central mechanism for integrating all existing and projected data sources.
- The MarineCadastre.gov Data Registry provides direct access to data currently available through MarineCadastre.gov. Filter the data by provider, thematic category, geographic region, and service type. If you are looking for a data set that is currently not available on MarineCadastre.gov, please email us.

Slide 12 Click To: Department of Commerce

http://www.commerce.gov/data
Slide 13 Click To: Data.gov Department of Commerce

Data.gov Department of Commerce

Slide 14 Click To: Department of Commerce APIs

http://www.commerce.gov/developer
Click To:
Department of Commerce APIs

http://www.commerce.gov/developer

---

Slide 15 Also Filter To: Data.gov NOAA

Data.gov for NOAA

Also Filter To: Data.gov NOAA

---

Slide 16 Also See Prototype: data.noaa.gov

https://data.noaa.gov/dataset
Slide 17 Filter to Excel in Prototype: data.noaa.gov

https://data.noaa.gov/dataset?res_format_limit=0

Slide 18 Also See: Environmental Research Division's Data Access Program
RESTful Web Services

http://coastwatch.pfeg.noaa.gov/erddap/rest.html

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web
Also See: Environmental Research Division’s Data Access Program RESTful Web Services

http://coastwatch.pfeg.noaa.gov/erdap/rest.html
http://semanticommunity.info/graphics/Files/21083/boxC9GAddigData-Ref.xlsx?origin=rest-web

Slide 19 Data Science for NOAA Big Data: Build Knowledge Base

Data Science for NOAA Big Data

Data Science for NOAA Big Data: Build Knowledge Base
Slide 20 Data Science for NOAA Big Data: Knowledge Base Contents

Data Science for NOAA Big Data: Knowledge Base Contents

- 6.1. Department of Commerce Data Sets and Information for Developers
  - 6.1.1. Complete catalogue of publicly-available Commerce data sets (EvC 12165 and NOAA 3160)
  - 6.1.2. Topics
  - 6.1.3. Theme Categories
  - 6.1.4. Data Set Type
  - 6.1.5. Tags
  - 6.1.6. Industry
  - 6.1.7. Organization Names
  - 6.1.8. Publisher

- 6.2. Department of Commerce Developer Application Programming Interfaces (APIs)
  - 6.2.1. Welcome
  - 6.2.2. Bureau of Economic Analysis
  - 6.2.3. Census Bureau
  - 6.2.4. International Trade Administration
  - 6.2.5. National Institute of Standards and Technology
  - 6.2.6. National Oceanoic and Atmospheric Administration (11 APIs)
  - 6.2.7. National Telecommunications and Information Administration
  - 6.2.8. Patent and Trademark Office

Slide 21 Data Science for NOAA Big Data: Spreadsheet Knowledge Base Index

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web

Slide 22 Data Science for NOAA Big Data: Spreadsheet Data.gov DoC Index

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web
Data Science for NOAA Big Data: Spreadsheet Data.gov NOAA Index

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web

Data Science for NOAA Big Data: Spreadsheet Data.gov NOAA Index

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web

Slide 24 Data Science for NOAA Big Data: Spreadsheet data.NOAA.gov Index

http://semanticommunity.info/@api/deki/files/31083/DoCNOAABigDataRFI.xlsx?origin=mt-web
Slide 25 Data Science for NOAA Big Data: Spotfire Cover Page

Web Player

Slide 26 Data Science for NOAA Big Data: Spotfire Data Ecosystem

Web Player
Slide 27 Data Science for DHS: Global Terrorism Database Knowledge Base

Global Terrorism Database

Slide 28 Data Science for DHS: Big Data Symposia Knowledge Base

Big Data Symposia
Slide 29 Global Terrorism Database Experience: Spotfire Cover Page

Web Player

Slide 30 Agenda

http://www.meetup.com/Federal-Big-Da...nts/213175262/
Research Notes

NOAA Request for Information on Increasing Access to Environmental Data

To increase access to and use of its vast library of environmental data, NOAA has issued a request for information (RFI) seeking comment from industry, non-profits, research laboratories, universities, and private-sector partners to help make NOAA’s vast data holdings available in a rapid, scalable manner to the public. Of the 20 terabytes of data NOAA gathers each day only a small percentage is easily accessible to the public. Through the RFI, American companies will be able to provide potential solutions for NOAA to turn this untapped information into usable products or services. Respondents have until March 24, 2014, to submit a written statement of interest, including a proposed way forward.

Esri: Providing Communities with Map-Based Planning Tools and Collaboration Platforms

Esri is unveiling a new two-part initiative to help communities more effectively build climate-resilience. First, Esri will develop and publish a series of free and open "maps and apps" developed in partnership with 12 cities that help address the most urgent climate-relevant needs shared among thousands of users of Esri’s ArcGIS platform—such as preparing for droughts, heat waves, or flooding. Second, Esri is announcing today a climate-focused geo-collaboration portal—an online destination to discover, contribute, and share resources critical to confronting the impacts of climate change. Additionally, Esri recently announced, during one of the largest gatherings of GIS developers and in response to the President’s call to action, a Climate Resilience App Challenge to inspire more than 2,500 developers to focus their
creative attention on creating mapping and analytical tools that help communities see, understand, and prepare for climate risks. Prizes will be awarded and the resulting apps will be openly shared in July.

Google Search for NOAA SHP
http://www.nws.noaa.gov/gis/shapepage.htm
http://www.nws.noaa.gov/regsci/gis/shapefiles/
http://www.nws.noaa.gov/regsci/gis/s...t_warnings.zip

---

NOAA embraces the business of big data

Source: http://fcw.com/Articles/2014/06/24/N...-data.aspx?p=1

By Colby Hochmuth

Jun 24, 2014

The National Oceanic and Atmospheric Administration generated some buzz with its February RFI seeking solutions for how to make its 20 terabytes of daily data available quickly and at scale. Now sharing only about 10 percent of that information, NOAA wanted to hear about ways to get more information into the hands of users – and maybe make a little money on the side.

"We wanted information, we wanted people's ideas -- 'what would you do, how would you expose the data, how would you unlock it, what would that cost and how could we make money from this,'" acting NOAA CIO Zach Goldstein said at a June 19 event hosted by the Office of Science and Technology Policy in Washington, D.C.

The request for information drew 70 responses from individuals, academia and industry organizations before closing March 31. But NOAA's market research is ongoing.
"It gives you a good idea of what they see as a potential for value," said David McClure, lead analyst for open
government services at NOAA and the man behind the agency's big data partnership business model, which built on the
longtime efforts of McClure and former NOAA CIO Joe Klimavicz.

McClure said the message was clear: "Start doing it." Which is why his team is now working to roll out the concept "as
fast as possible."

Besides providing a useful service, another motivation behind the effort is so NOAA can recover the costs it incurs when
opening up its data sets. When President Barack Obama signed the open data executive order in May 2013, no
additional funds were provided to support the effort. Money is being spent to open up government data, but nothing was
being done to cover that cost.

In the NOAA big data partnership business model that McClure conceptualized, there would be a small public fee for
accessing the data. The cost would be minimal -- in most cases zero. People would not be paying for the data, but for
accessing the data -- a fine line, McClure acknowledged.

He said the process will give NOAA an opportunity to spend more time enriching its data, providing more value to the
public and, theoretically, boosting NOAA's return on investment.

"We're not asking for a donation from private industry, we're asking them to invest," McClure said.

The rollout is incremental, but McClure said demand is high, with industry coming to NOAA with specific dataset
requests.

"If we can connect to that dataset, we already have customers lined up," McClure said.

Part of the business model is that NOAA agrees to keep improving both the data and data mining process. Finding
better processes -- the "better, faster, cheaper," way of doing things -- will result in additional savings, McClure said, and
serve as a model for other agencies interested in commercializing their data.

About the Author

Colby Hochmuth is a staff writer covering big data, cloud computing and the federal workforce. Connect with her on
Twitter: @ColbyAnn.
Government Data Hubs

http://project-open-data.github.io/data-hubs/

Department of Commerce

- International Trade Administration
- Coral Health and Monitoring Program - NOAA
- Integrated Ocean Observing System - NOAA
- Marine Cadastre - NOAA

Department of Commerce Data Sets and Information for Developers

Source: http://www.commerce.gov/data

The Department of the Commerce is committed to the President's Open Government Initiative and has submitted high value datasets in support of the Open Government Directive. These data sets span a wide range of functions of Commerce and represent areas of continuing interest by our stakeholders.

Learn more about our data:
- Complete catalogue of publicly-available Commerce data sets
- List of Department of Commerce Developer Application Programming Interfaces (APIs)

Questions or input for our data? Please use our U.S. Department of Commerce Open Government Portal.

Learn more about the Department of Commerce's Digital Strategy.

Complete catalogue of publicly-available Commerce data sets

22,365 datasets found for "Department of Commerce"

Topics
- AAPI (522)
• Education (224)
• Energy (183)
• Safety (183)
• Research (158)
• Climate (131)
• Ocean (120)
• Agriculture (106)
• Law (80)
• Consumer (66)
• Finance (65)
• BusinessUSA (60)
• Disaster (32)
• Ecosystems (22)
• Manufacturing (17)
• World Wide Human Ge... (13)

Topic Categories
• Asian (205)
• Criminal Justice (199)
• Total Energy (130)
• Hawaii (126)
• Pacific Islands (126)
• Civil Rights (89)
• Applied Science & T... (88)
• Guam (84)
• Pacific Islander (83)
• Education (76)
• Environment (60)
• Food Resilience (54)
• Ecosystem Vulnerabi... (51)
• K-12 (50)
• Elevation and Bathy... (41)
• Native Hawaiian (39)
• Compliance Violatio... (38)
• Physical and Oceano... (38)
• American Samoa (37)
• Corrections Data (37)
• Employment/Labor (37)
• Water (37)
• Northern Mariana Is... (34)
• State (34)
• Incidents and Crashes (33)
• Electricity (30)
• Military (29)
• Housing and Community (28)
• Higher Education (28)
• Agriculture & Food (27)
• Food Production (24)
• Veterans (22)
• Atmospheric (22)
• Natural Resources a... (22)
• Earth & Ocean Sciences (21)
• Health (21)
• Product Safety (20)
• Social & Behavioral... (20)
• Biodiversity (19)
• Biology and Habitats (18)
• Finance (18)
• Natural Hazards (18)
• Natural Resources a... (17)
• Workforce Diversity (16)
• Food Safety and Nut... (16)
• Atmospheric, Earth ... (15)
• Fatalities Casualti... (14)
• Exposure Data (13)
• Farm Damage (12)

Dataset Type
• geospatial (13285)
• non-geospatial (9080)

Tags
• oceans (2619)
• geoscientificinform... (2002)
• inlandwaters (1902)
• structure (1746)
• boundaries (1732)
• temperature (1502)
• noaa (1473)
• profile (1442)
• physical (1437)
• salinity (1429)
• earth science (1408)
• unknown (1373)
• wind direction (1307)
• wind speed (1303)
• barometric pressure (1298)
• aadt (1273)
• department of fishe... (1270)
• bathythermograph - xbt (1259)
• inspection (1246)
• nbis (1229)
• bridge (1228)
• nbi (1225)
• national-bridge-ins... (1224)
• structurally-deficient (1224)
• national-bridge-inv... (1224)
• functionally-obsolete (1224)
• deck (1224)
• global temperature-.... (1216)
• sea surface tempera... (1203)
• buoy - moored buoy (1200)
• air temperature - d... (1197)
• array for real-time... (1197)
• profiling floats (1196)
• relative humidity (1194)
• current direction (1188)
• current speed (1184)
• usgs (1166)
• meteorological (1160)
• u.s. geological survey (1111)
• environment (1070)
• coastal and marine ... (1067)
• cmgp (1064)
• time series (1058)
• buoy - drifting buoy (1050)
• world-wide distribu... (1047)
• tropical atmosphere... (1026)
• shoreline (1005)
• u.s. department of ... (1001)
• united states (998)

Formats

• XML (4836)
• HTML (4703)
• CSV (4182)
• ZIP (4177)
• JSON (3668)
• application/rdf+xml (3007)
• PDF (1482)
• TXT (1191)
• Excel (917)
• XLS (632)
• WMS (588)
• KML (523)
• Esri REST (471)
• WFS (461)
• application/vnd.lot... (442)
• gml (402)
• TAR (356)
• XYZ (291)
• application/rss+xml (202)
• application/octet-s... (81)
• application/unknown (75)
• Access (63)
• KMZ (53)
• application/x-sas (53)
• QGIS (52)
• shapefile (51)
• TIFF (40)
• raster digital data... (34)

http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data
Updated: Wed, 23 Sep 2015 08:08:44 GMT
Powered by mindtouch
• application/excel (34)
• RDF (31)
• xlsx (31)
• application/x-spss (29)
• other (27)
• OpenXML (24)
• personal geodatabase (24)
• ascii (23)
• application/xslt+xml (21)
• DOC (18)
• esri grid (17)
• kongsberg/simrad in... (17)
• WCS (16)
• application/x-stata (14)
• ArcGIS Online Map (13)
• winzip (13)
• application/doc (12)
• application/x-troff... (12)
• application/atom+xml (11)
• ArcGIS Map Preview (10)
• ArcGIS Map Service (10)

Organization Types
• Federal Government (13171)
• State Government (3656)
• State (3333)
• University (1499)
• City Government (616)
• County Government (48)
• Non-Profit (3)
• University (3)
• Other (1)

Organizations
• National Oceanic an... (3560)
• NSGIC GIS Inventory... (3333)
• U.S. Geological Sur... (2331)
• Kansas Data Access ... (69)
• Vermont Center for ... (60)
• Idaho State University (56)
• Department of the T... (54)
• City of Los Angeles (50)
• City of Philadelphia (42)
• Cook County of Illi... (39)
• City of Baltimore (22)
• International Trade... (20)

Publisher
• Federal Highway Adm... (1290)
• BJS - Bureau of Just... (650)
• Department of Veter... (605)
• Research and Innova... (568)
• data.illinois.gov (541)
• data.hawaii.gov (537)
• data.oregon.gov (528)
• data.ok.gov (514)
• Department of Defense (338)
• data.cityofnewyork.us (312)
• Department of Energy (179)
• U.S. Department of ... (156)
• US Department of Ho... (149)
• U.S. Fish and Wildl... (134)
• data.sfgov.org (126)
• data.mo.gov (114)
• U.S. Department of ... (113)
• U.S. Department of ... (106)
• Pipeline and Hazard... (104)
• data.cityofchicago.org (86)
• Farm Service Agency... (84)
• data.maryland.gov (73)
• Department of Healt... (68)
• National Highway Tr... (62)
• Office of the Secre... (62)
• Agricultural Market... (59)
• Economic Research S... (52)
Department of Commerce Developer Application Programming Interfaces (APIs)

Source: [http://www.commerce.gov/developer](http://www.commerce.gov/developer)

Welcome

On Commerce's Developer Resource page you can find and access featured data or web services from the Department's agencies and mission areas. This page will be updated as agencies make available new developer tools and data resources, so be sure to check back often. Questions and feedback about this page can be directed to Webmaster@doc.gov.

Bureau of Economic Analysis

[Bureau of Economic Analysis API](http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data) - The API provides all regional data and national economic data released by BEA.
Census Bureau

Census Bureau API - The API provides a range of popular data sets through a combined endpoint. The following data sets are currently available through the API: 2010 Census Summary File 1 (SF1), American Community Survey 5 Year Data, 2000 Census Summary File 1 (SF1) and Summary File 3 (SF3), 1990 Census Summary File 1 (SF1) and Summary File 3 (SF3).

Automated Export System - This API enables Internet application developers and providers to allow their users to pass data from their own third party applications into AESDirect. This makes it easier for end-users to enter data into AESDirect and eliminate double entry of data into both the third party application and AESDirect.

International Trade Administration

Export Trade Events - Events can be retrieved via http in html, json, and xml formats. To specify a format, json, for instance, add .json to the end of the url and before the query string. Events can also be filtered according to parameters such as term, start date, end date, source, countries, states and industries. Learn more at the BusinessUSA API developers hub.

National Institute of Standards and Technology

CDA Guideline Validation Tool - In addition to an on-line HTML formed-based submission page, the CDA Guideline Validation Tool is also available as a SOAP-based Web Service. By following the description of the web service contained in the WSDL, applications will be able to automatically query the service for validation of CDA-based documents.

National Oceanic and Atmospheric Administration

Meteorological Assimilation Data Ingest System - The MADIS Application Program Interface (API) is a library of subroutines, callable from Fortran, which provides access to all of the observation and quality control (QC) information in the MADIS database (referred to as the “FSL database” in the API documentation), and/or other supported meteorological databases (i.e., the NWS AWIPS netCDF database)

Legacy National Climatic Data Center RESTful Services - Station time-series data

GIS Map Services - Station and network locations, metadata search, climatological GIS products.

Severe Weather Data Inventory - NEXRAD Storm Cell Attributes, Storm Reports.

General National Climatatic Data Center THREDDS Data Server - Satellite/Radar/Model Data.

National Climatic Data Center Climate Data Online Web Services - Climate Data Online is a collection of climatic data that offers public access and consumption via discovery and ordering services.

National Digital Forecast Database SOAP Web Service - National Digital Forecast Database (NDFD) Extensible Markup Language (XML) is a service providing the public, government agencies, and commercial enterprises with data from the National Weather Service’s (NWS) digital forecast database.
**Center of Operational Oceanographic Products and Services** - CO-OPS SOAP Web Services were developed using industry standards. Each SOAP web service is offered with a sample request, sample response and sample Java Client code to our users to facilitate their work with seamlessly connecting to the services and to retrieve the data of interest.

**Center for Operational Oceanographic Products and Services** - The Center for Operational Oceanographic Products and Services (CO-OPS) Sensor Observation Service has implemented three operations that provide access to sensor observations and measurement data via a spatio-temporal query that can be filtered by phenomena, provide the means to access SOS service metadata, retrieve detailed information about the sensors and processes generating those measurements.

**Environmental Research Division’s Data Access Program** - The Environmental Research Division's Data Access Program is a data server that gives you a simple, consistent way to download subsets of scientific datasets in common file formats and make graphs and maps.

**Space Physics Interactive Data Resource Web Services** - Space Physics Interactive Data Resource (SPIDR) is designed to allow a solar terrestrial physics customer to intelligently access and manage historical space physics data for integration with environment models and space weather forecasts. SPIDR is a distributed network of synchronous databases and 100% Java middle-ware servers accessed via the World Wide Web.

**National Telecommunications and Information Administration**

National Broadband Map - Datasets are available for use. There is also a Developer page with tools, tricks and a showcase.

**Patent and Trademark Office**

**Trademark Status & Document Retrieval** - TSDR (Trademark Status Document Retrieval) is a web application that provides real-time access to the electronic file wrapper of U.S. Trademark applications and applications for Extensions of Protection, as well as U.S. Trademark Registrations. It also displays information contained in the USPTO records regarding International Registrations and applications for International Registration filed under the Madrid system through the U.S.A.

**NOAA Data Catalog**

Source: [https://data.noaa.gov/dataset](https://data.noaa.gov/dataset)

This NOAA Data Catalog is a prototype under active development. Availability and completeness are not guaranteed.

**Dataset Type**

- geospatial (55601)

**Tags**

- earth_science (24344)
• oceans (23515)
• oceanography (23278)
• noaa (19673)
• bathymetry (19665)
• nesdis (18403)
• mbes (18360)
• sonar (18335)
• doc/noaa/nesdis/ngdc (18216)
• seafloor topography (17895)
• water depth (17847)
• gps (17667)
• ships (17635)
• in situ/laboratory ... (17633)
• in situ ocean-based... (17628)
• profilers/sounders (17583)
• acoustic sounders (17583)
• earth remote sensin... (17558)
• msbs (17554)
• positioning/navigation (17553)
• gps/imu (17553)
• marine sediments (15073)
• sediment composition (15038)
• u.s. department of ... (15017)
• sound ... (15017)
• passive ... (15017)
• national geophysica... (15017)
• multibeam ... (15017)
• multibeam ... (15017)
• hydrographic survey... (15017)
• gps ... (15017)
• global ... (15017)
• bathymetry/seafloor... (15017)
• biosphere (5104)
• bathymetry/seafloor... (4403)
• corals (4254)
• zoology (4222)
• reef monitoring and... (3929)
• coastal processes (3760)
• u.s. department of ... (3742)
• united states of am... (3501)
• country/territory (3489)
• ocean basin (3488)
• coral reefs (3464)
• numeric data sets (3320)
• alaska (3242)
• national geophysica... (3198)
• pacific ocean (3025)
• multibeam mapping s... (2589)

Formats

• HTML (55327)
• XML (15062)
• PDF (11198)
• TIFF (10924)
• XYZ (7585)
• JPEG (2853)
• ZIP (1365)
• TXT (464)
• Esri REST (358)
• WMS (237)
• Excel (222)
• KML (88)
• KMZ (46)
• OpenXML (14)
• RAR (13)
• TAR (10)
• DOC (4)
• WFS (3)
• application/x-troff (2)
• audio/basic (1)
• CSV (1)

Communities

There are no Communities that match this search
You are receiving this e-mail as a courtesy on behalf of the NOAA Small Business Office. You have previously registered in our database as a vendor that is interested in working with us. In our on-going effort to ensure small businesses are given fair consideration for each of our requirements we would like to direct your attention to an Industry Day that we have scheduled for October 17, 2014 for Big Data. There is little available additional information at this time but please feel free to reach out to the Contracting Officer or Specialist listed in the announcement on FBO. The announcement can be seen by either clicking on the following link: https://www.fbo.gov/index?s=opportun...=core&_cview=0 or if you are uncomfortable clicking on a direct link simply go to FBO.gov and search for the solicitation number “ST-1330-14-RP-0039”. You should also be able to find the previous RFI that was posted on Federal Business Opportunities in the Archive of that site for this requirement. Should you have any questions or concerns I can help with please let me know, otherwise please direct your questions to the Contract Specialist. I look forward to seeing you at the event later this month!

Thank you

Small Business Office
Acquisition and Grants Office
National Oceanic and Atmospheric Administration (NOAA)

Register to do business with NOAA: http://www.ago.noaa.gov/quicklinks/s...rofile_reg.htm

Solicitation Number: ST-1330-14-RP-0039
Agency: Department of Commerce
Office: National Oceanic and Atmospheric Administration (NOAA)
Location: Acquisition and Grants Office
Big Data Industry Day Announcement

Source: https://www.fbo.gov/index?s=opportun...tabmode=list&=

Solicitation Number:
ST-1330-14-RP-0039
Notice Type:
Special Notice

Synopsis

Added: Sep 26, 2014 10:57 am
The purpose of this notice is to announce the government's intention to host an industry day for NOAA's Big Data initiative on October 17, 2014 at the NOAA Science Center located at 1301 East West Highway in Silver Spring, MD.

Additional Information will be provided at later date.
Contracting Office Address:
SSMC2 - 11th floor /OFA61
1325 East West Highway, 11th Floor
Silver Spring, Maryland 20910
United States
Place of Performance:
NOAA Science Center
1301 East West Highway
Silver Spring, Maryland 20910
United States
Primary Point of Contact:
Patrick Curry,
Contract Specialist
patrick.curry@noaa.gov
Phone: 3016281386
Secondary Point of Contact:
TANIA GATES,
Contracting Officer
TANIA.GATES@NOAA.GOV
Phone: 301-628-0024
Fax: 301-713-1024

Big Data Industry Day Information

Source: https://www.fbo.gov/index?tab=docume...155e2b0a4088a4

Type: Other (Draft RFPs/RFIs, Responses to Questions, etc..)
Label: Big Data Industry Day Information

Big Data Industry Day Cover Memo.pdf (228.78 Kb)

PDF
TO: All Interested Responders

FROM: Tania Gates, Contracting Officer

SUBJECT: REQUEST FOR INFORMATION (RFI) #10172014 – Comments to draft Statement of Objectives and Evaluation Criteria

This is the second RFI released by NOAA’s Office of Acquisition and Grants (AGO) with regards to the NOAA Big Data Initiative. RFI number ST-1330-14-RP-0039 was posted to the Federal Business Opportunities Website on February 21, 2014. This RFI is focused on the draft Statement of Objectives (SOO). At the Big Data Industry Day on October 17, NOAA will review the draft SOO and offer industry an opportunity to ask questions. All comments and suggestions to the SOO will be collected in the excel comment and suggestion sheet included in the Federal Business Opportunities posting and will be due on October 24, 2014 by 3:00 PM EDT.

If you would like to register to attend the industry day on October 17, please email Patrick Curry, Contract Specialist at patrick.curry@noaa.gov no later than October 15, 2014 at 12:00 PM EDT. When submitting an email to register for the event, please include “Big Data Industry Day Registration” in the subject line of the email and include your organization’s name, the names of the people attending, and their email address and phone number.

Due to seating limitations, there will be a limit of two attendees per organization. An agenda for the Industry Day is included as Attachment A. There will not be remote access available for the industry day, though any presentations and notes provided will be posted to the Federal Business Opportunities website by October 21, 2014. The NOAA building is located within walking distance of the Silver Spring metro station on the red line. Parking at NOAA is limited, but you are welcome to park in the NOAA parking garage. The Wayne Avenue parking garage is also available as parking in the NOAA parking garage is limited.

When reviewing the documents, NOAA asks that you consider the following challenges and keep these in mind when commenting and as we move forward in the process. These challenges serve as a framework for the larger problem we are trying to solve with this initiative. The expectation from NOAA is that by providing the requirements early in the process we can engage with industry and solicit their help for feedback.

1. Is a no-cost contract a viable business arrangement for this SOO?
   a. What risks or issues do respondents anticipate encountering in building this business model?
   b. Do the requirements specified in this SOO enable respondents to build a business model to recover
c. If you propose to recover your cost, is your cost efficiency in providing access an appropriate criterion for government selection?

2. The Government anticipates multiple partners, selected based upon a “best value” evaluation.

a. The Government is determining how to qualitatively assess the technical, business, and risk elements of proposals. Given the innovative nature of the potential partnerships, the Government is seeking insights to determine what may limit potential solutions in a “best value” evaluation, specifically in the areas of:

   i. Inciting creative uses and innovative approaches to spur economic growth;
   ii. Ensuring equal access on equal terms;
   iii. Supporting NOAA's compliance with the US Open Data Policy (Executive Order 13642 and OMB Memorandum M-13-13);
   iv. Providing a self-sustaining business model; and
   v. Data-handling and data curation policies for extracted data.

b. The Government anticipates open and active lines of communication amongst all partners to ensure fairness, transparency, cooperation, and implementation of industry standards. What recommendations do interested parties have for managing this communication?

I would like to thank everyone for their interest in supporting NOAA. The Big Data Initiative is an exciting project that will require industry to be creative and innovative. As you will read in the draft SOO, we hope that this initiative will create jobs and spur economic growth while allowing NOAA to be as transparent as possible. We look forward to seeing you on October 17, 2014.

The complete list of attachments is listed below:
Attachment A: Big Data Industry Day Agenda
Attachment B: Draft Big Data Statement of Objectives
Attachment C: Comment and Suggestion Spreadsheet

**Big Data Industry Day Agenda.pdf (221.92 Kb)**

PDF

Description: Attachment A - Big Data Industry Day Agenda

Date: October 17, 2014
Time: 7:45 AM EDT – 12:30 PM EDT
Location: NOAA Auditorium and Science Center
1301 East West Highway
Silver Spring, MD 20910

<table>
<thead>
<tr>
<th>Time (EDT)</th>
<th>Description</th>
</tr>
</thead>
</table>

http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data
Updated: Wed, 23 Sep 2015 08:08:44 GMT
Powered by mindtouch®
Draft_Big_Data_Statement_of_Objectives.docx

Word

Description: Attachment B - Draft Big Data Statement of Objectives

NOAA’s Big Data Partnership (BDP)

Statement of Objectives (SOO)

1. Background

In support of the Department of Commerce’s “Open for Business” agenda, the National Oceanic and Atmospheric Administration (NOAA) Office of the Chief Information Officer (OCIO) is seeking to create a vibrant private sector participating in the development and delivery of weather; water; ocean; and climate data, products and services to the public. Through expanded partnerships with private sector and other enterprise partners, the potential exists to develop and market enhanced products and services to the Nation. NOAA is looking for partners to incite creative uses and innovative approaches that will tap the full potential of its data, spur economic growth, help more entrepreneurs launch businesses and create new jobs.

NOAA’s goal is to establish a public-private partnership to create a self-sustaining technical and procedural infrastructure to: lower access barriers to NOAA’s open data by putting a copy of that data in the cloud - providing equal access to all on equal terms; and intelligently position the data near computing, analytic, and other value-added services - creating a new market space for economic growth and job creation.
Solutions must support NOAA’s compliance with the US Open Data Policy (Executive Order 13642 and OMB Memorandum M-13-13 dated May 9, 2014, which calls for making data available via APIs in machine-readable formats.

NOAA is seeking one or more partners to address the challenges in providing access to NOAA’s data. Currently, only some of NOAA’s data is posted on public servers and web sites, especially in the case of large numerical model outputs and large sets of observational data. In many cases these datasets are hosted on multiple public servers which may not use the same standard services and formats. Therefore, it is difficult to find and integrate data from these sources for cross-domain analysis and decision-making. Furthermore, NOAA’s large data volumes require users to have substantial network, storage, and computing capabilities of their own in order to interact with and fully exploit the value of these data.

2. Requirements

To establish a flexible public-private partnership that will improve access to NOAA’s data sets. The Offeror’s Performance Work Statement (PWS) shall meet the following requirements:

- 2.1. The Offeror shall position NOAA’s data with analytic, compute, and storage capabilities at no net cost to the Government.
  - 2.1.1. The Offeror’s solution shall not have an adverse impact on the Government’s mission.
  - 2.1.2. The specific subset of NOAA’s data to be included shall be determined in consultation between NOAA and the Offeror, and shall focus on datasets that are not presently available, or that would benefit from more scalable access methods, or that have high potential for enabling derived products and services.
  - 2.1.3. The Offeror shall provide public access to NOAA data.
  - 2.1.4. NOAA data moved to the cloud shall remain free to the public in its original, unaltered form.
  - 2.1.5. The Offeror shall enable the private sector to establish and perhaps charge for new value-added services and products.
  - 2.1.6. In order to enable the private sector to build business and create new jobs, the Offeror shall not constrain the number of value-added service providers that can produce and provide products based upon NOAA data.
  - 2.1.7. The government desires that the Offeror’s solution be extensible to other data sources (e.g. future NOAA data; data from other federal agencies; and data from non-Governmental sources).
  - 2.1.8. Data Extraction
    - 2.1.8.1. All data extraction methodologies require approval by the Government.
    - 2.1.8.2. The Offeror may request data sets from the NOAA data catalog.
2.1.8.3. The Offeror may “pull” data from public NOAA servers or repositories that already house the data (e.g. external web servers, ftp servers).

2.1.8.4. If the data is unavailable from public repositories, or if performance or security considerations warrant it, NOAA will transfer (“push”) the data to the Offeror from internal, secure locations.

2.1.9. Offerors shall allow for prototyping technology for data extraction, data distribution, storage, computation, and data analytics:

2.1.9.1. Solutions shall be agile, capable, and have the ability to be changed or terminated.

2.1.9.2. The Government desires that Partners will work with NOAA to provide improvements to their offerings and capabilities as additional NOAA datasets become available.

2.1.10. If an enhanced level of data extraction is required, the Offeror shall be responsible for all additional costs. This includes, but is not limited to:

2.1.10.1. Movement of data to the cloud

2.1.10.2. The Offeror’s own infrastructure

2.1.10.3. The Offeror’s data management, data services, and data security

2.1.10.4. Appropriate documentation and attribution of NOAA data

2.1.11. The Government shall not reimburse the Offeror for any costs associated with data extraction (“push” or “pull”) from NOAA.

2.1.12. The Offeror shall provide a primary point of contact for the Government.

2.1.13. The Offeror shall describe how the datasets are to be advertised.

2.1.14. The Offeror shall provide a means to communicate issues about NOAA datasets to the general community.

2.1.15. The Offeror’s SLA and constraints of this arrangement shall be made public.

2.2. Equal access on equal terms

2.2.1. Offerors shall provide open, timely access to NOAA data.

2.2.2. NOAA will make any data that has been provided to the Offeror available to all BDP Partners.

2.2.3. Offerors shall not utilize (e.g. perform analysis or computation on the data or make value added services) data until it is made available to the public and value-added service providers.

2.2.4. Offerors shall not provide any privileged private access to public data assets.
2.2.4.1. Partners may provide additional levels of access to NOAA's public data assets, but if they choose to do so the Partner must provide equal access on equal terms.

2.2.4.2. Offerors may use NOAA data for any lawful purpose which includes providing value-added services.

2.3. Compliance

2.3.1. Offeror's solutions shall support NOAA's compliance with:

2.3.1.1. Executive Order 13642; and

2.3.1.2. OMB M-13-13

2.4. NOAA shall ensure that existing mission services are not impacted.

2.5. The Government shall provide datasets on a best-effort basis.

2.5.1. The Government is providing data sets on an “as-is” basis. Public data sets are provided under their existing terms and disclaimers.

2.5.2. Inquiries to NOAA regarding datasets shall be handled on a best-effort basis.

2.5.3. The Government will not provide a Service Level Agreement (SLA).

2.5.3.1. This includes but is not limited to: access to metadata, defined latency, guaranteed throughput, reliability, assurances of continued access to a provided data, or a guaranteed response to an inquiry.

2.5.4. The Government shall retain the final determination of what data to provide.

2.5.5. The Government shall provide a point(s) of contact that will be responsible for providing feedback on risk, reasonableness of extraction methodology, concerns, and approval.

2.5.6. The Government reserves the right to cease generating data and providing specific data sets without notification.

2.5.7. The Government shall not incur any liability as a result of the proposed solution or partnership agreement.

2.6. Data Curation

2.6.1. NOAA shall remain responsible for retaining and maintaining the scientific stewardship of any data provided to industry partners.

2.6.2. Data Lifecycle Management
2.6.2.1. Proposed solutions shall include lifecycle management for NOAA data stored within the provider's system. (e.g. Acceptable latency in accessing infrequently accessed data.)

2.6.2.2. NOAA will continue to create data that support its missions and products and will be responsible for its data lifecycle management.

2.6.2.3. The Government copy of the data shall be considered the “master copy”.

   2.6.2.3.1. The Government shall provide a checksum or hash mechanism to enable verifiability of delivered data packages against the reference copy.

2.6.2.4. Refreshes and updates to the Government data in the cloud shall replace outdated copies.

2.6.2.5. The Offeror shall provide control mechanisms including but not limited to redacting or purging of data that is deemed faulty.

2.6.2.6. The Offeror shall provide Government all data assets under the same terms of use as the Government

   2.6.2.6.1. Reference http://www.weather.gov/disclaimer

   2.6.2.6.2. The Government may provide terms of use for specific data assets

2.6.3. Data Provenance

   2.6.3.1. Offerors shall make Government-provided data publicly available in its original form.

   2.6.3.2. Offerors may also distribute appropriately-attributed additional and derivative products.

   2.6.3.3. Offers shall provide appropriate attribution of Government-provided data.

   2.6.3.4. NOAA datasets with a NOAA-assigned Digital Object Identifier (DOI) shall be attributed or cited using that DOI.

      2.6.3.4.1. NOAA desires that derived products include metadata documenting provenance and lineage.

2.6.4. The Offeror shall utilize established Government data catalogs

   2.6.4.1. The Offeror shall utilize the NOAA data catalog (data.noaa.gov) to identify data sets.

   2.6.4.2. The Government may request input on prioritizing desired datasets.

   2.6.4.3. The Offeror may provide feedback on any requested enhancements to the NOAA data catalog.

   2.6.4.4. The Offeror shall provide support for interfacing to the existing (or future) NOAA Data Catalog in the following methods:
2.6.4.4.1 The proposed solution shall be able to provide reading of the Data Catalog through at least one of these existing methodologies:

2.6.4.4.1.1 Open Geospatial Consortium (OGC) Catalog Service for Web (CSW) interface

2.6.4.4.1.2 CKAN native API

2.6.4.4.1.3 Native CKAN user interface

2.6.4.4.2 The proposed solution shall provide methods to update the Data Catalog for the following outcomes:

2.6.4.4.2.1 The proposed solution shall provide access information for NOAA datasets available through the respondents cloud platform

2.6.4.5 The Offeror may work with NOAA to suggest improvements in dataset metadata.

2.7. The Offeror shall provide a reporting plan to include:

2.7.1. The type of information to be reported on, frequency of reporting (schedule), and methods for delivery of report information to NOAA.

2.7.2. Statistical analysis on the usage of NOAA data and any derived metrics to assess job creation or value to the economy.

2.7.3. The Government is interested in high level reporting on the Quality Assurance Surveillance Plan (QASP) (see section 2.8).

2.7.4. To avoid overly-cumbersome reporting, the Government will work with the partner.

2.7.4.1. The Government desires that the reporting be as lightweight and automated as possible.

2.7.5. The Offeror shall provide a Quarterly Status Report (QSR).

2.7.6. The Offeror shall provide an Annual Executive Review (AER) at a location to be determined by NOAA.

2.8. The Offeror shall provide a QASP

2.8.1. The Offeror shall provide appropriate systematic testing

2.8.1.1. The Offeror shall provide documentation of testing methods and results

2.8.2. The Offeror shall provide statistical sampling to appropriately survey the data

2.8.2.1. The Offeror shall provide documentation of statistical methodology and the survey results.
2.8.3. The Offeror shall provide verification that NOAA data is made available in accordance with requirement 2.2 Equal Access on Equal Terms

2.8.3.1. The Offeror shall provide documentation of the methodology used to provide this verification.

2.9. Project Plan

2.9.1. The Offeror shall maintain a project plan for:

2.9.1.1. Initial implementation

2.9.1.2. Data extraction for data the Government agrees to push to the Offeror’s infrastructure.

2.9.1.3. Deployment of ongoing improvements.

2.10. Information and Cyber Security Requirements

- The Offeror shall ensure compliance with GSA FedRAMP for hosting of data in cloud environment.

3. Period of Performance (POP)

3.1. 12 month base from time of award (CY1)

3.2. 12 month option (CY2)

3.3. 12 month option (CY3)

3.4. 12 month option (CY4)

3.5. 12 month option (CY5)

3.6. 6 month transition option

Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AER</td>
<td>Annual Executive Review</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>BDP</td>
<td>Big Data Partnership</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>CKAN</td>
<td>Comprehensive Knowledge Archive Network</td>
</tr>
<tr>
<td>CSW</td>
<td>Catalog Service for Web</td>
</tr>
<tr>
<td>CY</td>
<td>Contract Year</td>
</tr>
<tr>
<td>DOI</td>
<td>Digital Object Identifier</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NWS</td>
<td>National Weather Service</td>
</tr>
<tr>
<td>OGC</td>
<td>Open Geospatial Consortium</td>
</tr>
<tr>
<td>OMB</td>
<td>Office of Management and Budget</td>
</tr>
<tr>
<td>OSTP</td>
<td>The White House’s Office of Science and Technology Policy</td>
</tr>
<tr>
<td>POP</td>
<td>Period of Performance</td>
</tr>
<tr>
<td>PWS</td>
<td>Performance Work Statement</td>
</tr>
<tr>
<td>QASP</td>
<td>Quality Assurance Surveillance Plan</td>
</tr>
<tr>
<td>QSR</td>
<td>Quarterly Status Review</td>
</tr>
<tr>
<td>SLA</td>
<td>Service Level Agreement</td>
</tr>
<tr>
<td>SOO</td>
<td>Statement of Objectives</td>
</tr>
</tbody>
</table>
Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>master copy</td>
<td>Also known as the “gold copy” or “record copy”. The authoritative version.</td>
</tr>
<tr>
<td>Offeror</td>
<td>An organization (e.g. commercial, non-profit, or University) that formally responds to NOAA’s Big Data Partnership solicitation.</td>
</tr>
<tr>
<td>Partner</td>
<td>An organization (e.g. commercial, non-profit, or University) that is awarded a contract.</td>
</tr>
<tr>
<td>Privileged Private access</td>
<td>Exclusive access to the Government’s data that is unavailable to everyone. An enhanced tier of access is non-exclusive if anyone can obtain, on equal terms, that level of access.</td>
</tr>
<tr>
<td>Value-added service provider</td>
<td>An entity that produces a product, service, or insight based upon the Government’s data.</td>
</tr>
</tbody>
</table>

Comment_and_Suggestion_Spreadsheet.xlsx (11.33 Kb)

Excel

Description: Attachment C - Comment and Suggestion Spreadsheet

NOAA announces RFI to unleash power of 'big data'

Source: http://www.noaanews.noaa.gov/stories...4_bigdata.html

Agency calls upon American companies to help solve ‘big data’ problem

February 24, 2014

Today, NOAA, part of the Department of Commerce, announced a new effort to unleash the power of its data to foster innovation, create new industries and job opportunities, and spur economic growth. NOAA, through a Request for
Information, is looking to the private industry to help make NOAA's data available in a rapid, scalable manner to the public.

Of the 20 terabytes of data NOAA gathers each day -- twice the data of the entire printed collection of the United States Library of Congress -- only a small percentage is easily accessible to the public. Through the RFI, American companies will be able to provide possible solutions for NOAA to be able to turn this untapped information into usable products or services.

"From the surface of the sun to the depths of the ocean floor, NOAA works to keep citizens informed about the changing environment around them," said Kathryn Sullivan, Ph.D., acting NOAA administrator. "Our vast network of radars, satellites, buoys, ships, aircraft, tide gauges, and supercomputers keep tabs on the condition of our planet's health and provides critical information that's used to predict changes in climate, weather, oceans, and coastlines. As we continue to witness changes on this dynamic planet we call home, the demand for NOAA's data is only increasing."

"Quite simply, NOAA is the quintessential big data agency," said Joe Klimavicz, NOAA's CIO. "However, due to limited resources, only a fraction of that data is effectively and efficiently made available to the public. Imagine the economic potential if more of these data could be released. Unleashing the power of NOAA's data will take creative and unconventional thinking, and it's a challenge we can't tackle alone."

NOAA operates a number of operational platforms including: Doppler radar systems; weather satellites; buoy networks and stations; tide gauges; real-time weather stations; and ships and aircraft. This network provides valuable and critical data that are instrumental in protecting lives and property across the country. As demand increases for this data and information, NOAA must find ways to effectively and efficiently distribute it to decision makers and industries.

More information about the RFI, including how to participate can be found here.

NOAA's mission is to understand and predict changes in the Earth's environment, from the depths of the ocean to the surface of the sun, and to conserve and manage our coastal and marine resources. Join us on Facebook, Twitter, Instagram and our other social media channels.

---

Big Data RFI from NOAA

Source: here.
Solicitation Number: ST-1330-14-RP-0039
Notice Type: Sources Sought

Synopsis

Added: Feb 21, 2014 1:45 pm
BIG DATA RFI FROM NOAA
This is a request for information (RFI) only, and is not a solicitation for a contract or grant award. This RFI notice is for information purposes only, is not a request for proposals, and does not obligate the government in any way. The Government will not reimburse the respondents for any costs associated with the information submitted in response to this request. The Government will treat each submission as confidential.
Background
The Department of Commerce, National Oceanic and Atmospheric Administration (NOAA) Office of the Chief Information Officer (OCIO) is requesting information from industry and other organizations (e.g., non-profits, research laboratories, and universities) to determine whether capability and interest exists for establishing partnerships with NOAA for the purpose of intelligently positioning NOAA's vast data holdings in the cloud, to be co-located with easy and affordable access to computing, storage, and advanced analytical capabilities. NOAA anticipates these partnerships will have the ability to rapidly scale and surge; thus, removing government infrastructure as a bottleneck to the pace of American innovation and enabling new value-added services and unimaginable integration into our daily lives.

Only a small percentage of NOAA's data is posted on public servers and web sites. In many cases these data are hosted to multiple public servers, which may not use the same standard services and formats. Therefore, it is difficult to find and integrate data from these sources for cross-domain analysis and decision-making. Furthermore, NOAA's large data volumes require users to have substantial network, storage, and computing capabilities of their own in order to interact with and exploit the value of this data. NOAA is exploring opportunities to unleash the potential of its environmental data through alternative approaches to strategically position this data with analytic and computational capabilities.

A vibrant private sector participating in the development and delivery of weather; water; and climate data, products, and services to the public is of value to the U.S. economy and environmental awareness. Through expanded partnerships with private sector and other enterprise partners, the potential exists to develop and market enhanced products and services to the Nation. NOAA is looking for partners to incite creative uses and innovative approaches that will tap the full potential of its data, spur economic growth, help more entrepreneurs launch businesses, and to create new jobs.

Purpose of this Request for Information (RFI)
This RFI is intended to inform NOAA on the current status of industry sources, business practices, technical capacity, and operational capability. Furthermore, this RFI is intended to inform NOAA on the feasibility of partnering with one or more industry partners using no-cost agreements. These partners would be responsible for funding all costs related to the movement of data to the cloud as well as their own technology infrastructure, data management, data services, data security, and free data access to users. While this mode of delivery protects NOAA from the cost of provisioning additional infrastructure, it will require careful thought and controls to ensure NOAA's existing services are not impacted and NOAA remains compliant with statutory regulations. This mode of delivery should also provide the appropriate documentation and attribution of NOAA data. NOAA does not intend to use these partnerships as a replacement for existing services. NOAA will remain responsible for retaining and maintaining the scientific stewardship of any data provided to industry partners. The expectation is that all NOAA data moved to the cloud would remain free to the public in its original, unaltered form, but this model will enable the private sector to establish and perhaps charge for new value-added services and products. A proposed solution must also enable NOAA compliance with Presidential Executive Order: 'Making Open and Machine Readable Data the New Default for Government Information'. "As one vital benefit of open government, making information resources easy to find, accessible, and usable can fuel entrepreneurship, innovation, and scientific discovery that improves Americans’ lives and contributes significantly to job creation."
NOAA is providing all industry stakeholders an opportunity to comment and respond. NOAA encourages communication and a robust dialogue on this RFI.

**Information Requested**

Interested firms should submit a response with a written statement of interest or capability and discussion of the following enumerated points:

1. **Value Proposition**

   Including consideration of NOAA’s responsibilities by virtue of Executive Order 13642 and OMB M-13-13:


   a. Describe your vision of potential innovations, products, or opportunities that may become viable if NOAA releases its environmental data in a strategic location that has both analytic and computational capabilities.
   b. Describe how sharing NOAA’s data could create jobs or spur the economy.
   c. Provide metrics that should be considered to measure the level of impact or success from releasing this data through this partnership approach.

2. **Data Infrastructure**

   a. Provide architectural information and diagrams representing the intended role for any described capability (e.g., data storage, data ingest, data processing, analytics)
   b. Describe a viable methodology for extracting environmental data from NOAA's access locations, distributed across its enterprise in multiple geographic locations.
   c. Once the data are collected, describe how these data might be hosted in your environment and how it would be accessible to the public and users of your services.
   d. Describe your methods for ingesting data from source systems and open sources, automatic application of metadata and other tags, and delivery to analysts and users in "near real time".
   e. Describe the methods your organization could use to support user capacity and experience (look and feel) such as uptime guarantees, simultaneous user capability (e.g., crash-free user load capacity limits), download bandwidth, and computational capacity (including diverse computational architectures).
   f. Describe any data transfer, storage, or hosting limitations that should be considered.
   g. Identify the underlying open source technology and/or standard(s) or other standards (such as ISO, IEEE, industry) to which each capability or product adheres (including data and metadata standards).

3. **Data Services**

   a. Identify any catalog or description that NOAA would need to provide along with the data.
   b. Identify any NOAA data or datasets which you believe would be of highest value to initiate this partnership.
   c. Describe 1) relevant geospatial and other analytic technical capabilities that help data scientists and untrained end users alike identify patterns, trends, and data segments of interest in the context of dynamic and continuous correlation;
2) capabilities that inherently manage information relationships and models; and 3) agile analytics capabilities that can help data scientists and/or end users identify and dynamically exploit emerging trends in data sets and data flows.

d. Describe 1) relevant visual capabilities for helping users understand important relationships in large data sets (e.g., geospatial, graphs, time series, etc.); 2) information access in mobile environments; and 3) capabilities that enable information sharing through visual depictions of large/complex information.

e. Describe potential data services or any value-added data services and products.

f. Describe the methodology that you might use to determine the price paid by users for value-added services and products.

4. Data Management

a. Describe your methodology to ensure data integrity with the "reference copy" that will be retained within NOAA.

b. Describe data provenance methods that could be used to determine attribution and guarantee the reliability of the data.

c. Describe how your organization would transition the data to another provider, if necessary.

5. Data Security

a. Identify methods that are available in your solution for hardening the capability and/or security features that are intrinsic to your capability to meet FISMA security standards.

6. Partnership Methods

a. Describe any acquisition or partnering methodology that you recommend.

b. Recommend whether NOAA should engage a systems integrator or multiple cloud vendors to accomplish this objective.

c. Describe how your organization could be able to provide the services at no cost to the Government.

d. Describe the target customer audience and how you envision a sufficient revenue stream to be able to reliably provide the services.

e. Describe any risks that NOAA should consider or mitigate.

7. Capabilities

a. Provide a brief description of your Big Data capabilities, including storage, data analytics, and data visualization.

b. NOAA has 10s of petabytes of data and produces over 15 million products daily (in the order of 10s of terabytes). Does your organization have the ability to host major portions or all of NOAA's releasable data and the associated metadata? Describe your capabilities.

c. Provide an exemplar case study or success story where your Big Data solutions were integrated. Explain the alternatives considered and challenges faced.

d. Provide information on what Federal agencies and/or private entities have benefited from your Big Data capabilities and how your company is prepared to deliver and service Big Data capabilities to a Federal agency that must comply with IT laws, regulations and policies.

e. Responses from large and small businesses are encouraged. Please provide the size status of your company/organization based on NAICS 518210 - Data Processing, Hosting, and Related Services, with a size standard of $30 million.

http://semanticommunity.info/Data_Science/Data_Science_for_NOAA_Big_Data

Updated: Wed, 23 Sep 2015 08:08:44 GMT
Powered by MindTouch™
Instructions

Written comments, no longer than fifty (50) pages in length, should be submitted, electronically in PDF, no later than March 24, 2014, addressed to: David.Zhang@noaa.gov and Denise.e.Harper@noaa.gov. Please do not submit marketing material or business proprietary information.

Contracting Office Address

SSMC2 - 11th floor /OFA61
1325 East West Highway, 11th Floor
Silver Spring, Maryland 20910
United States

Primary Point of Contact

David Zhang,
Contract Specialist
david.zhang@noaa.gov
Phone: 301-628-1399

Secondary Point of Contact

Denise E. Harper,
Contracting Officer
denise.e.harper@noaa.gov
Phone: 301-628-1383