Summary of Questions & Answers following the 2/9/21 NOAA Environmental Leadership Seminar, "NOAA/NESDIS Transformation with Innovative Technologies" by Ms. Irene Parker, Assistant Chief Information Officer, NESDIS, NOAA

For more Information, contact Chad.Hamrlik@noaa.gov

1.) David: Would it be correct to assume that the PDA function would move to the cloud and that processing in the cloud could be increased so that when a continuous GSO stream is coming in it would not be slowed down by polar satellite stored mission data coming in.
ANSWER: Our migration roadmap includes moving PDA functionality to the Cloud in Dissemination and Access service of NCCF in the future.

2.) Ted: Being analysis ready is a real problem, data is messy, how are you going to deliver this?
ANSWER: Start by working early on with data providers to produce data in cloud optimized formats with enhanced populated metadata, and participating in ARD discussions with external partners and organizations that explore the best structure in cloud framework to analyze data in an efficient and effective way

3.) Michele: Is NCCF a medium or High system? How did NESDIS consolidate other Cloud efforts / tools that might have been based on different operating systems?
ANSWER: The NCCF is operating in two environments: The operational component is running in FISMA HIGH while a dev/science sandbox services are running in both high and moderate FISMA environments. For consolidation efforts, we looked for a solution that met a majority of our requirements. We standardized on Linux OS as the baseline and built a common toolset to support a majority of workflows. Please note that both environments are still utilizing the FedRAMP Moderate AWS cloud environment regardless of FISMA.

4.) Ted: Does the intellectual property rights apply to use by your taxpayers?
ANSWER: For each OTA, the Government must negotiate the intellectual property rights (IPR) of the OTA participants. Under this OTA agreement, by default, AI code will be developed jointly between NOAA and Google, and publicly released for anyone to use without restriction.

5.) Hernan: Unless data users know where to look for data at NOAA, external and also internal NOAA data users cannot readily discover and find data. We are one NOAA with different data streams. For example, there is no data search engine at the NESDIS or NOAA web sites. How could we improve data access at NOAA?
ANSWER: Absolutely, one of the near term goals of leadership is to improve the discoverability of the data from a centralized catalog and interface. One way to improve the discoverability and accessibility is to come together utilizing a common schema which populates metadata to a single catalog. To start we are working on merging the NOAA Data Catalog and OneStop so data being published to those two catalogs will have a single access interface.

6.) Tracy: Are you aware of how other NOAA line offices are using the cloud? Are Line offices working together at all, or is each LO working on their own cloud system?
ANSWER: All NOAA LOs are in the cloud in some fashion. There is a NOAA level committee, NOAA Cloud Committee where LOs are coordinating, collaborating and sharing lessons learned.

7.) Ted: Is NOAA having intellectual property rights, exclusive of your taxpayers?

ANSWER: For each OTA, the Government must negotiate the intellectual property rights (IPR) of the OTA participants. Under this OTA agreement, by default, AI code will be developed jointly between NOAA and Google, and publicly released for anyone to use without restriction.

8.) David: My AdminLan windows laptop has been "restarting" for the past 4 hours. 100% of my work is in the cloud. Why don't we start doling out chromebooks which are a lot less money and a lot more reliable.
ANSWER: NESDIS is looking at available solutions and working with NOAA OCIO to determine usability of Chromebooks going forward.

9.) Ted: How much control does GOOGLE have on the project list and execution?
ANSWER: Google and NESDIS worked together to determine the projects to be executed under our partnership. Google provides the AI experts and we provide the science experts. Note that the Government can not dictate the execution of a project(s), they must be agreed to jointly with Google.

10.) Jack: I would love to see a link to the Framework list of projects/ideas, as well as who are the reps from the NOAA line offices.
ANSWER: If you questions about OTA projects that were evaluated and what line office representative participated in the OTA planning - please contact Sid Boukabara (sid.boukabara@noaa.gov)

11.) Ted: My good science is being controlled by private company scientists, this is why line offices would be resistant.
ANSWER: You should contact your line office leadership to express your concerns. There are NOAA organizational goals and objectives related to science, artificial intelligence, OTAs, and migrating to the cloud which are driven by congressional legislation and executive branch directives.

12.) David: It has been mentioned in a different meeting that a copy of NOAA's data must have a "golden copy" be stored on-prem. why?
ANSWER: We are actively working with NOAA Chief Data Officer (CDO) and NARA to establish guidelines on "golden copy" storage in the cloud. Recently NOAA CDO has reached out to NARA's ERA2.0 chief architect to understand NARA's policies on having the "Golden Copy" stored in the cloud as well.

13.) Scott: What are the status and plans for CLASS moving to the cloud?
ANSWER: Currently we are moving our backup copy of CLASS data to the AWS cloud. We plan to complete the migration and have COOP/Disaster Recovery (DR) CLASS done via AWS cloud. We will finish the migration to AWS in FY22. The CLASS Asheville node will be the primary point of fulfillment for customer orders, at this time.

14.) Jason: Does "cloud optimized formats with enhanced metadata" imply some sort of NOAA-wide data standard?
ANSWER: Within NOAA we certainly will, at some point, identify recommendations from involvement with several agency partnerships and organizations like CEOS, ESIP, OGC, WMO, etc. and develop internal guidance.

15.) Sheryl: It's great to hear that you're negotiating an egress cost solution. Along those lines, programs would like to be able to track egress via cost tags in AWS. Currently cost tags aren't viewable by the programs. Would OCIO and NESDIS consider making it possible for programs to view cost allocation tags in order to track, manage, and assign egress costs?
ANSWER: The NOAA Cloud Utility Contract under SAIC uses CloudCheckr tool to provide cost information by tags/resources. The CORs and task monitors are provided access to CloudCheckr reports

and dashboards which provide this information already. Please open a service ticket with SAIC to get access to the tags and costing information in the cloud. Additionally, SAIC has enabled access to AWS Cost Explorer tool in the AWS console that provides a more near real-time access to costing information. NOAA has approved 5 mandatory tags so far but the line offices are free to include additional tags for a more granular costing. NESDIS is working with Cloud PMO to include additional tags

16.) We are having issues copy/pasting the link from the presentation.
ANSWER: Please use the below links which are referenced in the presentation.
Google OTA: https://www.nesdis.noaa.gov/nesdis-google-artificial-intelligence-prototyping-initiative
NOAA Cloud Strategy:
https://nrc.noaa.gov/Portals/0/Final%20Cloud%20Strategy.pdf?ver=2020-07-02-122459-813
NOAA Artificial Intelligence Strategy:
https://nrc.noaa.gov/LinkClick.aspx?fileticket=0I2p2-Gu3rA%3D&tabid=91&portalid=0

17.) Khadijah: Could you explain the process of data tagging?
ANSWER: Data tagging is part of the process of labeling data with additional information about that data. With Earth Data Records (for instance) this is accomplished through metadata elements and other attributes within that element. The goal is to ensure data is tagged appropriately to assist with efficient data discovery and dissemination.

18.) AJ: How is NESDIS employing AI/ML to automate routine but human intensive operations to optimize operations?
ANSWER: We aren't at this time. We are beginning to look at how we can use AI/ML for weather forecast improvements. Lessons learned from there may be applied to other aspects of the NESDIS mission.

19.) Is there something setup in Google to allow innovation by users to try things ahead of time before they move into a product?
ANSWER: It depends on the Line Office. Some offices have an infrastructure in Google and allow their users access. NESDIS does not have an environment in Google at this time.

20.) Melissa: Is NOAA able to leverage the expertise of the Google engineers to provide guidance and best practice about how NOAA should structure its data and services?
ANSWER: Yes, that is the secondary intent with the OTA. We will learn about AI/ML from their researchers and they will learn how to interpret and use NOAA data from us.

21.) Ted: How are you structuring training to get your scientists fully integrated with Goggles scientists, so that Google is just not a service bureau?
ANSWER: On the infrastructure side, we have Google engineers working on deploying the infrastructure in GCP being shadowed by NOAA staff. We are getting documentation on infrastructure deployment.<text>

22.) Richard: Will the AI/ML work be non cloud centric if multiple cloud vendors are used or specific to Google?
ANSWER: The Google OTA that has been negotiated will be done in Google Cloud Platform (GCP). The AI/ML work will leverage all of Google's tools. The concepts demonstrated and the techniques used in the OTA will be key for our NESDIS scientists to bring back and apply.

23.) Ted: If you move away from Google, what is that going to mean for AI/ML if your training is at Google now?
ANSWER: The AI/ML work will leverage all of Google's tools. The concepts demonstrated and the techniques used will be key for our NESDIS scientists to take back and apply. The problem becomes the data, without the data, training isn't possible. If we make a conscious decision to move to a different platform, the data must be in place on that new platform.

24.) Deke: You mentioned discoverability and potential for new usages of NOAA data. How can these technologies help NOAA better service NOAA's existing customers, who are more invested in continuity of services than perhaps novel users?
ANSWER: NOAA's goal is to emphasize the value of NOAA's data and realize its value as an asset. To accomplish this we need to improve on the discoverability and interoperability of the data for both external and internal users. When exploring new technologies we must consider continuity of service as a priority. Stakeholder(s) feedback and experience is a critical and valued component when implementing new technologies.

25.) Jeff: Is there an expectation that CLASS data will be accessible once the backup copy is migrated there? My understanding is that it uses Deep Glacier storage, with limitations on accessibility beyond defined backup / COOP needs.
ANSWER: CLASS data in AWS will be a dark copy. The Asheville node will fulfill user requests at this time.

26.) Richard: Is Google Anthos which work in Hybrid cloud be looked at or considered given that you are moving 15 terabytes to AWS for Class data and Google is developing the AI/ML?
ANSWER: We are moving ~15TB to AWS for CLASS needs. We will be copying other data sets to Google for the AI/ML work (less than 1PB currently). Google Anthos is currently not being considered.

27.) AJ: Is Google performing the Test and Evaluation of ML algorithms, in addition to developing them, or is their concern about having third party eval and tracking algo's for model shift and explainability?
ANSWER: Google researchers and NESDIS scientists will be working side by side in developing the algorithms.

28.) Dave: The cloud, AI, and ML work best in a flatten org. NESDIS and NOAA to some degree is a command and control hierarchical org. How can NESDIS promote innovation successfully?
ANSWER: Innovation is not easy, and traditionally NESDIS hasn't been willing to accept the associated risks. This risk averseness is changing and we are seeing the possibilities of accepting more risk with cloud projects. Innovation may come when we are less concerned with the infrastructure and who manages it, and shift the focus on what the data could produce. This takes time for teams to get past their desire for ownership of the systems and embracing change.

29.) Nazila: In the past, there has not been a cost associated with data retrieval and preparation.   Will there a way to reduce the costs of egress for internal to NOAA data users that use the data operationally?
ANSWER: Instead of moving the data from the cloud back to our on premises systems, NESDIS is providing scientists and analysts environments in the cloud where they can interact with the data. This approach significantly reduces egress of the data from the cloud to on premise systems.

30.) Ted: Is the US considering cloud service in the US only, offshore may not be so secure, politically this could be a problem depending on the administration, don't ask this online.
ANSWER: All NOAA cloud environments have to be compliant with NOAA and federal government rules and regulations.

31.) Melissa: Is NOAA able to leverage the expertise of the Google engineers to provide guidance and best practice about how NOAA should structure its data and services?

ANSWER: Absolutely. This is one of the stated goals of the Google OTA, it is a secondary intent with the arrangement. We learn about AI/ML from their researchers and they will learn how to interpret and use NOAA data from us.