

National Aeronautics and Space Administration



# NASA Big Data Working Group Meeting Summary

Office of the Chief Information Officer

## NASA Big Data Task Force Meeting Washington, DC

Tsengdar Lee, SMD HEC Program Manager

John D. Sprague, Associate CIO, T&I Division (Acting)

3/6/2017

[www.nasa.gov](http://www.nasa.gov)





## Agenda

- What T&I Does
- Big Data Working Group Meetings
- New Roles
- HEC and Big Data
- NASA Big Data Examples
  - » EVA
  - » QuAIL
- Deep Content Analytics and Deep Q&A



## ARCH

### Information Architecture

*The development and use of models, policies, rules or standards that govern how information is collected, stored, arranged, integrated, and put to use in data systems and in organizations.*

Standards and interoperability  
IA Reference Architecture  
Data Mining  
Data Contract Language  
Mission Support Tech Consulting

#### **Big Data**

NIAM.NASA.Gov

## Tech & Innovation Division--What We Do:



## DATA

### Data Science

*The collection, management and analysis of data in order to produce information and drive decision making.*

Data Strategy  
Data Governance  
Data Lifecycle Management  
Data Analytics Lab  
Data Fellows Program  
Data Stewards



## OPEN

### Open Innovation

*The development of new innovation frameworks and techniques, and the development and delivery of machine-readable instructions to access, arrange, and apply data.*

Agency Open Data Mgmt  
Digital Strategy Reporting  
Space Apps Challenge  
Innovation Incubator  
Data Innovation Pipeline  
Women in Data  
Open.NASA, Github/NASA  
Data.NASA, Code.NASA  
API.NASA



## DIGITAL

### Digital Integration

*The delivery of enhanced digital capability to support data interoperability and accessibility to enable data insights and discovery.*

Data inventory/Registry  
Tagging/discoverability  
Data usability/APIs  
Computing/Coding  
Mission-focused tech applications



## TECH

### Tech Infusion

*Research, prototype and assess the creation, modifications and usage of processes or tools to solve a problem, to meet stakeholder needs*

Data Centric Architecture  
Internet of Things  
Software as a Service  
Virtual Desktop Infra  
Collaboration (Secure Fileshare, NASATube)



## Big Data Working Group Meetings

- Co-Chaired by OCIO and SMD
- Meets monthly, 2nd Thursday of each month, typically 1 hour
- 80+ Members
- Notes are captured on the <http://niam.nasa.gov/big-data-2/> website (only accessible from the NASA corporate network or VPN)
- Big Data Face to Face Work Sessions (1 or 2 a year, rotating Centers)
- Suggested Training
  - MIT's ***Tackling the Challenges of Big Data*** online course is considered the gold standard
  - NASA Intro to Big Data, in the works: Big Data SATERN course
  - Various Open Source Big Data tools by Apache
  - Massive Open Online Courses (MOOCs)
    - Coursera, Udacity, Udemy, Big Data University and edX



## Big Data Working Group Meetings Typical Agenda

- Roll Call, New Members Introduced
- Info on Upcoming NASA, National and World Big Data Events
- Latest Hot Projects, Demos, New Regulations
- Resource Sharing
- Guest Speaker (rarely)
- Around the Horn (Mission Directorates, Centers)
- Future Topics





## NASA New Roles with Big Data at Office of the CIO

- New NITRD Subcommittee member: Deputy CIO Terry Jackson becomes NASA representative on the NITRD subcommittee
- Data Scientist: NASA has several including one in Office of the Chief Information Officer
- Big Data Administration
- Chief Data Officer/Chief Knowledge Officer: Becoming very popular in Federal Government and in Industry
- Big Data Stewards
- Data Evangelists
- Labs:
  - » Data Analytics Lab
  - » Internet of Things Lab

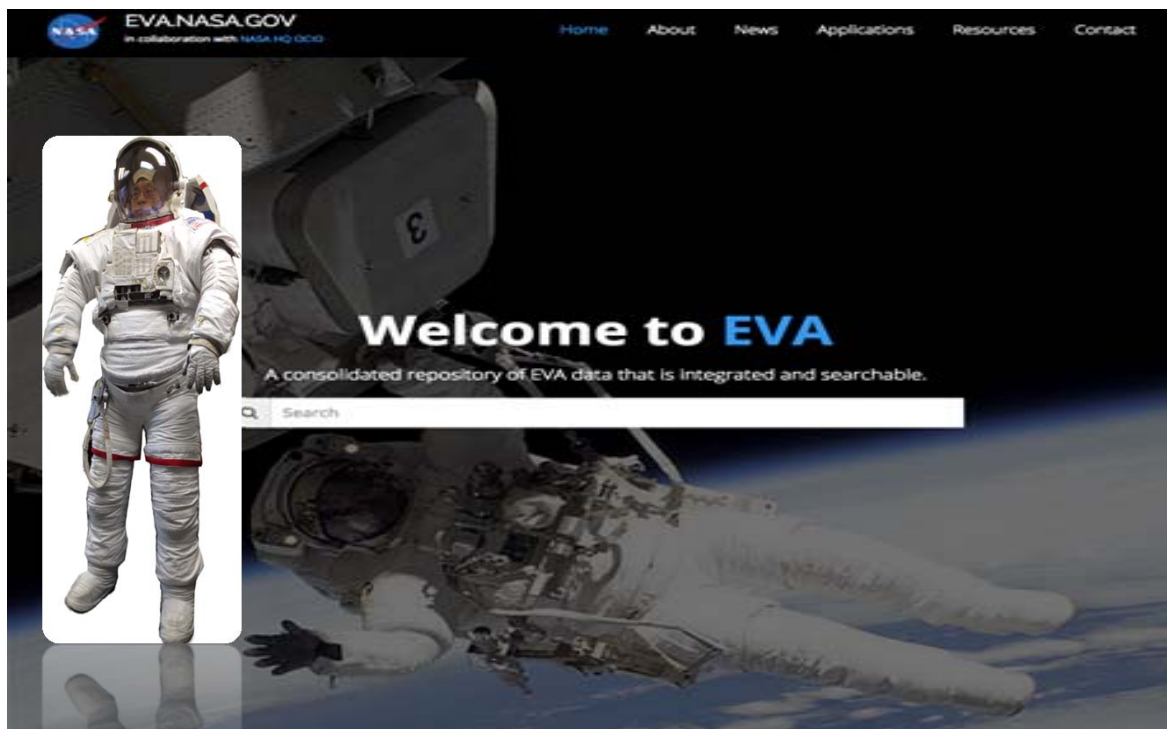


## High End Computing (HEC) and Big Data

- Newly amended HPC Act [As Amended Through P.L. 114–329, Enacted January 6, 2017] provides clear definition of high-end computing to include big data:
  - » ‘high-end computing’ means the most advanced and capable computing systems, including their hardware, storage, networking and software, encompassing both massive computational capability **and large-scale data analytics** to solve computational problems of national importance that are beyond the capability of small- to medium-scale systems, including computing formerly known as high-performance computing.



## NASA Examples – Extra Vehicular Activity (EVA) Data Integration



Access to all EVA data – Search, Browse, Navigate & Analyze

### Just Completed Phase 2

#### State-of-the-art Platform using Tools:

- Core platform based on ELK Stack
- PreAlert for Anomaly Detection
- Hosted in AWS Government Cloud
- Various databases including RethinkDB, Neo4J, Amazon RDS, & PostgreSQL
- APIs deployed as micro services using Docker container
- Amazon S3 for document storage
- Integrated with jBPM workflow engine
- UI framework: HTML5, Angular, Bootstrap, Materialize CSS
- 2D and 3D models for browsing





## Quantum Artificial Intelligence Laboratory (QuAIL)

- QuAIL is the space agency's hub for an experiment to assess the potential of quantum computers to perform calculations that are difficult or impossible using conventional supercomputers.
- “The D-Wave took about a hundredth of a second; with a classical computer it'd take about 100 days”, Google’s Director of Engineering, Hartmut Neven
- Beginning with the D-Wave Two™ quantum computer, NASA’s QuAIL team is evaluating various quantum computing approaches to help address NASA challenges.

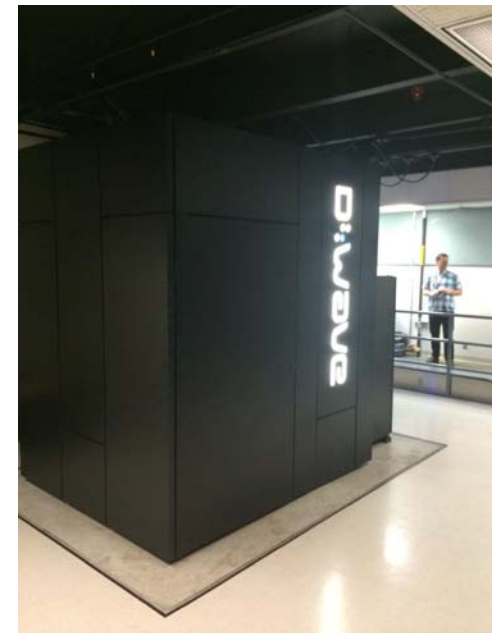


Photo by John Sprague taken 1 May 2014, with Ed McLarney from Langley Research Center.

<https://ti.arc.nasa.gov/tech/dash/physics/quail/>



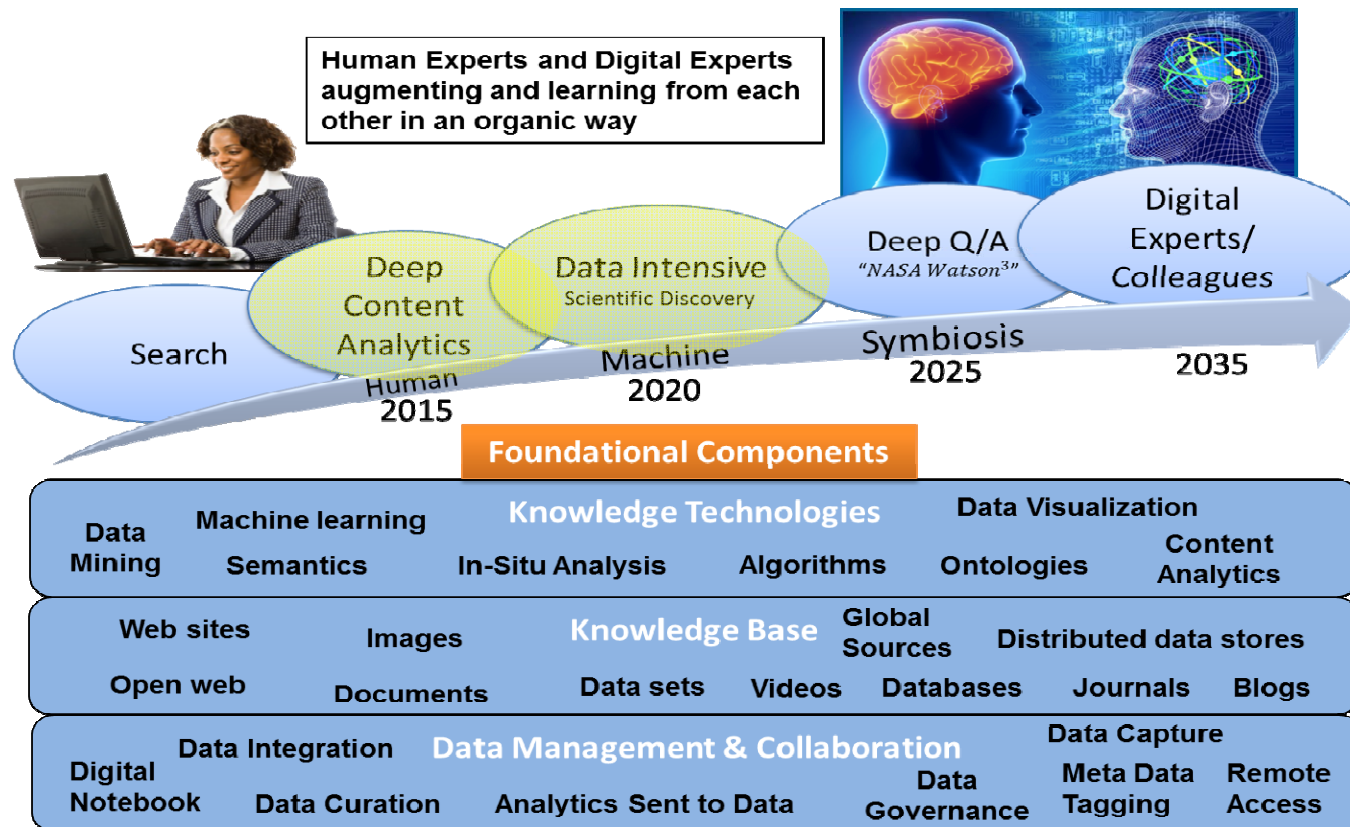
## Langley Research Center's Deep Content Analytics and Deep Q&A

- Deep Content Analytics: using Watson Content Analytics Technology
- Goal is to provide community with natural language processing technologies that will quickly make sense of internal/global knowledge by identifying trends and experts, aiding in discovery, and finding answers to questions with evidence
- Key Collections Analyzed:
  - Space Radiation
  - Aerospace Vehicle Design
  - Carbon Nanotubes
  - Autonomous Flight
  - Model Based Engineering
  - Human-Machine Teaming
  - NASA Lessons Learned
  - Uncertainty Quantification
  - Select NASA Technical Reports
  - Space Mission Analysis
- Deep Q&A: Using Watson Discovery Advisor Platform



## Big Data Analytics & Machine Intelligence Capability Vision: Virtual Research and Design Partner

**Enable NASA employees to achieve greater scientific discoveries and systems innovations**





# Questions?



# BACKUP SLIDES



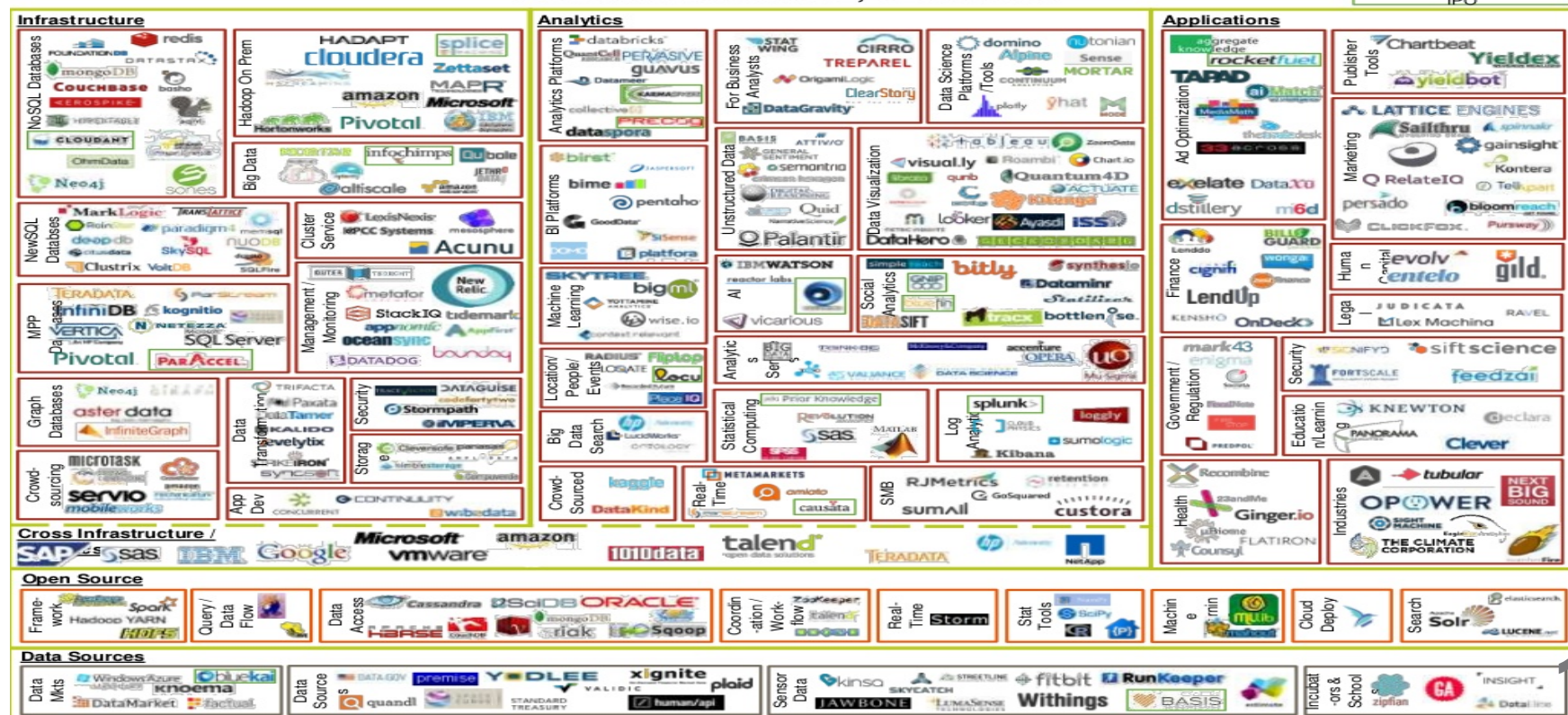


# Big Data Tools & Technologies

## Big Data Landscape

**BIG DATA LANDSCAPE, VERSION 3.0**

Exited: Acquisition or IPO



© Matt Turck (@mattturck), Sution Dong (@sutiandong) & FirstMark Capital (@firstmarkcap)