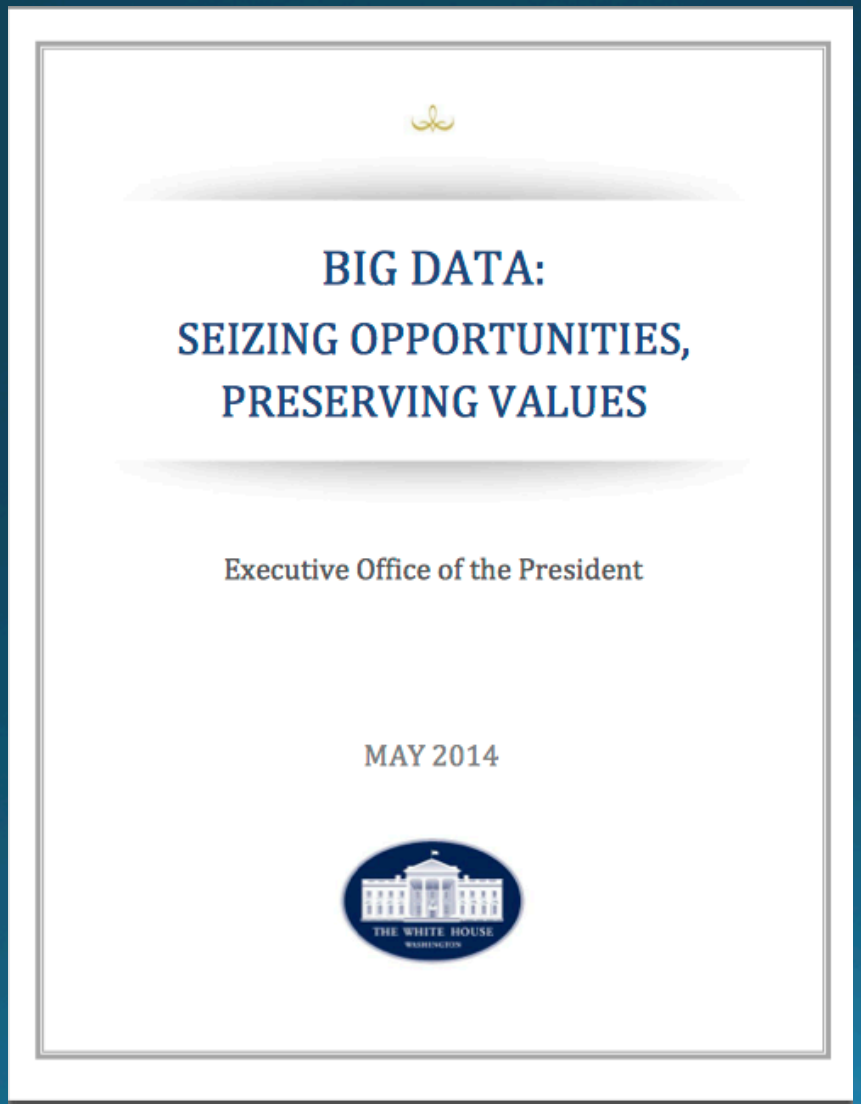
Big Data Task Force

Dr. Erin Smith
Executive Secretary, Big Data Task Force
NASA Science Mission Directorate
February 16, 2016

Ad Hoc Task Force on Big Data

- Authorized via NASA Advisory Committee (NAC) Charter part 13.1.vi, will report to NAC Science Committee (SC)



Big Data Motivations

- NASA has been charged with making NASA Earth observations data interoperable with other agencies such as NOAA and USGS.
- Providing interoperable data sets facilitates potentially new science, better use of data for decision support and potentially actionable science information.
- Making data more accessible can result in its use for completely new purposes.
- Large data sets demand understanding of architecture and tools for current and future needs.

Big Data Applications

- **Astrophysics:** supernova models, galaxy dynamics. Interoperability: application of planetary and Earth atmospheric models to exoplanets

- **Heliophysics:** multiple operating missions need data synthesis; modeling of solar events; radiation protection for spacecraft, human exploration

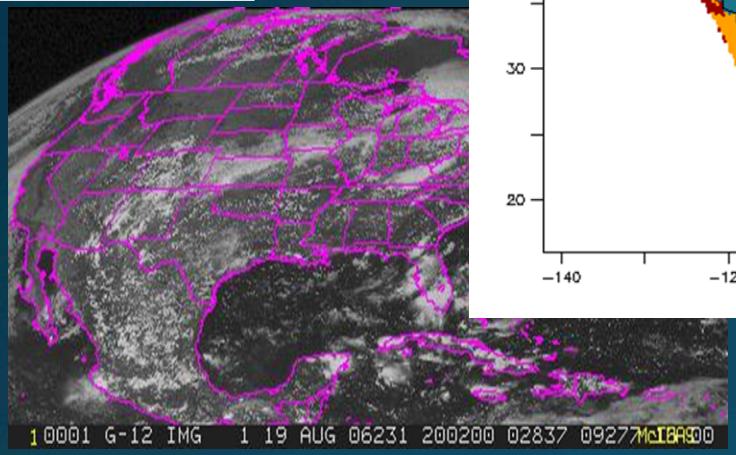
- **Earth Science:** climate models, use of observations by non-scientists (e.g. farmers); drought forecasting, disaster response

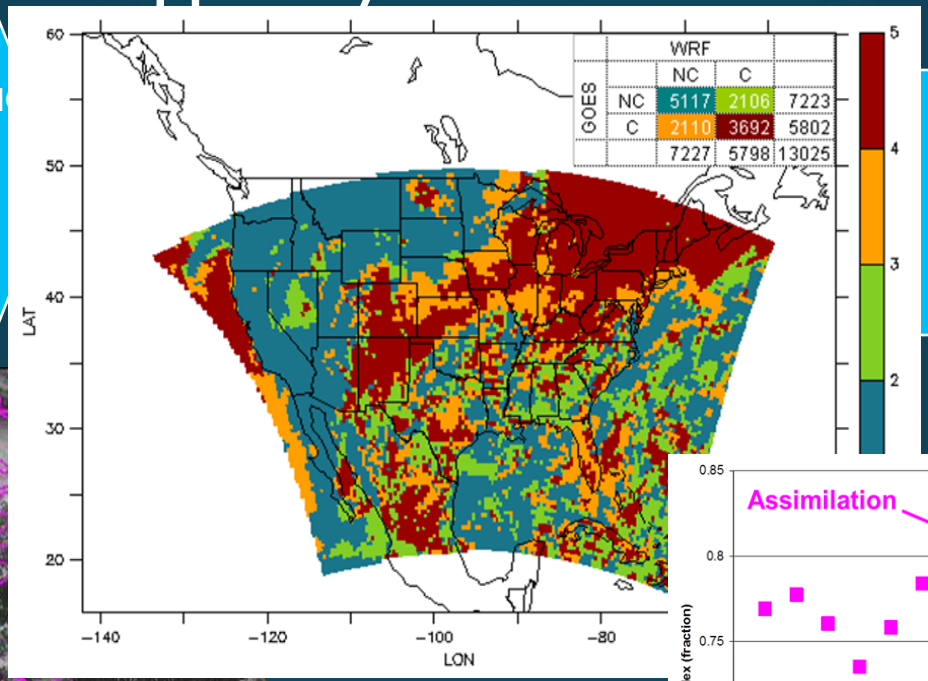
- **Planetary Science:** modeling of interconnected systems, formation models, database and archival projects

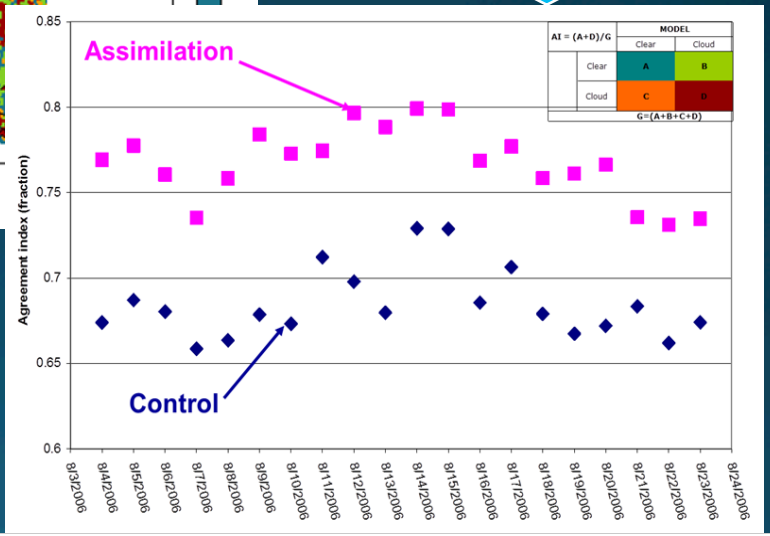
Incorporating Space-borne Measurements to Improve Air Quality Decision Support Systems for Texas

CLOUD LOCATIONS & TIMIN FROM SATELLITE INGESTED INTO THE AIR QUALITY MODEL USED TO PLAN ACCEPTABLE EMISSIONS

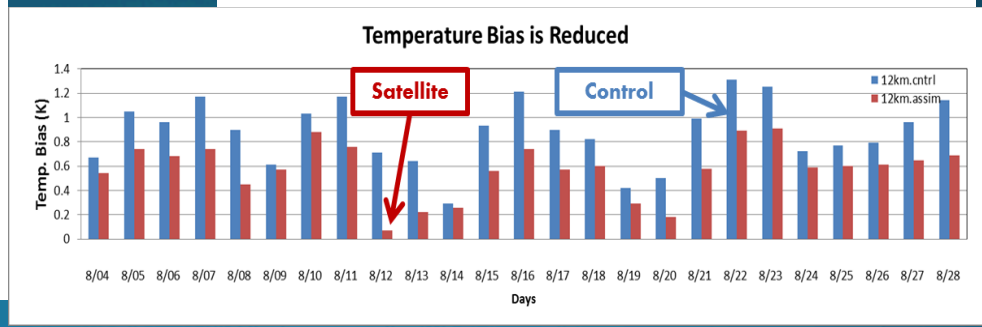
IMPROVED AIR QUALITY PLANNING AND REGULATORY DECISIONS







- The temporal and spatial location of clouds have a large impact on the projected air quality given a set of emissions. This tool is designed to provide accurate cloud information.
- Texas Commission for Environmental Quality (TCEQ) used this tool in their latest State Implementation Plans (SIPs)
- The State of Texas joined NASA Applied Sciences in funding 30% of the Project (~$310K)



Integrated Radiation Protection Strategy Enables Human Mars Exploration

Integration across Research and Technology Required…

National Aeronautics and Space Administration

Mission and Architecture Systems Analysis

Near Earth Asteroid Systems

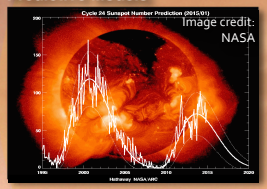
Mars MTV-02

In-situ Resource Utilization

Active Shielding Concepts

Environmental Modeling, Monitoring, and Prediction

Predictive Models



Precursor Data — MSL RAD

On-board Dosimetry– ISS TEPC

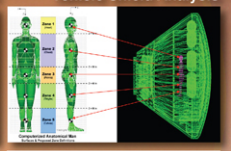
Crew Selection and Operations

Individualized Risk Assessment

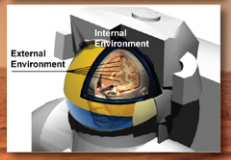
Integrated Radiation Protection System Design and Analysis

Design and Optimization Tools

Crew Exploration Vehicle Shield Analysis



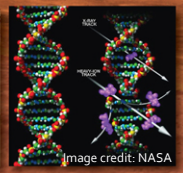
Images credit: OLTARIS.nasa.gov (top) and NASA (lower)



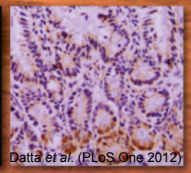
High Energy Nuclear Physics and Transport

Radiobiology and Biological Countermeasures

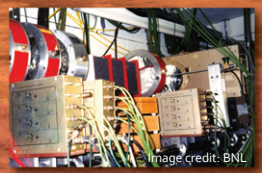
NASA Space Radiation Lab at Brookhaven National Laboratory



X-ray vs. Heavy ion Track Damage to DNA


Datta *et al.* (PLoS One 2012)
Leukemia induction with GCR — Mouse Model

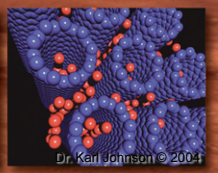
Innovative Multi-Purpose Shield Solutions



Heavy Ion Testing of Inflatable Shield Prototype

Water Filled Composite Shield Sections

Reconfigurable Personal Shielding



Hydrogen Storage BNNT

Background Image credit: NASA

Terms of Reference

NASA ADVISORY COUNCIL

SCIENCE COMMITTEE

AD HOC TASK FORCE ON BIG DATA

TERMS OF REFERENCE

The Ad Hoc Task Force on Big Data (hereinafter, "Task Force") is an ad hoc task force of the NASA Advisory Council's (NAC) Science Committee supporting the advisory needs of the NASA Administrator, the NASA Science Mission Directorate, the other NASA Mission Directorates, and the NASA Chief Information Officer, as required. The scope of the Task Force includes all NASA Big Data programs, projects, missions, and activities. The Task Force will focus on such topics as exploring the existing and planned evolution of NASA's science data cyber-infrastructure that supports broad access to data repositories for NASA Science Mission Directorate missions; best practices within NASA, other Federal agencies, private industry and research institutions; and Federal initiatives related to big data and data access.

Per NASA Policy Directive (NPD) 1150.11, *Federal Advisory Committee Act Committees,* the Task Force will be managed under procedures that ensure the same spirit of openness and public accountability that is embodied by the Federal Advisory Committee Act (FACA). This includes public meetings as appropriate and public access to Task Force records.

MEMBERSHIP

The membership of the Task Force will consist of leading authorities with relevant expertise drawn from industry, academia, independent researchers, and Government institutions. The Associate Administrator for Science Mission Directorate, after consultation with the Administrator, will appoint the Task Force Chair and the members. Appointments will be for a limited duration not to exceed a two-year term, to be determined by the Associate Administrator for Science Mission Directorate in consultation with the NASA Administrator. The Task Force Chair, in consultation with the Science Committee Chair, will select a Vice Chair from among the members of the Task Force. The Task Force will have between five to ten members.

MEETINGS

The Task Force will meet approximately three times a year, usually in advance of the Science Committee meetings. Task Force meeting agendas will be approved pursuant to NPD 1150.11, by the Task Force Executive Secretary, after coordination with the Task Force Chair, and will be responsive to requests from the Administrator, the Associate Administrator for Science Mission Directorate, and the Council Chair. The Task Force Chair is not a Science Committee member but does report on activities of the Task Force at Science Committee meetings, upon the invitation of the Science Committee Chair. Recommendations of the Task Force shall be transmitted to the Science Committee and the NAC in the same way as subcommittee recommendations under the NAC Charter. In the event of a necessary absence by the Task Force Chair, the Vice Chair will represent the Task Force in the Science Committee meeting. The Vice Chair is not a member of the Science Committee. The Task Force Executive Secretary will publish notices of upcoming Task Force meetings in the Federal Register at least 15 days prior to each meeting.

REPORTING

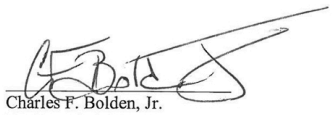
As appropriate, the Task Force Chair will report on the Task Force's findings and draft recommendations to the Science Committee at the Science Committee's public meetings, for deliberation. Records of each Task Force meeting, including meeting agenda, list of attendees, minutes, and presentations, will be kept by the Task Force Executive Secretary. The meeting agendas, presentations and minutes from Task Force public meetings will be posted to the Science Committee website for public access. After certification of the minutes by the Task Force Chair and Task Force Executive Secretary (not to exceed 90 days following the meeting), the minutes will be posted to the Science Committee website. In addition, the Task Force Executive Secretary will keep detailed financial records, member appointment records, and other pertinent records throughout the year, and provide annual summary input on the Task Force activities to the Science Committee Executive Secretary for NASA's required annual fiscal year report on its FACA advisory committees to the General Services Administration (i.e., FACA Database). The Task Force Executive Secretary will also coordinate with the members to ensure that all legal/ethics requirements pertaining to Special Government Employee appointments are met. The requirements include, but are not limited to, timely filing of annual financial disclosure reports and annual ethics training by a representative of NASA's Office of General Counsel (OGC), and seeking advice from the NASA OGC Ethics Team for any questions concerning Government ethics requirements that may arise.

ADMINISTRATIVE PROVISIONS

The Task Force Executive Secretary will be appointed by the Associate Administrator for Science Mission Directorate, following coordination with the Science Committee Executive Secretary and Council Executive Director. The Task Force Executive Secretary will work with the Task Force Chair to coordinate meetings, agendas, speakers, etc. for Task Force meetings. Staff support and travel funds for the Task Force Chair and members will be provided by the Science Mission Directorate. Other NASA Mission Directorates may provide support for specific activities, as appropriate.

DURATION

The Task Force Terms of Reference will terminate at the discretion of the Administrator, or at the end of two years from the date of signature noted below, whichever comes first. If the Council's Charter is terminated or expires, the Task Force will terminate. If the Task Force terminates, the Task Force terms of reference also terminate, and all appointments to the Task Force terminate.



Charles F. Bolden, Jr.
NASA Administrator

1/8/2015
Date

Membership

- Broad call for nominations that produced an initial list of 70 potential candidates
- Consultation from
  - 1) SMD Divisions

    Earth Science, Planetary Science, Astrophysics and Heliophysics and Office of Chief Information Officer
  - 2) NAC Committees

    Science, Aerospace, Technology Innovation and Engineering, Institutional, Human Exploration and Operations
  - 3) Science Committee Subcommittees

    Earth Science, Planetary Science, Planetary Protection, Heliophysics, Astrophysics

Terms of Reference

- Explore existing and planned evolution of NASA's science data cyber-infrastructure
- Catalogue best practices within NASA, other Federal agencies, private industry and research institutions
- Investigate federal initiatives related to big data and data access
- As appropriate, the Task Force Chair will report on the Task Force's findings and draft recommendations to the NAC Science Committee at its public meetings for deliberation
- The Task Force may produce work products such as white papers or reports as appropriate

Challenges

- Definitions: "Big Data" means different things to different groups.
- Engineering: Architecture, infrastructure and logistics of storing/accessing/using large data sets
- Data rights: IT security, Intellectual Property
- Communication: Visualization, training
- Large data sets demand understanding of architecture and tools for current and future needs

Status

- Terms of Reference signed
- Executive Secretary: Dr. Erin Smith, NASA Ames Research Center
- Nominees:
  - 6 appointed
  - 2 well into clearance process
  - 2 chosen more recently, beginning clearance process
- Ethics Training completed for members

Discussion Topics for Science Committee

1. What are the opportunities you see in finding synergies in the application of big data across NASA's scientific fields?
2. Within your scientific disciplines are there any particular areas you can identify as benefiting from the work of the Big Data Task Force?
3. What kinds of linkages between scientific disciplines should be established?
4. What topics would the NAC Science Committee want the Task Force to address?
5. What points should be focused on as Task Force is stood up?
6. What kinds of products should the Task Force produce (talks, white papers, formal reports, etc.)?

Feedback from Science Committee

- Pull membership from commercial entities, other (non-NASA) sciences that deal with big data
- Include feedback from cutting edge of data analytics, storage, access, etc. (e.g. Google, Microsoft…) as well as ground-based sciences (e.g. LSST)
- Topics to focus on: interoperability, data visualization, data permanence, and data usage
- Act as a point of contact for community and other federal entities (e.g. NSF)

NAC SC Requests Feedback

- NAC SC at its July 2015 meeting asked the Subcommittees to discuss and provide feedback on the charge of the Task Force

- Presentations and discussion occurred at the 4 Subcommittee meetings held (HPS, PSS, APS, ESS)

- Question posed:

Regarding the Ad Hoc Big Data Task Force (BDTF), what would your Subcommittee like to see as the task force's

1) Needs Statement – the need or problem the task force is addressing
2) Goal
3) Objectives
4) Tasks
5) Deliverables

Subcommittee Feedback: BDTF Linkages/Dialogues

- Link with existing efforts on big data (e.g. Planetary Data System workshops, virtual astronomy model, STScI work, LSST work, National Climate Assessment data online, Community Coordinated Modeling Center).
- Invite persons involved in those efforts to participate (data system managers, modelers, simulation experts, industry working with NASA Ames, NASA archives people)
- Important to leverage with industry partners, and learn methods from universities and government science centers

Subcommittee Feedback: BDTF should Assess/Inform

- Within-discipline - Understand big data first within subfields of discipline and ascertain if there is leveraging/cooperating. Big Data Task Force could facilitate identification of within-discipline synergies, needs and gaps.

- Cross-pollination - Inform and learn across disciplines/NASA SMD divisions/Federal agencies as *big data is happening in all sciences*. Big Data Task Force could facilitate the exchange of knowledge of how different practices and tools are used, insights, developments and lessons learned.

Task/Deliverable: Big Data Task Force could identify the best way to gather data on big data needs and activities and have a feedback mechanism so the Subcommittees and disciplines benefit from this effort (e.g. survey to industry members, AAS town hall)

Subcommittee Feedback: Big Data Aspects to Address

Usability of Data, Datasets and Databases
- Management and Access
- Utilization (including near real-time)
- Interoperability (e.g. integration of heterogeneous datasets, harmonization of data with different timeframes)

Analysis
- Data mining/analysis of large datasets
- Algorithm and statistics development

Storage
- Data curation
- Archiving tools and technology

Visualization (e.g. hyperwall)

Utilizing State-of-the-Art IT Systems and Tools

Task Force Goals

1. What are opportunities in Big Data, of particular interest to NASA and science stakeholders, that this task force should investigate?
   - What kinds of linkages between scientific disciplines should be established?
   - What capabilities already exist?
   - What capabilities exist for interoperable data sets? What can be developed?

2. What kinds of products should the Task Force produce (talks, white papers, formal reports, etc.)?

3. What entities/subject-matter experts should be invited to work with the task force or asked for input?