

# The Potential Value and Impacts of a Data Decadal Survey

A Panel Discussion

ESIP Data Decadal Survey Cluster

Anne Wilson, Moderator

ESIP 2013 Summer Meeting

July 10, 2013

# What is a Decadal Survey?

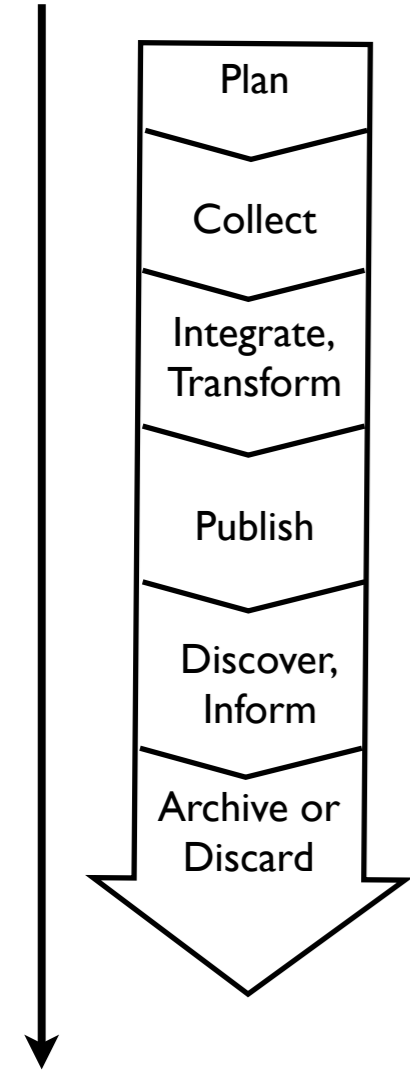
- Broad study of a topic or area
- Requested, funded by government agencies, other organizations
- Coordinated by National Research Council (NRC), conducted under a Board(s)
- Provides community consensus on research goals and priorities for moving forward over next 10 years
- Used to prioritize research areas and to focus efforts, resources accordingly
- Has been shown to be a robust method for developing goals and supporting objectives in pursuit of high priority science questions

# A Decadal Survey for Data?

- Idea raised last fall by Dan Baker at Boulder Earth and Space Science Informatics Group (BESSIG) panel discussion on “Science Policy for Scientists and Engineers”
- Session at ESIP 2013 winter meeting, “Is it Time for a Data Decadal Survey?”, a cluster was formed
- What would it mean?
- What might it enable?
- What scope?

# Data life cycle activities, concerns, and perspectives

## Data Life Cycle Model



time

## Activities

Planning and Production

Reqs Def  
Planning  
Development  
Deployment  
Operations

Data Management

Collection  
Processing  
Quality Control  
Documentation  
Cataloging  
Dissemination  
Preservation  
Stewardship  
Usage Tracking  
Final Disposition

Usage

Discovery  
Reception  
Understanding  
Analysis  
Value-Added Products  
User Feedback  
Citation  
Tagging  
Gap Assessment

## Overarching Concerns

Principles  
Governance  
Resources  
Standards  
Architecture  
Assessment



## Perspectives (stakeholders?)

Scientific  
Technical  
Organization/Institutional  
Economic/Financial  
Policy/Legal  
Socio-cultural

# Our Panelists

- **Stan Ahalt**, Director, Renaissance Computing Institute (RENCI)
- **Dan Baker**, Director, Laboratory for Atmospheric and Space Physics (LASP), Professor, Astrophysical and Planetary Sciences, Chair of 2013 - 2022 Heliophysics Decadal Survey
- **Michael Tiemann**, Vice President of Open Source Affairs, Red Hat Inc.
- **Todd Vision**, Associate Director for Informatics, National Evolutionary Synthesis Center (NESCent), Professor of Biology, UNC

Please see full biographies in the program

# Questions for the panelists

- Do you have ideas for science that could be done with improved data management and stewardship (DMAS) that cannot be done now? If so, what?
- Do you know of serious gaps in DMAS that negatively impact science? If so, tell us about some.
- Do you see a need for a Data Decadal Survey? If so, why? If not, why not?
- What is your vision for the future regarding scientific data? Please be bold and include fanciful, idealistic, lofty, and even utopian ideas.

# More Panelist Questions

If in support of a survey:

- What do you think are the big questions around DMAS that must be addressed? What are the highest priorities?
- Is a Decadal Survey the right vehicle, or should it be something else?
- What should be the scope of the survey?
  - Given that either extreme of depth or breadth is of less general use, what exactly should we target?
  - Earth Science only or broader? For example, does it make sense to start "small" in the Earth Sciences and then generalize? Or, as data management problems across all domains are basically similar, should we start with the more general and perhaps go into greater detail later?
  - Data? Software? Methodologies?
- What would be metrics for assessing survey success?
- A Data Decadal Survey is risky because the topic is extremely broad, the community of data users is vast and heterogenous (and can include commercial interests), and the outcome will not be a focused mission, facility, or research initiative. How should that risk be managed?

# Schedule

- 12 minutes per panelist (no Q & A)
- 32 minutes Q & A
- 5 minutes closing