A Maturity Model for Earth Science Data and Information Services

Bruce R. Barkstrom
&
John J. Bates

OWGDI S
Objective

• Reduce difficulty and confusion in community about scientific data stewardship
• Produce an easily understood way of identifying maturity of data products and science data stewardship approaches
• Help identify areas needing improvement
OWGDIS Requirements

- Interagency views of data maturity and science data stewardship needs reflect wide diversity
  - Vocabulary
  - Experience
  - Background

- Need understandable approach
A Simple Maturity Model

- Represent data maturity in terms of three separate dimensions:
  - Scientific Maturity
  - Preservation Maturity
  - Societal Impact
- Use CMMI-like levels
- Total maturity is simply length of vector
CMMI Maturity Levels

1. Initial – Unpredictable results
2. Managed – Repeatable performance
3. Defined – Cross-project interoperability
4. Quantitatively Managed – Improved performance + Compliance with Federal Enterprise Architecture
5. Optimized – Rapidly configurable performance + Continuous Process Improvement
Component Maturity

- Identify key attributes of maturity in each dimension
- Develop maturity ranking for each attribute on scale of 0 to 5
- Summarize component maturity by weighting each attribute
  - Simplest weight = 1/Number of attributes
  - Develop more complex weightings after experience with approach
- Advantage: can do much of work with simple spreadsheet
Scientific Maturity Key Attributes

- Physical Understanding of Measurement Process
- Measurement of Key Instrument Characteristics
- Public Accessibility of Data Processing
- Rigorous Validation
Preservation Maturity Key Attributes:

- Systematic Approach to Guaranteeing Preservation of Data Understanding
- Systematic Reduction of Threats to Preservation
- Assurance of Preservation Cost Effectiveness
Societal Benefit Key Attributes

• Bibliometric Metrics
  - Publications and Citations

• Scientific Community Knowledge
  - Data use, including interdisciplinary data fusion and statistical studies

• Economic and Policy Utility
  - Market valuation increase
  - Reduction in time to influence policy
  - Benefit/Hazard Reduction resulting from data use
Some Caveats

• Using a Maturity Model will be exploratory – and iterative
  – No expectation we’ll get it “right” the first time through

• Community Diversity must be incorporated
  – Different views of data processing, calibration, validation, need for knowledge preservation
  – Different vocabularies

• Deep Uncertainty needs to be incorporated
  – Diversity of opinions on areas of scientific controversy and value need common framework and disciplined discussion – openness a key
  – Including “societal benefit” is very difficult and risky
Key Benefits

- Allows OWGDIS (and OWG) to develop an approach consistent with NRC Recommendations on Metrics
- Open Process
  - Can surface divergent needs and opinions
  - Can provide disciplined forum for discussion and resolution of differences
- Periodic Evaluation is required
  - Incorporate new information and deeper thought
  - Evaluation allows new directions
Proposed Process

• Community Discussion
  – Disseminate notional maturity model to interested community (scientists, agencies, data managers)
  – Solicit community comments within 60 days [end ~Aug. 1]

• Edit Model
  – John Bates, Bruce Barkstrom as editors; seek up to 7 contributors (volunteers) to help

• Revised Model Distribution and Final Document
  – Disseminate revised document [~Sept. 1]
  – Face-to-Face discussion at Fall Workshop [~Nov. 15]

• Suggest Annual Revision Cycle
  – Spring deadline may be needed to align with agency budget work