The PDS Approach to Data Quality Assurance

Anne Raugh

Planetary Data System Small Bodies Node

University of Maryland

ADASS XXVIII

14 November 2018

A Brief History of the PDS

- CODMAC, volume 1 (1982):

   "The present report on data management and computation was prepared in response to our perception that data problems were pervasive throughout the space sciences."

- Planetary Data Workshop Proceedings (1984):

   "It is noted with increasing alarm by many in the science community that valuable data sets are disappearing."
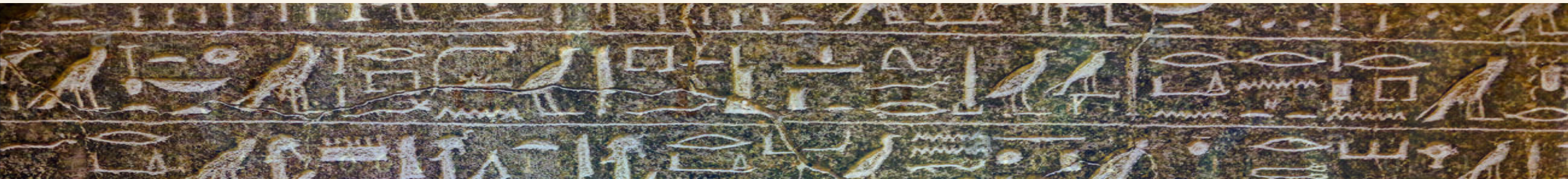
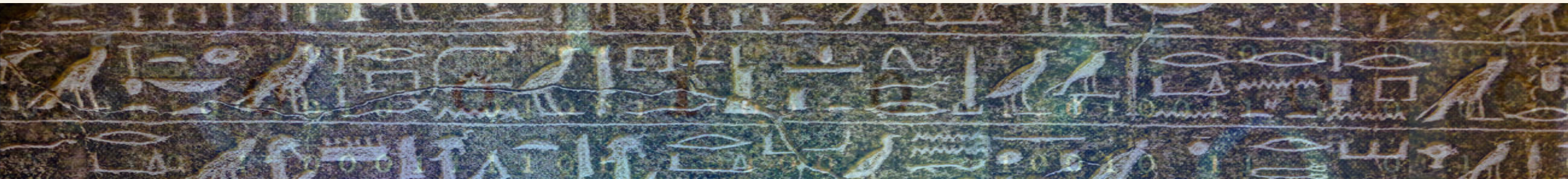   *The PDS was established c. 1984 to address this crisis.*

The PDS Mandate

Maintain the availability and usability of data resulting from NASA's planetary missions for future generations.

What is "Data Quality"?

Data must be:
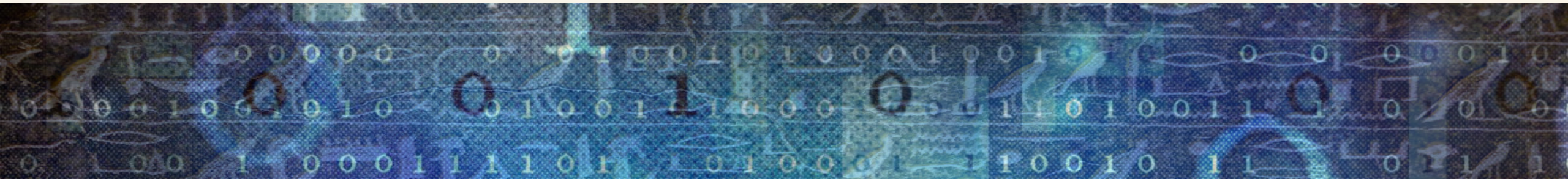
- Understandable
- Usable
- Complete

Quality Assurance Tools

1. PDS4 Information Model (IM)

2. External Peer Review (Refereeing)
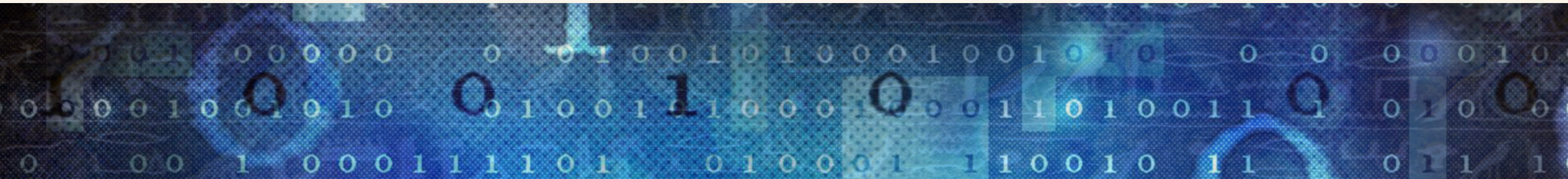
PDS4 Information Model

- Content Management
  - Minimum metadata requirements
  - Metadata templating (best practices)
- Schematic Validation – programmatic enforcement
- Programmatic accessibility

External Peer Review (Refereeing)

- Review Panel
    - Independent discipline experts for science review
    - PDS for standards compliance
- Process
    - Exercise the data
    - Review meeting and lien generation
    - Revision and acceptance testing

In summary

- PDS4 information model systematizes the metadata and formats needed to support completeness and usability of the candidate submission within the archive context.

- External peer review provides an independent verification and demonstration that the candidate submission is, in fact, complete and usable at time of acceptance.

~ ensuring data quality at the point of archive acceptance.