

Criteria for Assessing Repository Trustworthiness: An Assessment

John Steven Hughes¹ and Robert R. Downs²

¹ steve.hughes@jpl.nasa.gov

NASA Jet Propulsion Laboratory (JPL),
California Institute of Technology

² rdowns@ciesin.columbia.edu

NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network
(CIESIN), The Earth Institute, Columbia University

SCIDATACON 2016

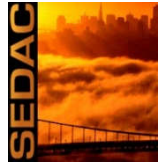
Denver, CO 11-13 September 2016

Session 100: Auditing of Trustworthy Data Repositories
Monday 12 September 2016, 14:00

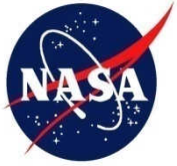




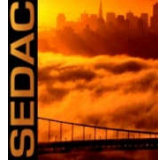
The OAIS Reference Model



- ISO 14721, Reference Model for an Open Archival Information System (OAIS)
- Framework for repositories to provide long-term stewardship of data and other digital information
- Used as a basis for assessment instruments and activities
- Developed by Consultative Committee for Space Data Systems (CCSDS) in 2003
- Published by ISO in 2003, reviewed and revised in 2012
- <https://public.ccsds.org/Pubs/650x0m2.pdf>
- ISO 14721:2012 **5-year review started in August 2016**



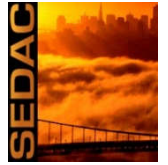
ISO Repository Trustworthiness



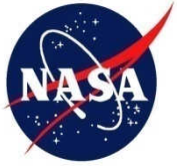
- ISO 16363, Audit and Certification of Candidate Trustworthy Digital Repositories
- Based on OAIS Reference Model
- Instrument for assessing trustworthiness of repositories
- Developed by CCSDS in 2011
- <https://public.ccsds.org/Pubs/652x0m1.pdf>
- Published by ISO in 2012
- ISO 16363:2012 5-year review started in August 2016
- **Simultaneous review with ISO 14721:2012 (OAIS RM)**



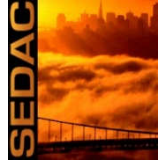
Pre-Review Assessment of ISO 16363



- Scope of preliminary assessment
 - Overall suitability of ISO 16363 to assess data repository trustworthiness
- Approach
 - Analysis of the ISO 16363 organization and metrics for measuring the trustworthiness of a data repository
- Method
 - Independently assessed ISO 16363 metrics for *Mandatory Responsibilities*
 - Attained consensus based on agreement of independent assessments
- Details examined
 - **Organization of ISO 16363:2012**
 - **ISO 16363 Metrics for measuring the *Mandatory Responsibilities* for compliance with the Open Archival Information Systems (OAIS)**



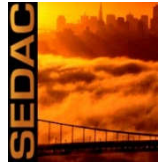
Organization of ISO 16363:2012



- Hierarchical organization of metrics into three sections*
 - Organizational Infrastructure
 - Digital Object Management
 - Infrastructure and Security Risk Management
- Hierarchical organization within sections
 - Subsections under each section
- Hierarchical organization of metrics
 - Sub-metrics under selected metrics
- Structure within metrics
 - Supporting text, Examples, and Discussion for each metric

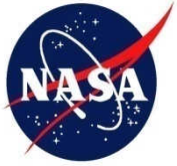


OAIS Mandatory Responsibilities*

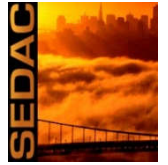


- **Negotiates for and Accepts Information**
 - Negotiate for and accept appropriate information from information Producers.
- **Obtains Sufficient Control**
 - Obtain sufficient control of the information provided to the level needed to ensure Long Term Preservation.
- **Determines Designated Community**
 - Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided, thereby defining its Knowledge Base.
- **Ensures Information is Independently Understandable**
 - Ensure that the information to be preserved is Independently Understandable to the Designated Community. In particular, the Designated Community should be able to understand the information without needing special resources such as the assistance of the experts who produced the information.
- **Follows Established Preservation Policies and Procedures**
 - Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, including the demise of the Archive, ensuring that it is never deleted unless allowed as part of an approved strategy. There should be no ad-hoc deletions.
- **Makes the Information Available**
 - Make the preserved information available to the Designated Community and enable the information to be disseminated as copies of, or as traceable to, the original submitted Data Objects with evidence supporting its Authenticity.

*Derived from: CCSDS. 2012. Reference Model for an OAIS.



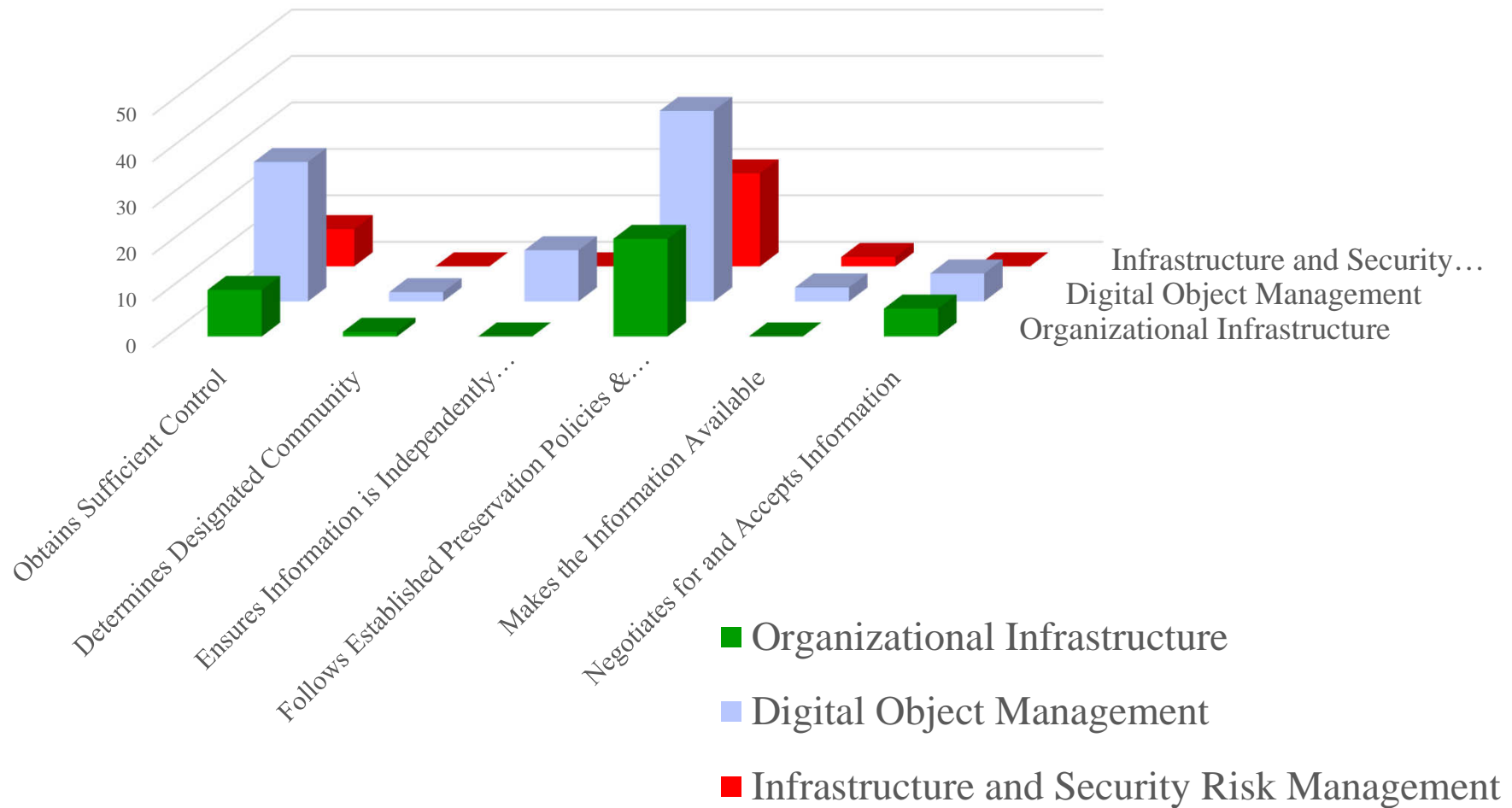
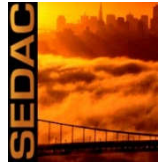
Results - ISO 16363 Organization



- Hierarchical organization of metrics into 3 sections
 - Organizational Infrastructure, Digital Object Management, and Infrastructure and Security Risk Management are **logical groupings**
- Hierarchical organization within sections
 - Subsections within each section reflect **logical organization** of standard
 - Metrics pertain to the section in which they appear
- Hierarchical organization of metrics
 - Sub-metrics under selected metrics introduces **potential redundancy**
- Structure within metrics
 - Metrics are described concisely, but **some could be explained further**
 - Supporting text, Examples, and Discussion provide context for each metric

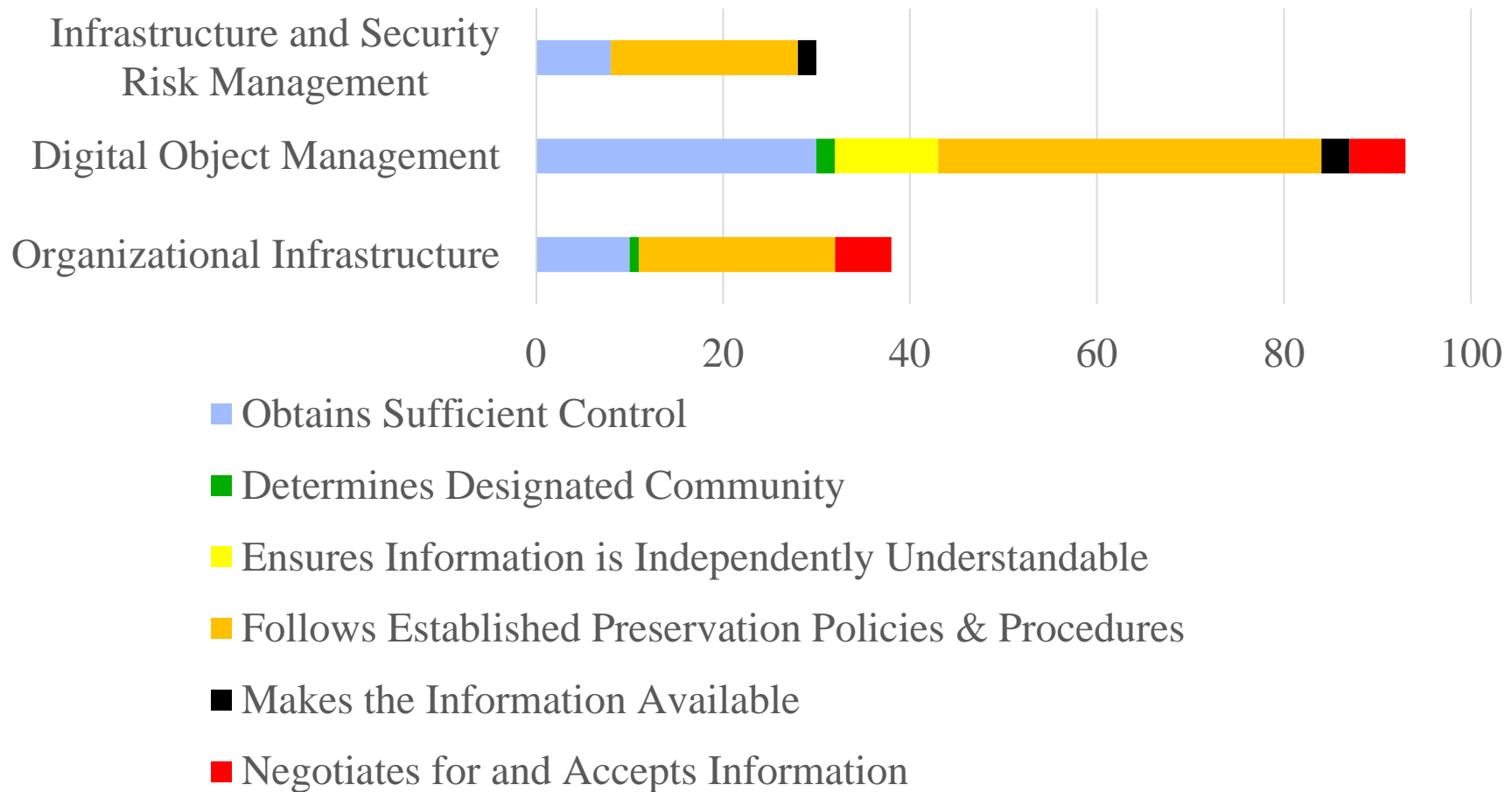
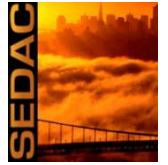


Requirements by Organizational Structure



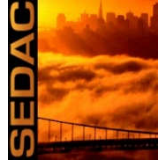


Requirements within Organizational Structure





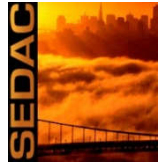
ISO 16363 Measurement of Mandatory Responsibilities



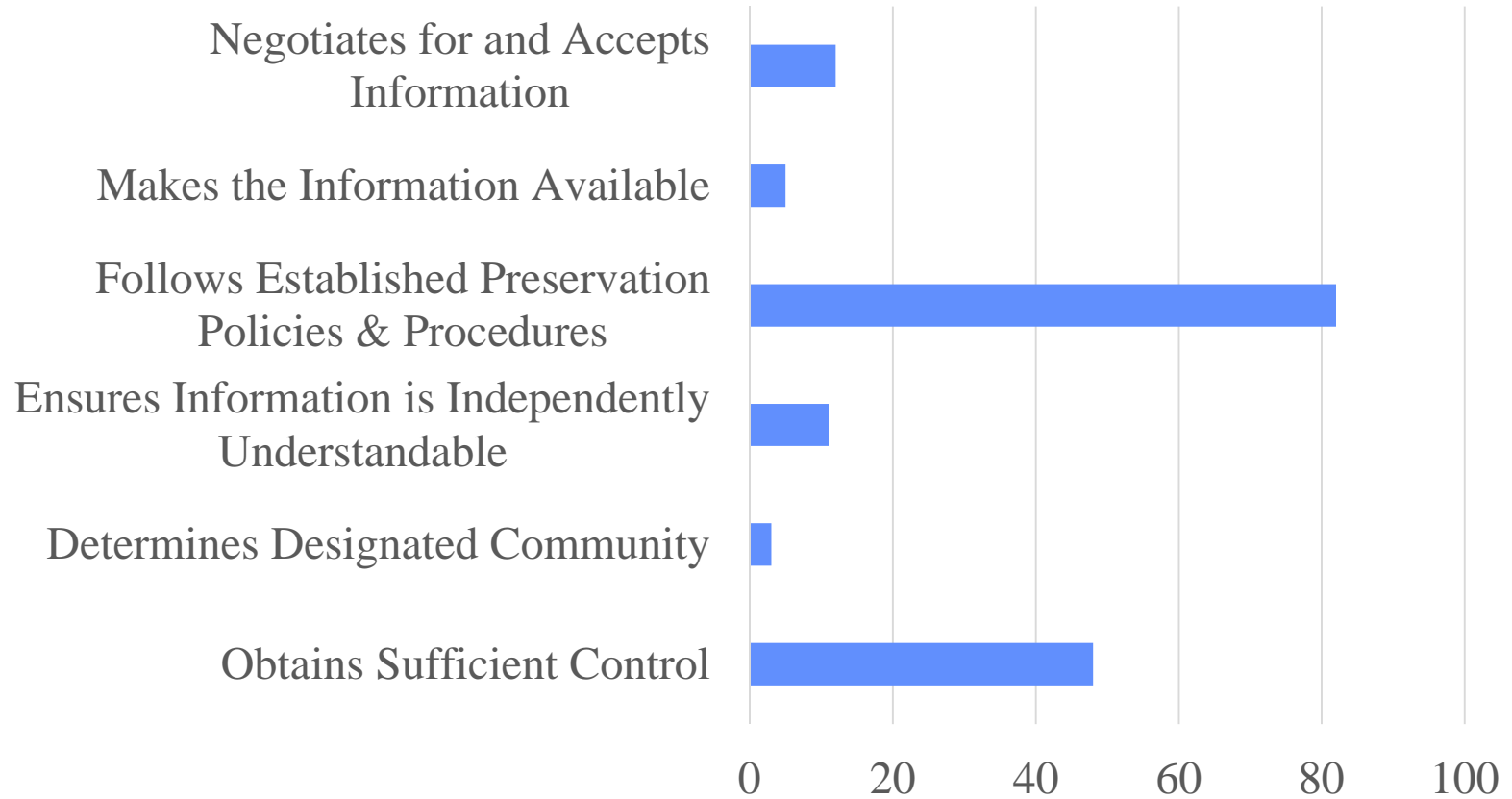
- 109 metrics and sub-metrics
- Multiple metrics measure each Mandatory Responsibility
- Many metrics measure multiple Mandatory Responsibilities
- **Mandatory Responsibilities measured more than others**
 - Follow Established Preservation Policies and Procedures
 - Obtains Sufficient Control
- **Mandatory Responsibilities measured less than others**
 - Negotiates for and Accepts Information
 - Ensures Information is Independently Understandable
 - Determines Designated Community
 - Make the Information Available



Metrics Measuring Mandatory Responsibilities



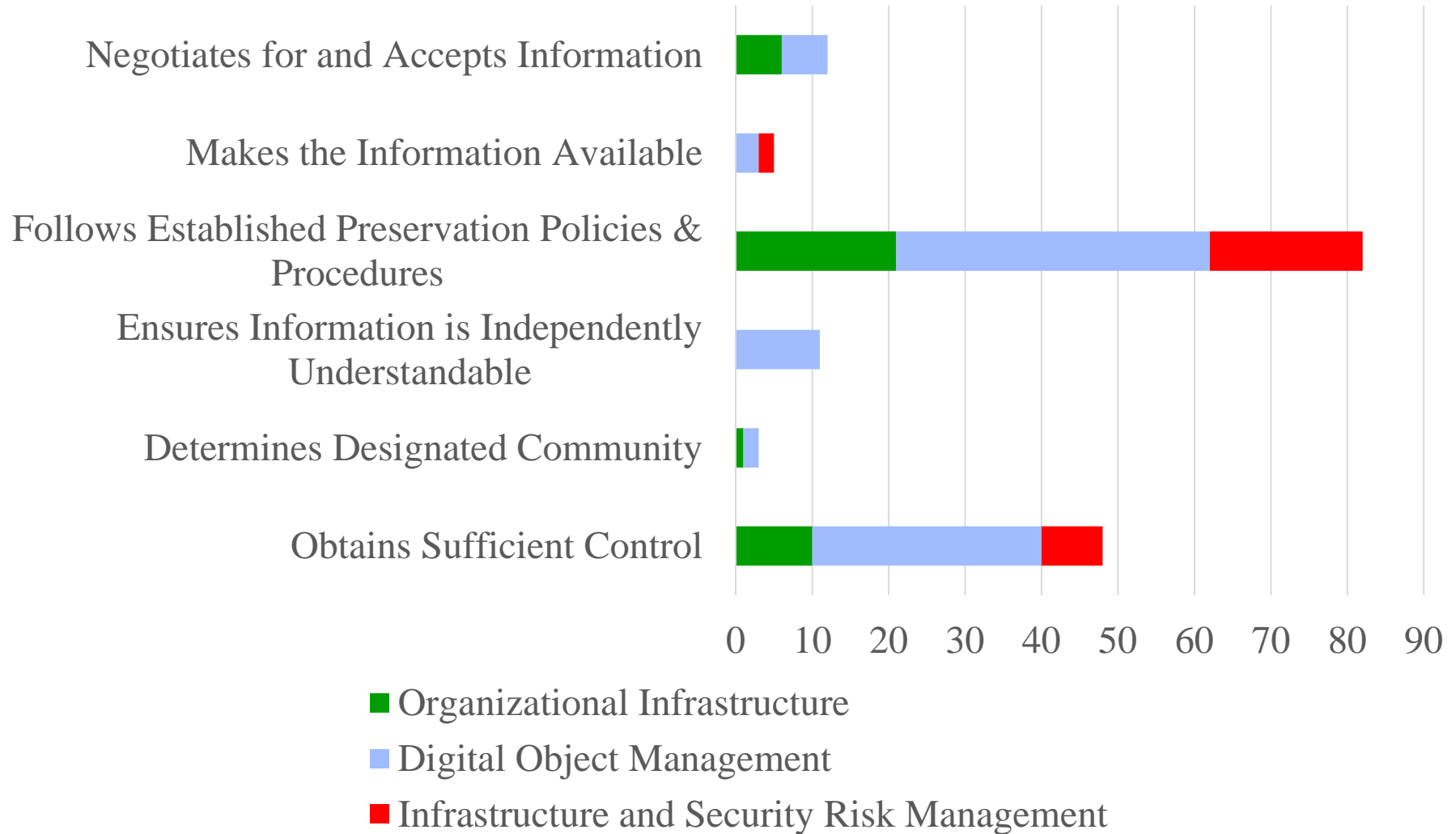
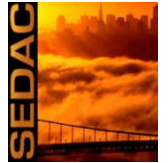
Mandatory Responsibilities

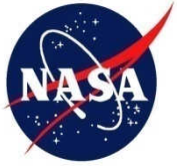


Number of Metrics Found that Measure Mandatory Responsibilities

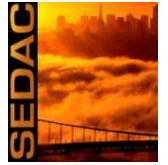


Metrics Measuring Mandatory Responsibilities within Organizational Structure of ISO 16363





Recommendations (3 Examples)



- Generally, **revise or update language** in metrics, including explanation, discussion, and examples for selected metrics to **reflect current capabilities and recent improvements** in data management practices
- Revise terms in metrics to **enable consistent interpretation**
 - "Obtain sufficient control for Preservation" implies legal issues (metric 3.5.1); business planning (metric 3.4.1); and tracking 4.1.6
- Combine or revise metrics for **consistent level of detail**
 - For example, the submetrics for Identifiers (4.2.4.x) and for identifying and managing risks (5.1.1.x) appear to be much more detailed than most submetrics.