

Group 1

Archival Information and Technology

Long Term Sustainment of Digital Information for Science and Engineering: Putting the Pieces Together

Team Report/Recommendation Template

Breakout Team ID	1 – Archival Information and Technology	
Identify: Problem or Issue	End user perspective: Identify data and metadata required, extent, depth, format and data types for specific users and scenarios	
Analyze: Root Cause	We don't know enough about who the end users are and how they will use the data.	
Recommendation	Collect end-user use cases	
Benefit	Predict current and future usage from end user perspective. Also will help us identify what should be saved and what should be discarded.	
Plan: Action(s) to implement	Owner/Time Frame	
Identify end users and collect use cases and/or survey current end users	Paul – engineering Bill – manufacturing Kate – govt. pubs SK Gupta – NSWC Gerry Graves – aero, Elec. Boat, Honeywell, GD	
Capture lessons learned from existing (including paper-based) archives (historical point of view)	Crispin Hales – eng. forensics	
Mechanism for sharing lessons learned, use cases (use wiki)	Josh (NIST)	

Long Term Sustainment of Digital Information for Science and Engineering: Putting the Pieces Together

Team Report/Recommendation Template

Breakout Team ID	1 – Archival Information and Technology	
Identify: Problem or Issue	Getting the right information from the creator for archiving from an archivist's point of view (which is to serve the end user)	
Analyze: Root Cause	Not having rationale, rich documentation required for archival packaging of data for future use	
Recommendation	Create guidelines/requirements for creators to follow, while recognizing that they won't be universally followed. Potential enforcement through Contractual mechanisms, peer review. Support creators through (semi-)automated tools.	
Benefit	Archival task is facilitated for the goal of supporting the end-user	
Plan: Action(s) to implement	Owner/Time Frame	
Examination of existence or lack of current guidelines in different communities	Vicky - crystallography	
Educating the creators of information – the need for development guidelines with archives and end-users		
Explore metadata extraction tools and watch for opportunities to influence vendors		
Establish Communities of Practice for archive submission requirements and guidelines – meta-guidelines and dynamic standards will emerge from this		

Long Term Sustainment of Digital Information for Science and Engineering: Putting the Pieces Together

Team Report/Recommendation Template

Breakout Team ID	1 – Archival Information and Technology	
Identify: Problem or Issue	How do we encode traditional metadata and unstructured metadata (annotations, social tagging)? What metadata is required for every object in the archive? What format-specific metadata are required (e.g., binary image versus text, content (image could be xray or CT-Scan)?	
Analyze: Root Cause	Different levels of detail in metadata needed depending on format and context.	
Recommendation	Follow guidelines and best practices for creation of structured as well as informal metadata. Classify different kinds of data objects and their metadata requirements. Look at NARA work done in WVU college of engineering (w/other partners).	
Benefit	Better searching, characterization of objects, validation of archives, metrics.	
Plan: Action(s) to implement	Owner/Time Frame	
Review LC sustainability criteria to determine need for metadata for various formats, and do gap analysis for missing formats.	Josh et.al	
Develop classification for engineering and science domains for different data formats.	NIST with input from COIs and NARA	
For each classification, articulate preferred set of data elements and preferred encoding.	COIs	