

Data Management Plans as a Research Tool

ACRL 2015 | Portland, Oregon, 25-28 March 2015

Amanda L. Whitmire — Oregon State University Libraries
 Jake Carlson — University of Michigan Library
 Patricia M. Hswe — Pennsylvania State University Libraries

Susan Wells Parham — Georgia Institute of Technology Library
 Lizzy Rolando — Georgia Institute of Technology Library
 Brian Westra — University of Oregon Libraries

What is the DART Project?

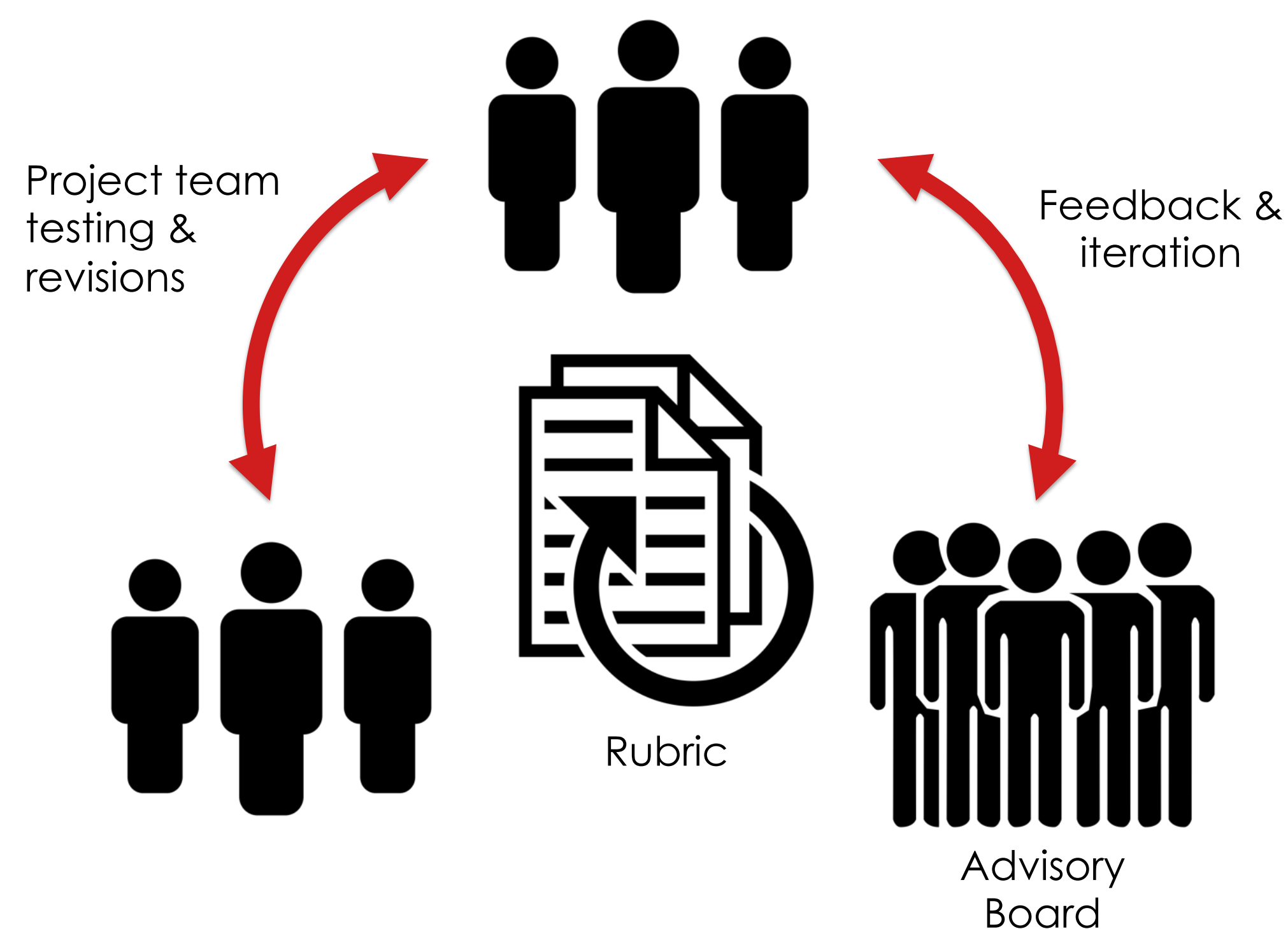
An IMLS-funded, multi-university study of data management plans that have been submitted to the National Science Foundation (NSF). This will be accomplished through the creation and use of a rubric, which will enable robust, consistent assessments across the research team. The goal is to turn the DMP into a research tool for academic librarians.

Project Outputs

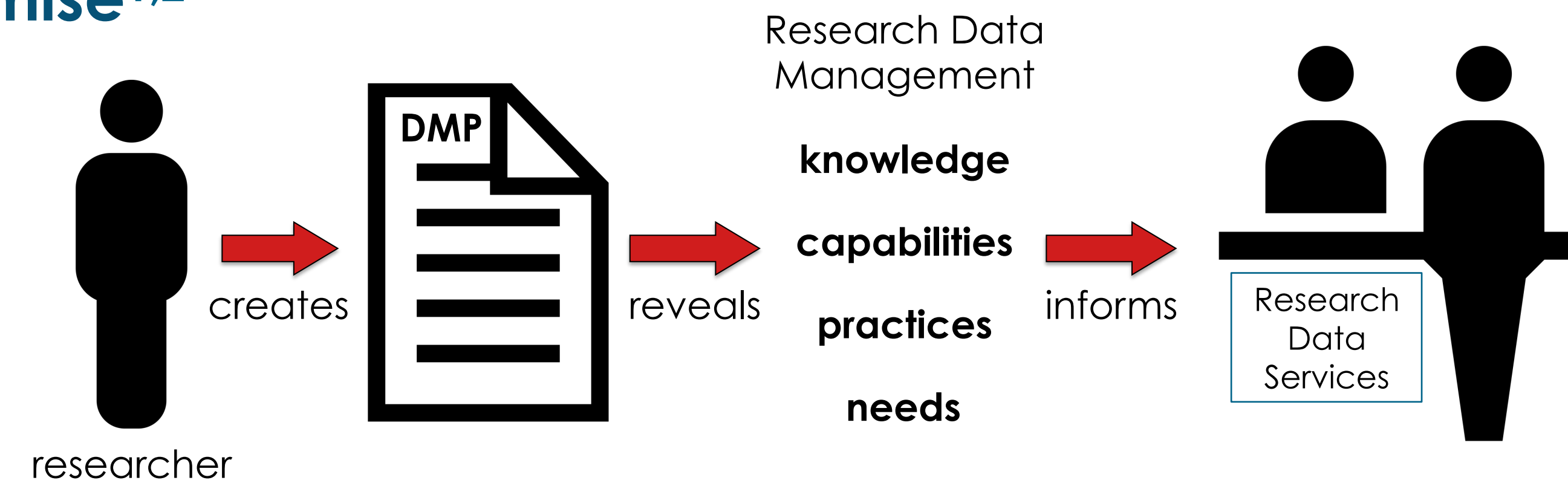
- 1) An **analytic rubric** to standardize the review of data management plans as a means to inform targeted expansion or development of research data services at academic libraries;
- 2) A study utilizing the rubric that presents the results of **data management plan analyses** at five universities.

Project Phases

- 1) Acquisition of DMPs
- 2) Iteratively develop analytic rubric
- 3) Use rubric to analyze DMPs
- 4) Finalize & share rubric
- 5) Share results of DMP analyses at 5 research intensive US universities



DART Premise^{1,2}



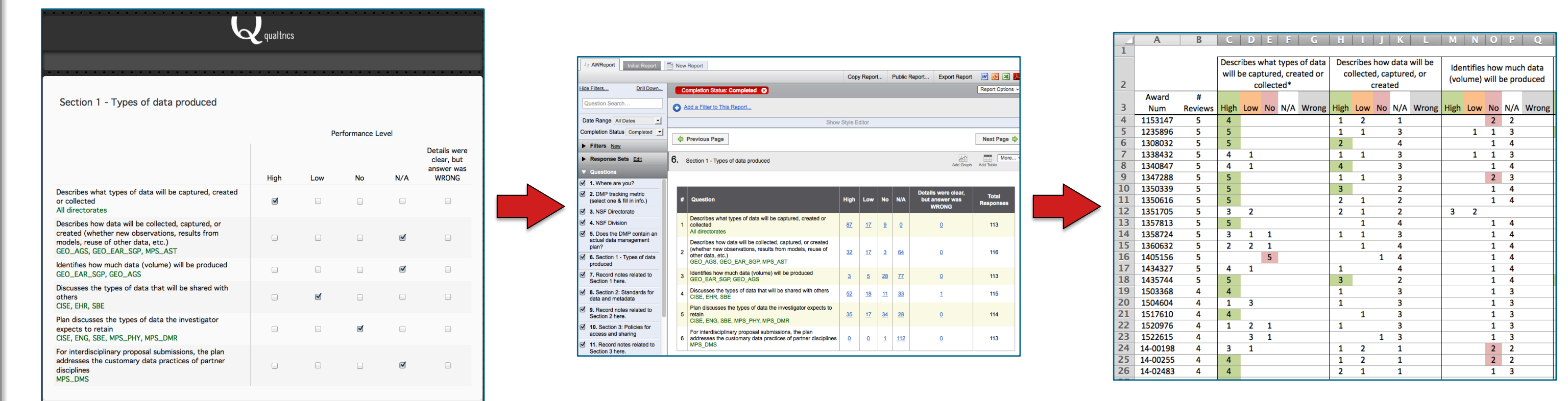
DART Rubric

A rubric is a descriptive scoring tool. An analytic rubric provides detailed information about sub-components of the work being assessed through independent evaluation of various criteria. An example of a rubric for evaluating NSF DMPs is shown below. This example is for the section regarding data and metadata formats (*please note*: this is not the finalized rubric).

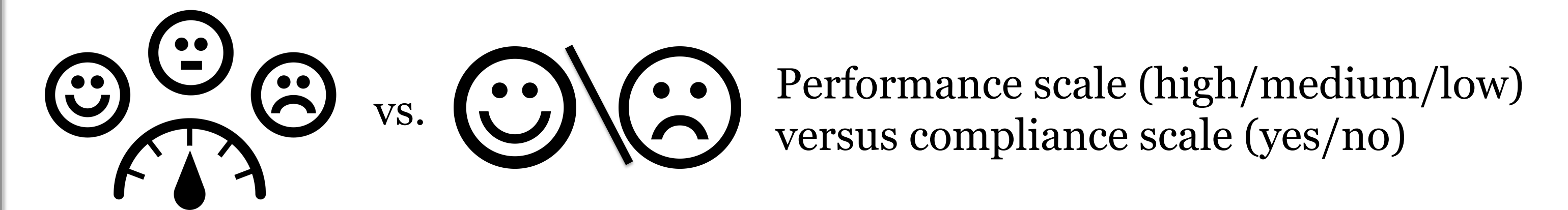
	Performance Level				
	Performance Criteria	Complete / detailed	Addressed criteria, but incomplete	Did not address issue	Directorates with this requirement
General assessment criteria	Identifies metadata standards or formats that will be used for the proposed project	The metadata standard that will be followed is clearly stated and described. If no disciplinary standard exists, a project-specific approach is clearly described. <i>Ex.: Data will be described using Darwin Core Archive metadata, and accompanied by readme.txt files providing information on field methods and procedures.</i>	The metadata standard that will be followed is vaguely stated. If no disciplinary standard exists, a project-specific approach is vaguely described.	The metadata standard that will be followed is not stated and no project-specific approach is described.	All
	Describes data formats created or used during project	Clearly describes file format standard(s) for the data. Examples: plain text, Excel, SPSS, R, Matlab, Access DB, ESRI Shapefile, TIFF, JPEG, WAV, MP4, XML, HTML, or other software-specific file formats <i>Ex.: Soil temperature data will be collected via datalogger and are exported from the logger in tab-delimited text files.</i>	Describes some but not all file formats, or file format standards for the data. Where standards do not exist, does not propose how this will be addressed.	Does not include any information about data format standards.	All
Directorate- or division-specific assessment criteria	Identifies data formats that will be used for storing data	Clearly describes data formats that will be used for storing data and explains rationale or complicating factors. Examples: hardcopy logs and/or instrument outputs, ASCII, XML files, HDF5, CDF, etc. <i>Ex.: NMR data will be saved in proprietary format to preserve embedded information, and converted to JCAMP files for ease of access and in case the proprietary systems fail or become unavailable.</i>	Only partially describes data formats that will be used for storing data and/or the rationale or complicating factors.	Does not describe data formats that will be used for storing data and does not provide a rationale or discuss complicating factors.	GEO AGS MPS AST MPS CHE
	If the proposed project includes the use of unusual data formats, the plan discusses the proposed solution for converting data into more accessible formats	Explains how the data will be converted to a more accessible format or otherwise made available to interested parties. In general, solutions and remedies should be provided. <i>Ex.: Still Images from microscopes will be converted from proprietary formats to Open Microscopy Exchange format (OME-Tiff) for preservation and sharing.</i>	Vaguely explains how the data may be converted to a more accessible format or otherwise made available to interested parties.	Does not explain how the data will be converted to a more accessible format or otherwise made available to interested parties.	MST AST MPS CHE

1. Parham, S. W. and Doty, C. (2012). "NSF DMP content analysis: What are researchers saying?" *Bul. Am. Soc. Info. Sci. Tech.*, 39: 37–38. doi: 10.1002/bult.2012.1720390113
 2. Mischo, W. H.; Schlembach, M. C.; and O'Donnell, M. N. (2014). "An Analysis of Data Management Plans in University of Illinois National Science Foundation Grant Proposals." *Journal of eScience Librarianship* 3(1): Article 3. <http://dx.doi.org/10.7191/jeslib.2014.1060>

Current Workflow



Challenges



Balancing granularity with what we could reasonably expect from PIs using vague agency guidance



Need to build a greater path for consistency - reduce areas of having to make a decision on something

What's Next?

- Full analysis using rubric:
- Review >100 DMPs each
- Intra- & inter-institution analyses

Share the rubric & our research results

