



EOSDIS

NASA'S EARTH OBSERVING SYSTEM
DATA AND INFORMATION SYSTEM

Big Earth Data Initiative (BEDI) Metadata Improvement: Case Studies

John Kozimor, Ted Habermann and John Farley

This work was supported by NASA/GSFC under
Raytheon Co. contract number NNG10HP02C

Big Earth Data Initiative (BEDI)

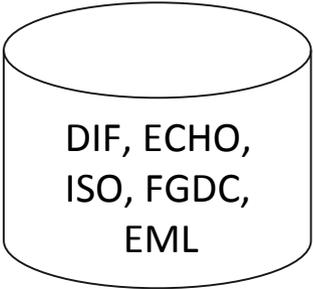
The Big Earth Data Initiative (BEDI) invests in standardizing and optimizing the collection, management and delivery of U.S. Government's civil Earth observation data to improve discovery, access use, and understanding of Earth observations by the broader user community.

Complete and consistent standard metadata helps address all three goals.



Many agencies included.

Projects/Communities

Projects	Big Earth Data Initiative
Communities	 <p>NASA, NOAA, USGS, USDA, EPA</p>
Dialects	 <p>DIF, ECHO, ISO, FGDC, EML</p>

Overview

- Tools and techniques has been developed to:
 - Compare recommendations and dialects
 - Identify the structure of metadata collections
 - Compare the structure of metadata collections
 - Evaluate and measure metadata completeness with respect to recommendations
 - Evaluate and measure metadata completeness with respect to specific organization goals
- These tools and techniques were applied to NASA metadata collections.
- The results of this analysis identify how well CMR metadata collections comply with UMM-Collection Profile.

Case Study 1:

Do my dialects support my requirements?

Overview: Metadata recommendations change as new communities and needs emerge. Metadata management tools are driven by dialects. Changing those tools and training people are difficult, so adoption of new dialects is relatively slow.

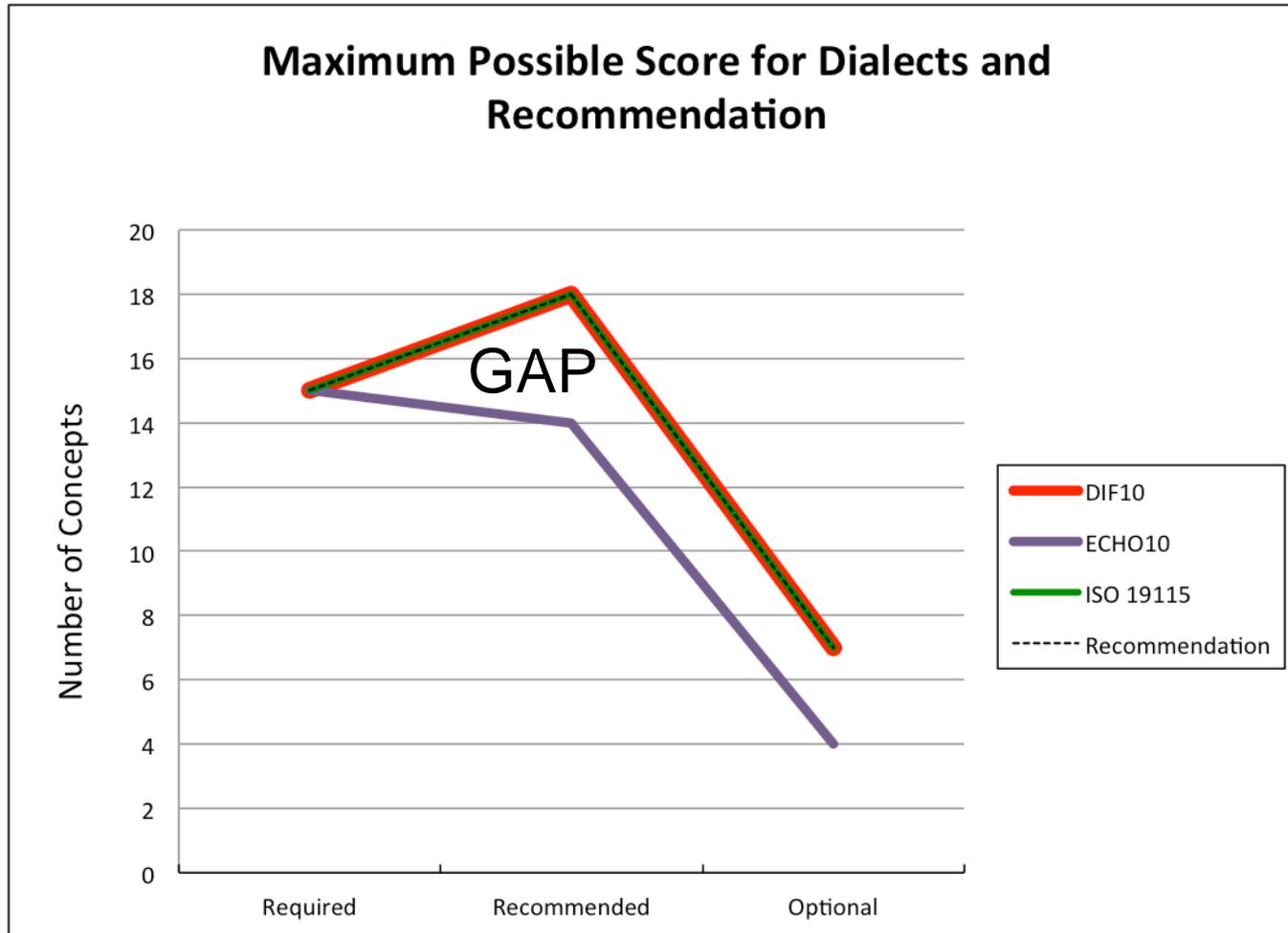
Purpose: Identify gaps between existing organizational capabilities (dialects) and new recommendations (requirements).

Scope:

- ISO 19115, ECHO, DIF10 and DIF dialect support with respect to the UMM-Collection Profile

Case Study 1: Results

UMM-Collection Recommendation / NASA Dialects



Case Study 2:

How complete is CMR metadata with respect to the UMM-Collection Profile

Overview: The UMM-Collection Profile specifies metadata concepts (required, recommended and optional) for use with CMR collection metadata. Required concepts are mandatory in CMR metadata.

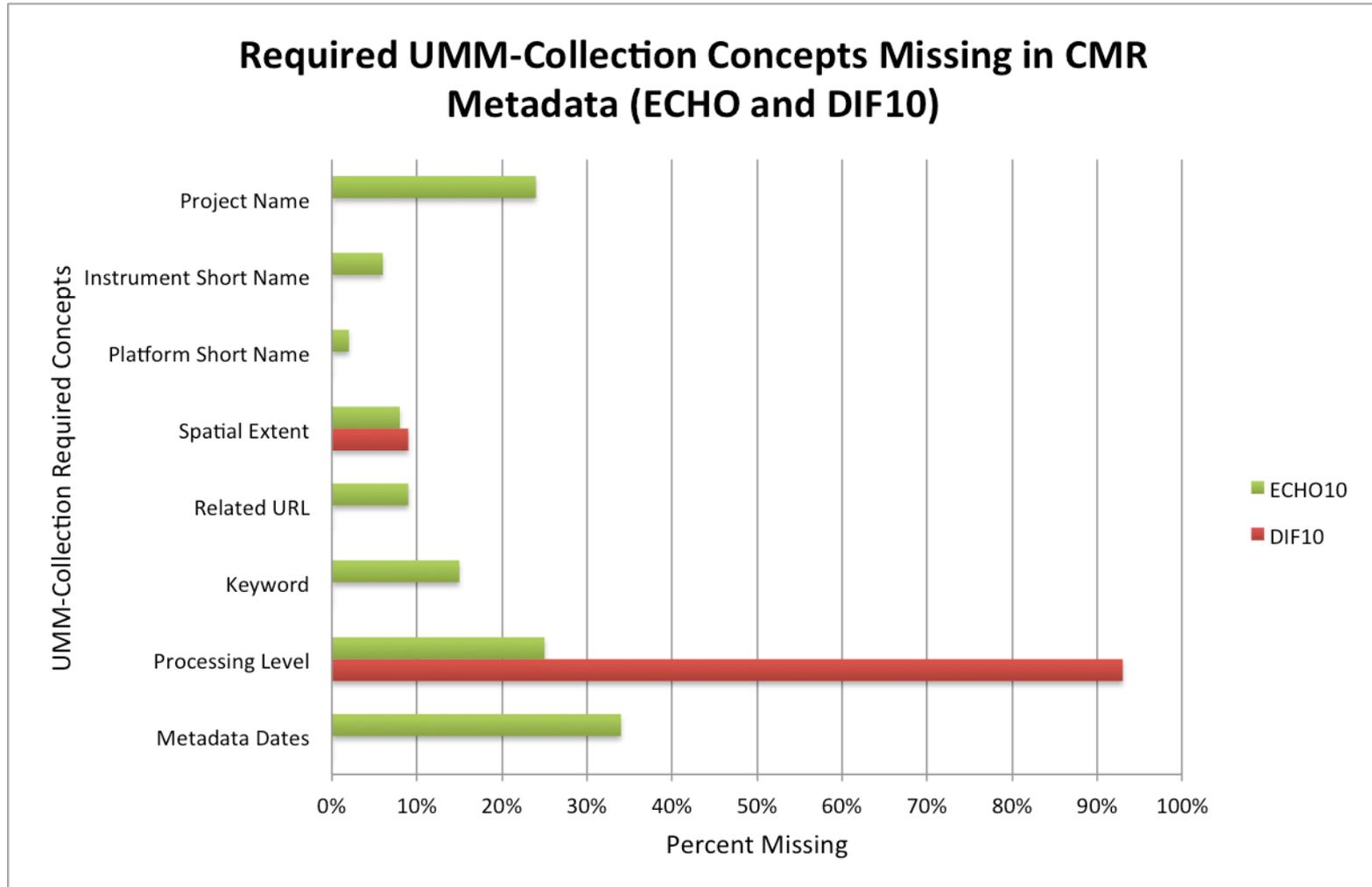
Purpose: Evaluate and measure the completeness of NASA metadata collections with respect to the UMM-Collection profile.

Scope:

- 4,000 ISO, ECHO, DIF10 and DIF records from 17 CMR metadata collections

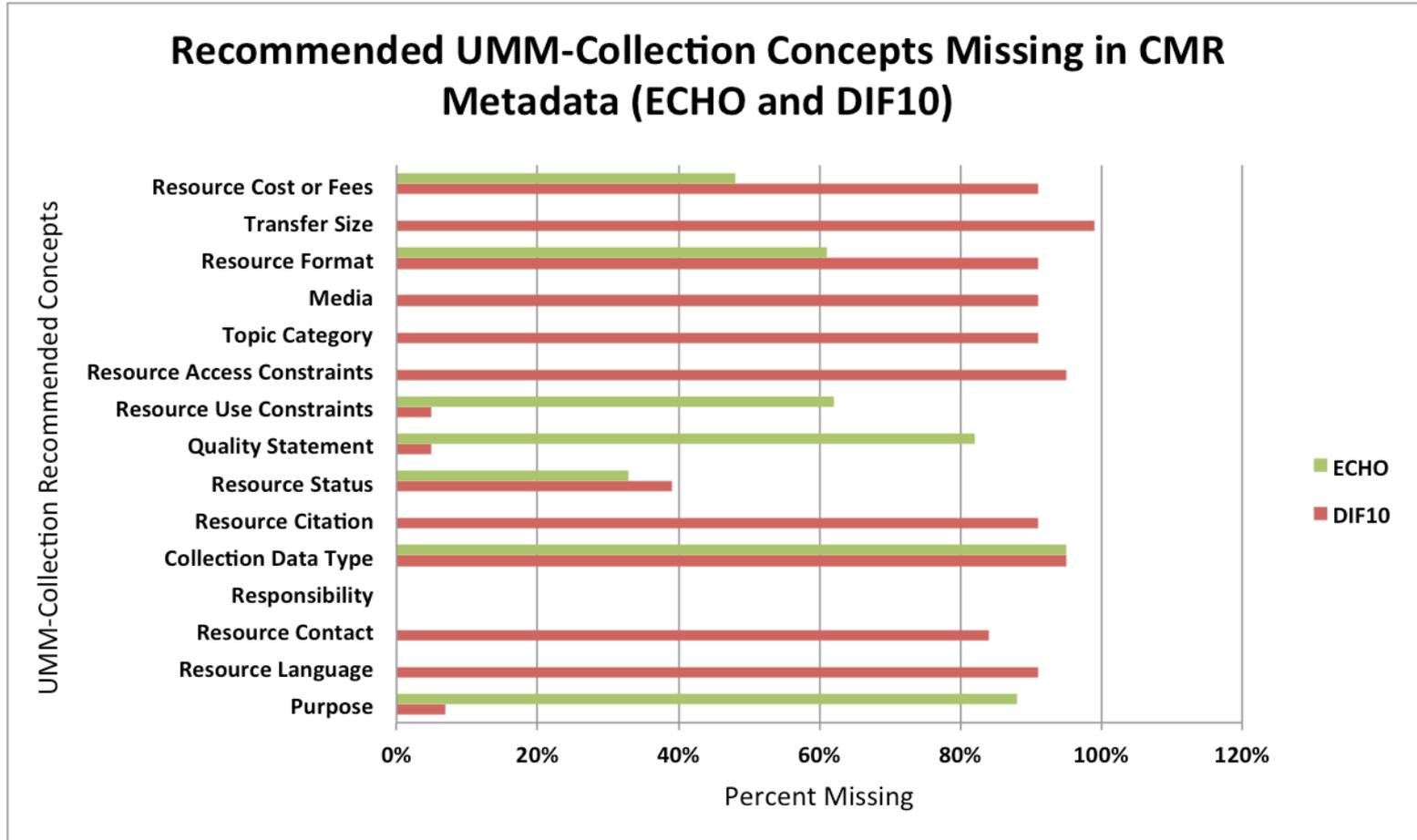
Case Study 2a: Results

UMM-C Required



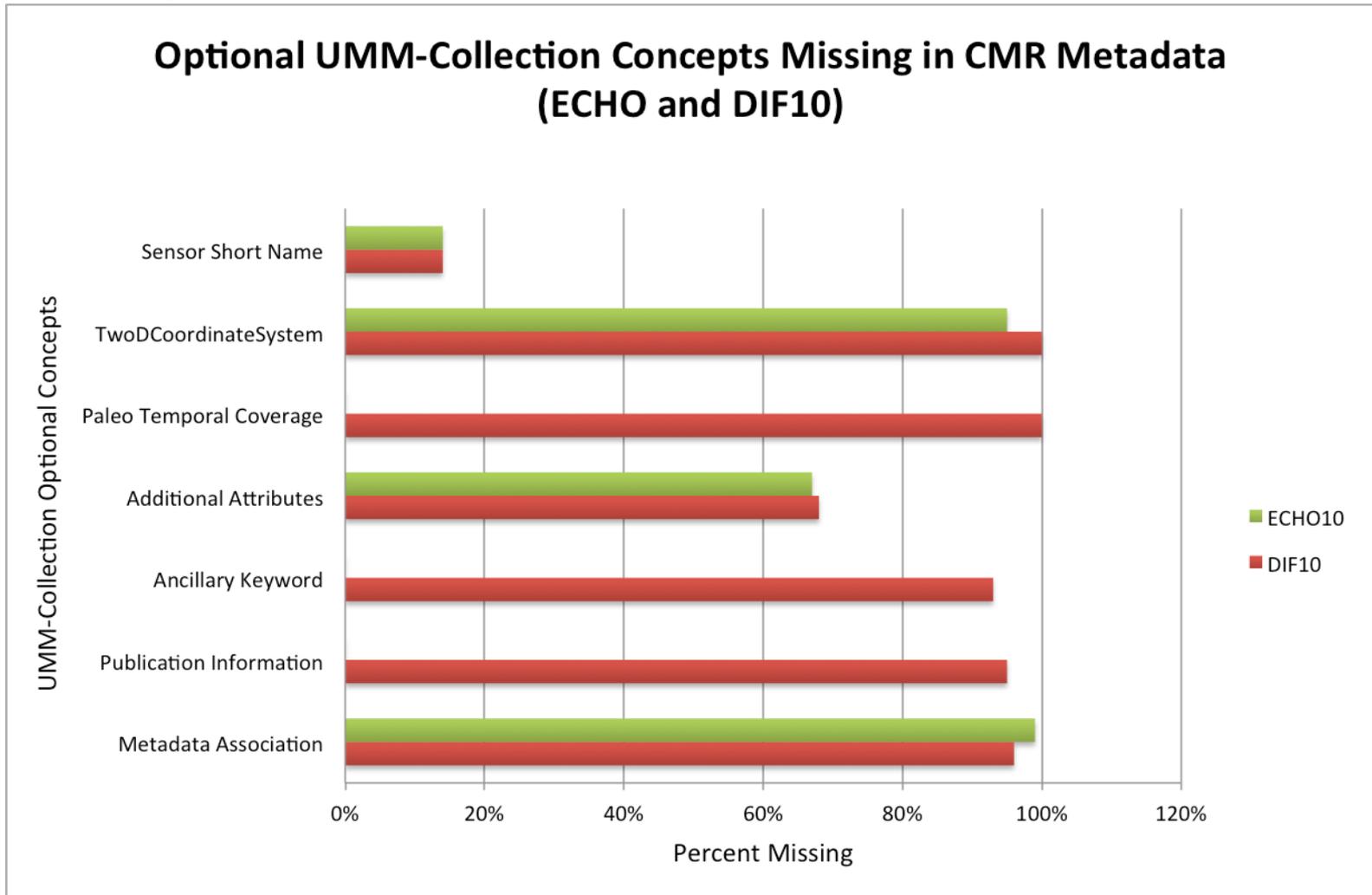
Case Study 2b: Results

UMM-C Recommended



Case Study 2c: Results

UMM-C Optional



Case Study 3:

What metadata records require updating?

Overview: Metadata completeness (signature) scores can be used to identify groups of metadata records that are often missing the same concepts.

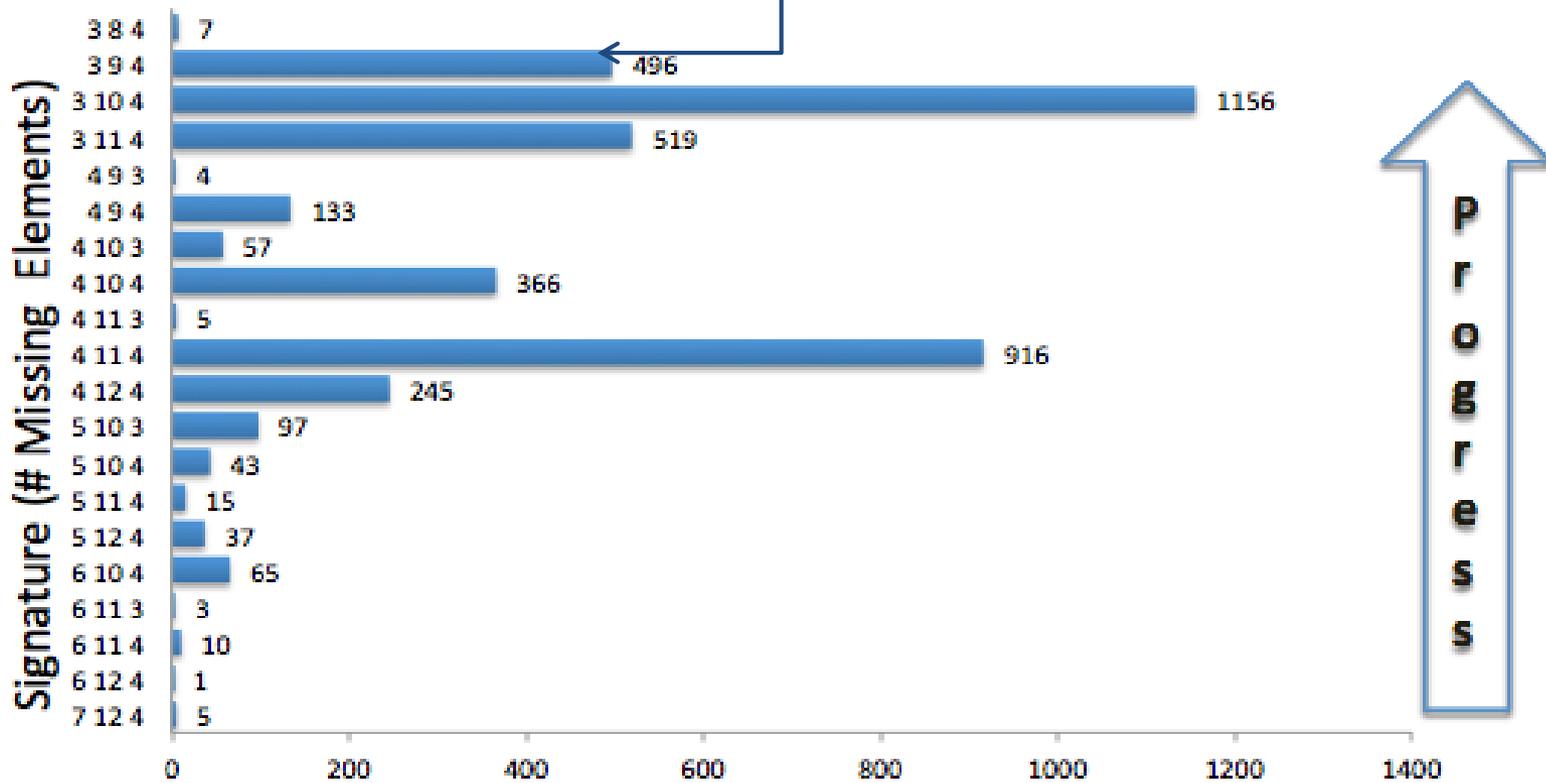
Purpose: Utilize signature scores to identify and prioritize metadata improvement work.

Scope: 16,000 CMR metadata records.

Case Study 3: Results

DIF Signatures Groups

	Resource Id	Resource Version	Keyword	
LAADS	X	X	X	78
LANCEMODIS	X	X	X	76
LARC	X	X	X	186
GSFC4PA	X	X	X	148
GESDISC	X	X	X	7



This work was supported by
NASA/GSFC under Raytheon Co.
contract number NNG10HP02C

Raytheon