DRAS-TIC Fedora Requirements Gathering Interviews – Summary

Introduction

This document is a summary of information collected during requirements gathering interviews conducted with DRAS-TIC Fedora institutional partners as part of the user research component of the DRAS-TIC Fedora project. The partners interviewed were University of Maryland Libraries, Georgetown University Libraries, and the Smithsonian Institution. The interviews were conducted on 2/27/2018 (UMD Libraries), 9/3/2018 (Georgetown University Libraries), and 4/10/2018 (Smithsonian Institution) and a report on each interview was compiled. Some key findings are presented below.

Summary

The three institutions under study presented diverse use cases, but all the use cases had some key common characteristics. Concerns around scalability were expressed by all three institutions with varying degrees of urgency. UMD and SI reported scale as an immediate challenge, while Georgetown cited it as a potential challenge in the future. All three institutions were also interested in systems with open standards; all three currently use DSpace for at least a subset of their collections, and two of the three currently use versions of Fedora.

UMD and SI both described significant challenges with their current systems, while Georgetown's current software successfully meets the library's needs. Although requirements were complex, each institution had some defining characteristics regarding its requirements from digital repository software. For UMD Libraries, the desire to consolidate collections under one versatile system was a main feature. For SI, the ability to compute/perform analysis functions on large and complex objects inside the repository was a high priority. Georgetown cited the stability and reliability of DSpace as an important feature.

The remaining sections present comparative information on the current systems, collection descriptions, and desired digital repository functions of the three partner institutions.

Current Systems

	UMD Libraries	Georgetown Libraries	Smithsonian Insititution
Digital Repository Software	Fedora 2 & 4; DSpace 5	DSpace 5.8	Fedora 3; DSpace; Commercial Enterprise System
Current Challenges	Scalability; fragmentation	No major challenges	Scalability, difficult interoperability, complexity of systems

Table 1: The digital repository systems in use at each institution and challenges encountered while using them.

Scalability challenges:
For UMD Libraries, the most immediate scale challenges are with DSpace as an institutional repository. For SI, storage is required in systems that can manage complex objects.

Fragmentation:
UMD Libraries use three different systems for two use cases. Since all collections present similar challenges, addressing those challenges with one system for both library collections and UMD scholarly production (the two use cases) would be simpler and more efficient.

Difficult Interoperability
In SI, lack if interoperability among the various systems - each system or institution is a silo - used in its 19 units, or museums, means that it is difficult to organize collaborative projects.

Descriptions of Collections

	UMD Libraries	Georgetown Libraries	Smithsonian Institution
Estimated size of collections	200+ TB	~590,000 items	6 – 10 PB
Most common file types	.tiff and .wav files	PDF and JPEG files	Varies depending on which unit
Rate of collection growth	30 TB a year	Roughly 10 collections a year	Not answered

Table 2: Size of collections and most common file formats (where available) in each institution.

Desired Digital Repository Functions

The six digital repository functions in the table below were cited by representatives of two or more of the three partner institutions. All institutions require systems that scale easily and prioritize systems that are open source. All are also interested in linked data and a system that supports AV formats.

Desired Function	Cited By		
Easily Scalable	UMD	GT	SI
Interest in Linked Data	UMD	GT	SI
Open Source	UMD	GT	SI

Supports AV formats	UMD	GT	SI
Flexible metadata requirements	UMD		SI
Community support	UMD		

Table 4: Desired digital repository functions in common among the three institutions.