

# Comparing Staff Time for Maintaining Access to Ephemeral Resources vs Archived Data

The BTAA Geoportal provides **discovery** of geospatial resources via a catalog of aggregated metadata. This metadata includes **access** to the resources via external links. If the source organization removes, replaces, or migrates the resources, our links break, and we fail to provide **access**.

One solution to this problem is to archive the data. This would involve depositing the resources into a university-managed spatial repository and providing stable linkages.

This document compares the staff time required for maintaining access to these resources as external, ephemeral links or by archiving them in a spatial repository. It does not estimate costs for building a spatial repository or additional staff or training that it might require.

## Intangible costs

- **User Attrition**
  - Users may not revisit the site if they click broken access links
  - Users may attempt to save records that they cannot find later
- **Community Reputation**
  - We cannot share majority of records back to OpenGeoMetadata in good faith, because of the maintenance required
  - BTAA records are not seen as complete by some community members (few web services or stable download links)
  - The project may be contributing to the vexing phenomenon of “link rot”

# 1. Maintaining Access to Ephemeral Resources

## Method 1: Batch replacement of collection

*Comparison Re-accessioning (data portals with metadata APIs)*

- Developing & maintaining comparison query scripts (Metadata Coordinator, Spatial Analyst & Curator)
- Running query scripts
  - Monthly re-accession of ArcGIS Open Data Portals (Graduate RA)
  - Quarterly re-accession of Socrata (Graduate RA)
  - Periodic re-accession of CKAN Data Portals and custom portals (Metadata Coordinator)
- Metadata cleanup and enhancement (Graduate RA, Metadata Coordinator)
- Transformations (Metadata Coordinator)
- Deleting orphaned records (Metadata Coordinator)
- Uploading replacement records (Metadata Coordinator)
- Validation and error checking (Metadata Coordinator, Selected Task Force Members)

## Method 2: Manually fixing broken links

*Broken Link Scans (web services, maps, resources listed only on file servers)*

- Development of the URI Analysis script (App Developer)
- Running script monthly (App Developer)
- Creating report (App Developer)
- Analyzing report (Metadata Coordinator, Spatial Analyst & Curator)
- Distributing some investigative work to Task Force (Metadata Coordinator)
- Fixing links manually, usually one by one (Selected Task Force Members)

## Estimate of time per staff member devoted to maintaining 6085 ephemeral records\*

Staff	Batch replacement of collections Estimated Time	Manually Fixing Broken Links Estimated Time	Annual Total Hours	5 year forecast
Metadata Coordinator	8 hours per week	2 hours per week	500	2500
Graduate RA	12 hours per week	-	600	3000
Spatial Analyst & Curator		12 hours per year	12	60
App Developer	-	1 hour per month	12	60
Assigned Task Force Members	10 hours per year	10 hours per year	200	1000
			Total: 1324	Total: 6620

\* Calculations based on reported time

## Annual and 5 year Totals

- ❖ Total number of ephemeral metadata records: 6085 records
- ❖ Total number of staff hours devoted to maintaining: 1324 hours per year (0.64 FTE)
- ❖ Five year forecast without growth: 6620 hours
- ❖ Combined labor required: 12 minutes per record per year

## 2. Archiving Resources in a Spatial Repository

### Labor Requirements

#### *Initial ingest*

- Harvesting resource (Digital Preservation Analyst, Graduate RA)
- Converting to preservation format if necessary (Digital Preservation Analyst, Spatial Analyst & Curator)
- Writing metadata (Digital Preservation Analyst, Graduate RA, Task Force Members)
- Depositing in spatial repository with automatic publishing to geospatial server and geoportal (Digital Preservation Analyst)

#### *Maintenance*

- Format conversion as needed (Digital Preservation Analyst)
- Metadata updates as needed (Digital Preservation Analyst)

### Considerations

#### *Spatial Data Infrastructure*

- Note that this solution would necessitate the construction of an Spatial Data Infrastructure technology stack, whether homegrown or contracted out

### Estimate of anticipated time devoted per item for 6085 archived resources\*

Staff	Initial Ingest (Year 1)	Maintenance (subsequent years)
Digital Preservation Analyst	10 minutes	4 minutes
Graduate RA	15 minutes	
Spatial Analyst/Curator	5 minutes	
Assigned Task Force Members	5 minutes	1 minute
	35 minutes/record	5 minutes/record

\*Calculations based on anticipated time spent per item.

### Annual and 5 Year Totals

- ❖ Total number of archived resources: 6085
- ❖ Total number of staff hours devoted to archiving (Year 1): 3550
- ❖ Total number of staff hours devoted to maintenance (subsequent years) = 487
- ❖ Combined number of staff hours devoted to maintenance (Years 2-5):  $487 \times 4 = 1948$
- ❖ Combined labor over five year forecast without growth: 5498 (0.52 FTE)