



NATIONAL
ARCHIVES

NATIONAL HISTORICAL
PUBLICATIONS
& RECORDS COMMISSION



COLORADO
State Archives



SGER Grant to Preserve Colorado's Born Digital Legislative History

Project Plan

First Choice for Service

This documentation was made possible through the financial support of the National Historical Publications & Records Commission.

Table of Contents

Introduction	3
Purpose of Plan	3
Overview	3
Project Approach	4
1. Project purpose.....	4
2. Goals and objectives	4
2.1 Performance Objectives for NHPRC Project RG-50010-16	4
2.2 Project objectives.....	4
3. Scope	5
3.1 Inventory	5
3.2 Migration.....	5
3.2.1 Migration tracking procedures	5
3.3 Audio Edits	6
3.4 Embed Metadata	6
3.4 Collection Policy	6
4. Roles and responsibilities.....	6
5. Assumptions and constraints.....	6
5.1 Project Assumptions	6
5.2 Project Constraints.....	6
6. Terms of Reference	6
7. Migration workflow	7
8. Naming Convention	7
9. Descriptive and technical metadata guidelines	8
9.1 Re:Discovery Collection Hierarchy	9
10. Project timeline.....	10

Introduction

Purpose of Plan

The SGER Grant to Convert Freedom System (GCFS) project plan will provide a definition of the project, including the project's goals, objectives, terms of reference, and migration guidelines. In addition, the plan will supplement the GCFS User Guide as reference documentation throughout the project lifecycle.

The Project Plan defines the following:

1. Project purpose
2. Goals and objectives
3. Scope and expectations
4. Roles and responsibilities
5. Assumptions and constraints
6. Terms of reference
7. Migration workflow
8. Migration file naming convention
 - o Version control
9. Migration technical and descriptive metadata guidelines
10. Project timeline

Overview

In 1973, the Colorado General Assembly began recording the meetings of the Senate and House in accordance with the State's newly enacted "Sunshine Law" (C.R.S. 24-6-401). The individual committees are recorded (such as House Education Committee, Senate State Affairs Committee), as well as the Chambers belonging to each legislative body. These recordings are vitally important as they are the only records that can prove legislative intent. They are often requested by Supreme Court Justices, Legislators, the District Attorney's Office, private lawyers, and historians.

From 1973-1997, analog audio tapes were used. The years between 1997-2002 were recorded on digital audio tapes (DATS), in DDS (Digital Data Storage) format. From 2002-2011, software was used called the Freedom System (FS), which recorded the audio in digital format. The software also allowed for searching and playback as well.

The software was originally designed for use with Windows XP and is quickly becoming out of date. The General Assembly's IT has developed a patch for Windows 7 - and State Archives was awarded a State Electronic Records Grant (SGER) to transfer these files to archival format and to provide online access through some sort of interface.

The goal of this project is to transfer these nine years of recordings to archival format, edit the files so they are cohesive, make user access copies, catalog, and publish online. We are also expected to work with the Records Management team, General Assembly’s IT, and the Colorado Channel to streamline the process for future recordings (which are now digital video).

Project Approach

There will be four phases in the migration of the Freedom System audio records.

High level milestones

Phase 1	Migrate Freedom System to WAV format
Phase 2	Catalog converted files into RE:Discovery
Phase 3	Upload master WAV files to OCLC “Dark Archive” - create access MP3s
Phase 4	Upload MP3s to CONTENTdm (access digital repository)

1. Project purpose

The Colorado State Archives (CSA) seeks a solution to provide resources regarding the conversion and preservation of nine years of born-digital audio records created by the Colorado General Assembly. The project will help foster cross-agency partnerships, and assist with the creation of a robust state collections development policy including updates to agency retention schedules that address permanent born-digital audiovisual collections. Policies, procedures, and work flows developed as part of this project will lead to the creation of a comprehensive statewide model to ingest future large born-digital content collections from all state departments, municipalities, and stakeholders involved with statewide programs/collections.

2. Goals and objectives

2.1 Performance Objectives for NHPRC Project RG-50010-16

1. Submit all reports by the dates specified in the Award Reporting Requirements of the Notice of Action.
2. Acknowledge the National Historical Publications and Records Commission in all print and electronic products that result from grant support. The Commission strongly encourages grantees to use the NHPRC logo in grant products and publications and to link to its website. The logo is available at: <http://www.archives.gov/nhprc/about/logo.html>. The final narrative report must include link to grant products.
3. Amount of cost share meets or exceeds the amount approved by the NHPRC.
4. Convert approximately 150,730 audio files covering ten years of legislative activity.
5. Provide access to the converted digital files through the Colorado State Archives website.
6. Complete a records schedule for the electronic records of the Colorado General Assembly.
7. Publish policy, procedure, and work flow documentation in CoSA’s PERTTS portal.
8. Use existing State, as well as professional and community, communication venues to publicize the project. Promote the project via regular social media posts, including the Archives’ Facebook and Twitter accounts.
9. Submit proposals to present on the project to at least two professional conferences and publish articles about the project in at least two professional publications.

2.2 Project objectives

1. Preservation of 14TB of permanent records of the Colorado General Assembly.

2. Collaborative development of policies that relate to permanent born-digital collections.
3. Nine years of legislative history will be made available to researchers.
4. Increased collaboration and sharing of information between branches of state government.

3. Scope

The general work plan for this project will be to ingest digital files, migrate them to archival format, and embed metadata before creating access copies and uploading to the web. Simultaneously, collections policies, standards, and retention schedules will be developed for records created by the Colorado General Assembly.

3.1 Inventory

A complete inventory of all FS audio files will be completed as part of the migration. The inventory will serve as a tracking documentation of the FS audio file migration. The inventory will contain quality assessment checks which are to be completed once the migration is complete.

Inventory and migration procedures: S:\SGER_GRANT\SGER_documents\Working\FSMigrateTrackguide.working.docx

In addition to the manual inventory, we are investigating a means to extract all metadata from the FSM system.

3.2 Migration

Process Freedom System audio files and convert to an uncompressed WAV format with the standard PCM codec.

Specifications:

WAV - PCM (standard WAV)

- File extension: .wav
- Internet media type: audio/wav
- Sample rate: 44.100 khz
- Bit rate: 16 bit

Migration of FS files will include all files, regardless of content.

3.2.1 Migration tracking procedures

The purpose of the migration tracking procedures will be to assign FS files to Scott or Eamon for migration and to identify the stage of each migration within the migration workflow.

Information to capture:

-
- **Migration date - YYYY/MM/DD**
- **Bit Rate Check - Initials of individual conducting assessment YYYY/MM/DD**
- **Metadata Check - Initials of individual conducting assessment YYYY/MM/DD**
- **Audio Check - Initials of individual conducting assessment YYYY/MM/DD**
- **Metadata captured - YYYY/MM/DD**
- **Cataloged in RE:Discovery - YYYY/MM/DD**
- **Access MP3 created - YYYY/MM/DD**
- **MP3 added to RE:Discovery - YYYY/MM/DD**
- **Checksums added to Re:Discovery record - YYYY/MM/DD**

- Quality control check - YYYY/MM/DD

3.3 Audio Edits

Due to FS settings/parameters a number of calls were split into multiple WAV files. The SGER project team investigated editing and merging these split files. It was determined that the split files names still reflected accurate recording time date and time stamps. Due to this, merging and editing files is not required for search purposes and has been determined to be out of scope of this project.

3.4 Embed Metadata

Create technical and descriptive metadata as part of the migration process. Embed the metadata in the migrated WAV file.

The procedures we used to embed the metadata into the created Master WAV files is documented in the BWF MetaEdit Guide:

"S:\SGER_GRANT\SGER documents\Working\BWF_MetaEdit_Guide.pdf"

3.4 Collection Policy

This project is in accordance and compliance with the Colorado State Archives Collection Policy: P:\Admin\Policies & Procedures\FINALcollectionspolicy_2016.pdf

4. Roles and responsibilities

Project team	Role	Responsibilities
Aly Jabrocki	Image and Media Archivist (Team Lead)	Project Management, Quality control
Scott Lawan	Digital Media Archivist	Migration, processing, metadata capture
Eamon Smallwood	Digital Media Archivist	Migration, processing, metadata capture, documentation
Kevin Smith	Project Archivist I	Quality assessment

5. Assumptions and constraints

5.1 Project Assumptions

- Add FS files will be converted to WAV PCM
- Project team members will work in conjunction to complete project
- FSM will be functional and accessible during the project
- The S: server will be functional and accessible during the project

5.2 Project Constraints

- Limited functionality of FS, i.e. inability to easily bulk export metadata from system.
- Limited technical support from State of Colorado IT department.

6. Terms of Reference

Term	Definition
Freedom System	The Freedom System is a proprietary digital audio recording system purchased by the

General Assembly for the recording of legislative sessions. The System uses compressed audio files with codec G.723.1 (which is unsupported open source). In 2016 a SGER grant was awarded to convert these files to uncompressed WAV format with the standard PCM codec.

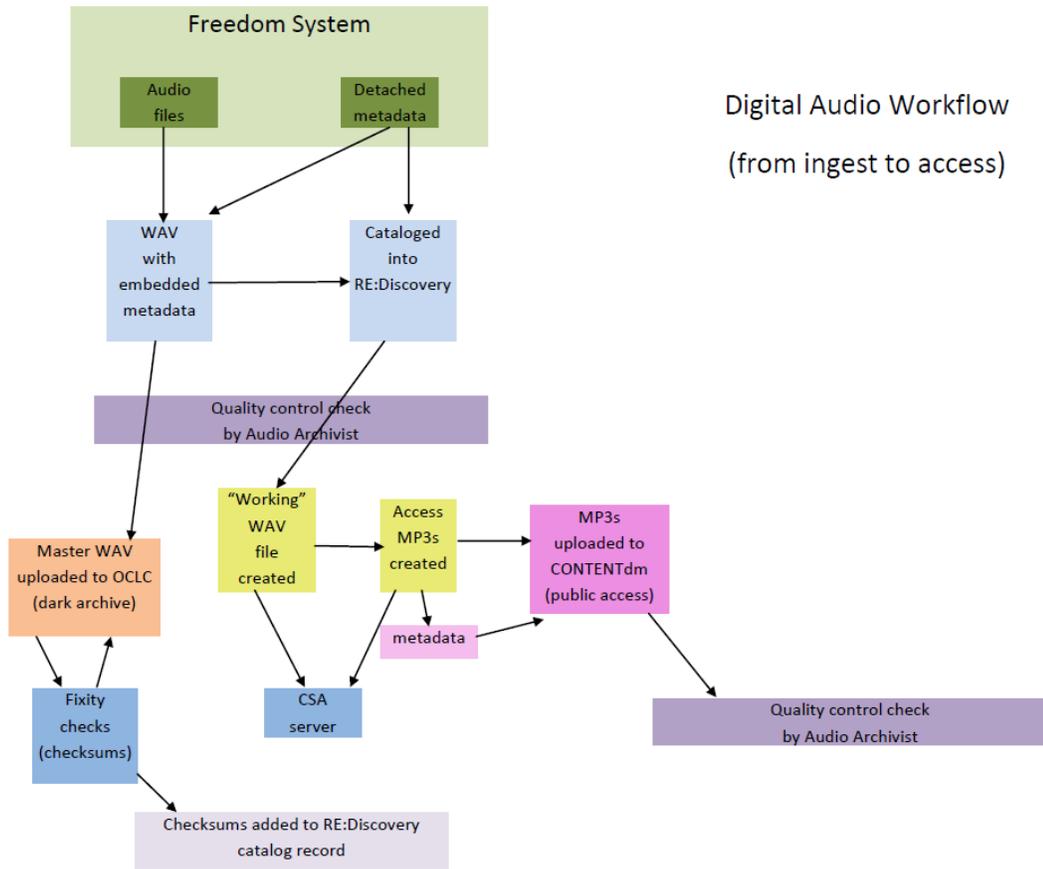
State Electronic Records Grant (SGER)

Approved grant to transfer Freedom System audio files to an archival format and enable online access

Waveform Audio File Format (WAV)

Archival uncompressed audio format

7. Migration workflow



Digital Audio Workflow
(from ingest to access)

8. Naming Convention

The naming convention will use the original file names generated by FS. These file names are effective as they are unique to each file and contain channel number and timestamp descriptive information.

File names are to remain as-is to minimize edits to existing filenames.

Information to include in file names¹:

1. Channel Name/Abbreviation - physical location of recording, expressed as “cXXX or cXXXX”
2. Date - YYYYMMDD format
3. Timestamp or range - 24 hour format, e.g. 0800_1210

Current: SCR353_April12_800am_1210pm_2006_HouseChambers_February2_952am_200pm_2010

Final: c202_20060412_0800_1210.wav

9. Descriptive and technical metadata guidelines

Descriptive and technical metadata will be captured for each migrated file in the following formats.

Format/Schema	Type
BWF	Technical
ID3	Descriptive
RIFF	Technical
DC	Descriptive
MODS	Descriptive

BWF:

Description: This recording has been transferred into a WAV format with PCM codec from the Freedom System. Original format was WAV with G.723.1 codec.

Originator: Colorado General Assembly

Originating Date:

Date of recording in *yyyy-mm-dd* format

ID3:

Artist: Colorado General Assembly

Copyright: State of Colorado for public use. Credit
Colorado State Archives.

RIFF:

Original Medium: WAV codec G.723.1, with Freedom System
software.

Technician: *Your Name*

Archival Location: Colorado State Archives

DC:

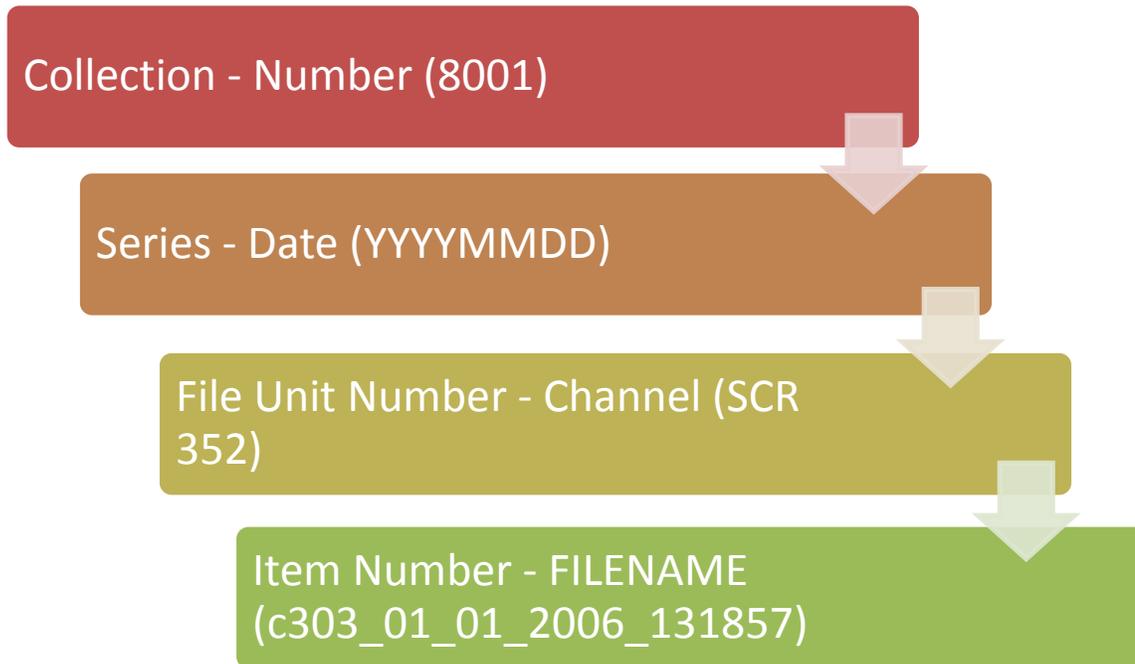
<S:\SGER GRANT\SGER documents\Working\SGER sample DC.xml>

MODS:

¹ <https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming>

9.1 Re:Discovery Collection Hierarchy

Migrated records will be cataloged into the Re:Discovery database using the proposed collection hierarchy.



File Unit Information tab - Main Information tab

Field Name	Information
Collection NBR	8001
Series Nbr	YYYYMMDD
File Unit Nbr	Channel
Title	Channel, YYYY-MM-DD Timestamp (HH:MM:SSAM/PM)
Category	Nonmusical sound recording
Location	
Dates	YYYY-MM-DD
Extent	HH:MM:SS
Creator	Colorado General Assembly
Creator Role	Creator
Summary Note	Channel 7, HCR0107, StartTime: 2006-01-01 01:18:57 PM, EndTime: 2006-01-01 01:19:00 PM, Duration: 0:00:03
Media Type	audio/x-wav
Language	English (eng)
Phys Char	N/A
Notes/Prov	This recording has been transferred into a WAV format with PCM codec from the

Freedom System. Original format was WAV with G.723.1 codec.

Proc By	Eamon Smalwood
Topic Term	Legislative materials--lcsh
Geogr Name	Denver (Colo.)--lcsh
Genre/Form	Nonmusical sound recording
Web Ready?	Y
User 3 (digital origin)	audio/x-wav, reformatted digital
User 4 (rights statement)	This record is intended for public use and is designated as Copyright Undetermined, http://rightsstatements.org/vocab/UND/1.0/ . Please credit "Courtesy of Colorado State Archives." For more information please contact Colorado State Archives.

10. Project timeline

Pre-Grant (FY-16)

Interview and hire two Archivists (A and B), set up work stations and install all software.

July - December, 2016 (FY 17)

Archivists A and B begin employment, Audio Archivist trains new hires on file conversion/migration, metadata, cataloging, sound editing, and CONTENTdm. Archivist A and B start conversion/migration of audio files, embedding metadata and editing files once trained. In October, cataloging of files into RE:Discovery starts, along with first round of master WAV files being uploaded to OCLC. The Audio Archivist will begin quality control checks weekly, beginning in October 2016.

January - June, 2017 (FY 17)

Archivists A and B continue with all tasks from the previous period. Audio Archivist continues with weekly quality control checks and ensures all work is performed to approved standards. If number of files processed does not meet goal trajectory, the Audio Archivist will work with the Legislative Council Library and the University of Denver to recruit interns for metadata embedding and cataloging.

July - December, 2017 (FY 18)

Archivists A and B continue with all tasks, with assistance from interns if necessary. Migration of Freedom System files into WAV format with metadata will be complete by December, 2017. Along with ongoing quality control checks, the Audio Archivist will meet with the General Assembly's Legislative Council and Office of Information Technology to plan for ongoing data growth for ongoing audio-visual file ingests and transfers. The Audio Archivist will also collaborate with other State Archives staff and the General Assembly to draft retention schedules regarding audio-visual files.

January - June, 2018 (FY 18)

Archivists A and B will complete RE:Discovery cataloging of audio files in March, 2018. Upload and transfer of master WAV files to OCLC will be completed in June, 2018 as well;

along with access copy MP3s uploaded to CONTENTdm. The new retention schedule and ingest policies will go into effect in March, 2018. The Audio Archivist will begin compiling final figures and weighing them against performance objectives in January, having all figures and objectives met by the end of the grant cycle on June 30, 2018.



COLORADO
Division of Central Services
Department of Personnel
& Administration