

MATRIX

The Center for Humane Arts, Letters, and Social Sciences Online

Mission

- Networking and Coordination
- Publication
- Teaching
- Research and Study
- Research and Development



Creating a **National Gallery** of the **Spoken Word** Online

The National Gallery of the Spoken Word is funded by Michigan State University, it's partnering institutions, and the National Digital Library Initiative Phase II spearheaded by the National Science Foundation.



Bring Historical Voices Online

I. Prioritization

II. Digitization Issues

III. Metadata

IV. Storage

V. Management/Preservation

VI. Access and Permissions

VII. Delivery

I. Prioritization



- Audience
- Rights and permissions
- Accessibility
- Ease of digitization
- Preservation

II. Digitization of speech recordings for archival purposes

- 1. Prepare the originals**
- 2. Hardware-based A/D Conversion**
- 3. Capture digital signal on a PCI sound card**
- 4. Record input in a sound editor**
- 5. Process the recorded signal**
- 6. Save the file**
- 7. Store the file**

Prepare the originals

- **Identify the media**
- **Choose hardware for playback**
- **Connect the playback device to an A/D converter**
- **Use correct settings and connections for mono and stereo recordings**
- **Avoid using filters such as Dolby NR**
- **Adjust output levels to avoid signal overload**
- **Monitor the output**

Hardware-based A/D Conversion

Choose the appropriate A/D converter

It is recommended to use a professional-level, hardware platform. Such hardware should include 16-bit input/output, auto calibration, sophisticated signal conditioning and anti-aliasing filters to take full advantage of the 16-bit signal representation.

Additionally, the hardware should be able to offer a wide assortment of frequency ranges, XLR (balanced) inputs, DAT pass-through, and powerful digital signal processing circuitry for fast signal analysis.

Unlike systems built around generic plug-in multimedia sound cards, such hardware usually offers input signal-to-noise performance typically 20-30dB superior to generic plug-in sound cards.

Capture digital signal on a PCI sound card

The analog signal is converted to digital by the external module. Then it is captured by an internal PCI sound board. ***It is important to avoid using standard multimedia sound cards.*** Professional-level sound cards that meet the following specifications are better:

- PCI Interface
- 8 to 24 bit resolution
- Variable sample rates, including 11.025kHz and 44.1kHz
- Analog two channel in/out via 1/4" TRS connectors
- S/PDIF digital in/out via gold-tipped RCA connectors
- +4/-10dB balanced/unbalanced operation
- 4 channel operation using both analog and digital ins and outs simultaneously
- Analog Input Dithering

Record input in a sound editor

A sophisticated sound editor which meets the following specifications is recommended:

- **Multiple Document Interface for editing multiple files in one session**
- **Unlimited file size editing**
- **Real-time graphs (amplitude, spectrum, bar, and spectrogram) (optional)**
- **Real-time fast forward and rewind playback**
- **Numerous effects (distortion, Doppler, echo, filter, mechanize, offset, pan, volume shaping, invert, resample, transpose, noise reduction, time warp, pitch, and more)**
- **Supports many file formats (WAV, AU, IFF, VOC, SND, MAT, AIFF, MP3, OGG, and raw data) and can convert to/from these formats**
- **Drag-and-drop cue points**
- **Direct sample editing with the mouse**

Process the recorded signal

For strictly archival purposes, it is recommended not to use any processing beyond the following A/D conversion settings:

- Sample rate: 44.1 kHz
- Resolution: 16 bit
- A/D converter settings: use plain signal with no correction. Use normalization if necessary.
- Input level: adjust and monitor levels to avoid signal overload. Digitization is very prone to clipping on overload. The process of setting levels is different for digital and analog.
- Save as: *.wav

Save the data

The WAV file format is recommended for storing digital versions of speech recordings. WAV files are uncompressed, thus preserving all bits of information recorded in the AD process. It is also widely used and easy to process and convert to a variety of streaming formats.

The WAV file format is a variant of the RIFF format for data interchange between programs. This format was designed so that data in a file is broken up into self-described, independent "chunks".

A RIFF file contains a single chunk of type "RIFF", with other chunks nested inside it. Therefore, the first four bytes of a WAV file are "RIFF", and the four bytes after that contain the size of the whole file minus 8 bytes.

After the RIFF header is the WAV data, consisting of the string "WAVE" and two important chunks: the format header and the audio data itself. There may also be other chunks in a WAV file that contain text comments, copyrights, etc., but they are not needed to play the recorded sound. The header and the audio data **can be easily separated** to facilitate migration to any other format, if necessary.

Data Storage

- To facilitate long-term preservation, store digitized WAV files as data, not as audio files.
- The files can be stored on a hard drive with FTP access and a regular back-up plan
- An additional copy of original *.wav files should be stored on CD ROM or DVD as a preservation copy

III. MetaData

Metadata is machine understandable information, usually for the web. Files can be digitized, collected and stored, but what remains missing is information about the data - labeling, cataloging and descriptive information structured in such a way that allows web pages to be properly searched and processed by a computer. Several options currently exist for encoding metadata, among them the Dublin Core initiative, W3C's metadata initiative, XML, and several other emerging technologies. At their most base level, each initiative has the same objective: to describe data in a way which facilitates organization, searching, and retrieval in a logical manner.

Extensible Markup Language (**XML**) has emerged as a viable markup language to combine the searching functions of a database with the mark-up functions of solutions like HTML.

XML tags describe the data in a document; stylesheets then tell browsers how to display each tag.

Encoded Archival Description (**EAD**) is a Document Type Definition (**DTD**) of XML for encoding archival finding aids using the Standard Generalized Markup Language (**SGML**).

Transcriber

Transcriber 1.3b7

File Edit Signal Segmentation Options Help

report - Fireside chat: On National Security. Sunday, December 29, 1940

Franklin Delano Roosevelt, President

MY FRIENDS: This is not a fireside chat on war.

It is a talk on national security,

because the nub of the whole purpose of your President is to keep you now, and your children later, and your grandchildren much later, out of a last-ditch war for the preservation of American independence

and all of the things that American independence means to you and to me and to ours.

Tonight, in the presence of a world crisis, my mind goes back eight years to a night in the midst of a domestic crisis.

It was a time when the wheels of American industry were grinding to a full stop, when the whole banking system of our country had ceased to function.

[b] [NOISE]

I well remember that while I sat in my study in the White House, preparing to talk with the people of the United States, I had before my eyes the picture of all those

fourfreedoms

Resolution

Fireside chat: On National Security. Sunday, December 29, 1940

Franklin Delano Roosevelt, President

MY FRIENDS: This is not a fireside chat on war. It is a talk on... security, because the nub of the whole purpose of your President is to keep you n... preservation of American independence

0 5 10 15 20 25 30

Cursor : 9.66 sec Selection : 9.66 sec - 13.7 sec (4.03 sec)

Transcript - imported or typed in to text window

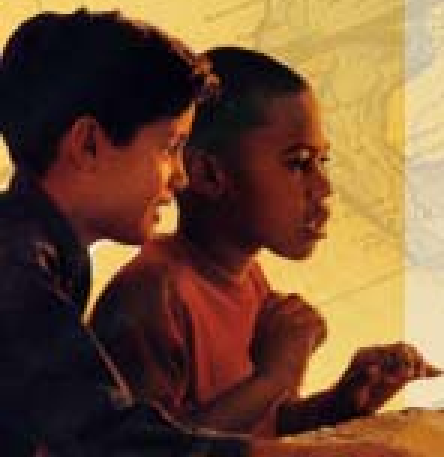
Playback controls

Visual representation of audio material

Linked layers: segmentation (pink), speaker (grey), transcript (blue), timecode (white)

IV. Storage

- **Sound**
- **Metadata**
- **Images and Text**
- **How to associate files**



V. Management / Preservation

- **Managing the System**
- **Backups**
 - metadata, not audio
 - archival storage of audio on CD
- **Migration**



VI. Access and Permissions

Specifying permissions

- **Flexibility in Choices**
- **Locating permission data in metadata**



Access and Permissions

Usage restrictions on certain files and collections should be addressed as copyright clearances and other guidelines are negotiated with rights holders.

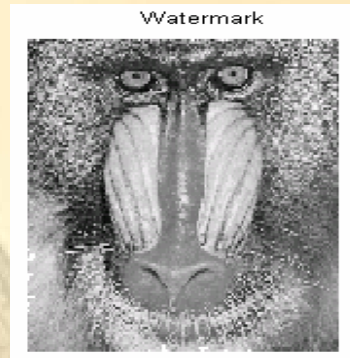
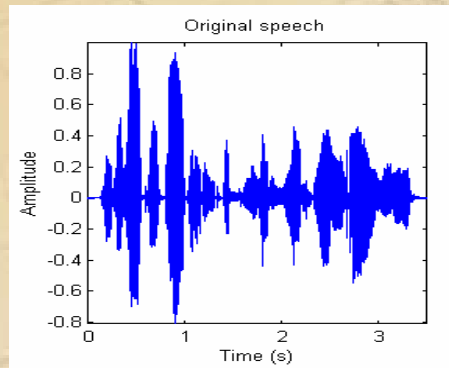
Restrictions may be implemented on servers:

- **Directory level security**
- **IP address restrictions**
- **Broadcasting restrictions (time limited access)**

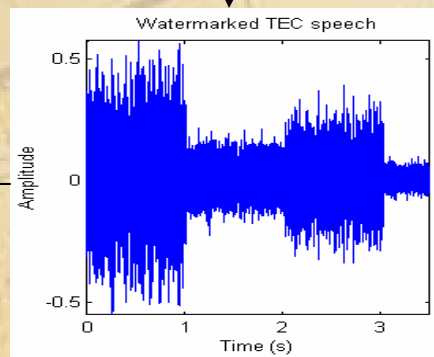
Specific solutions to copyright restrictions will have to be negotiated on a per archive or per producer basis.

Access and Permissions

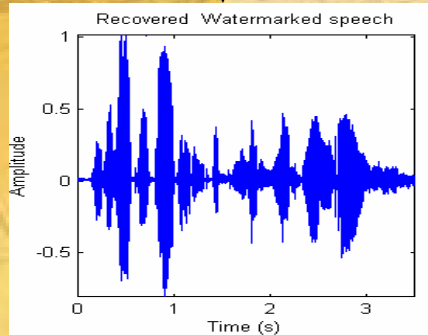
Encryption/Copyright Protections; Novel Watermarking Techniques



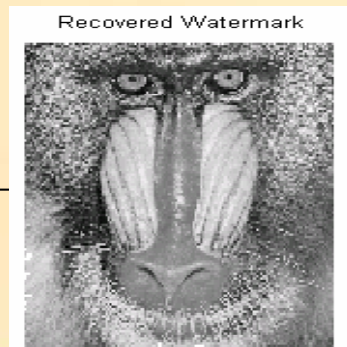
**Watermark and encrypt speech
for transmission**



**Decrypt watermarked
speech for playback**



**Recover watermark
if necessary**



VII. Delivery

Online Streaming

- RealAudio
- QuickTime
- MPEG files

Downloads

CD distribution of uncompressed media

