Electronic Record Preservation

Topics

8. Wrap Up

7. Methods

1. Challenges

6. Goals

2. Fragility
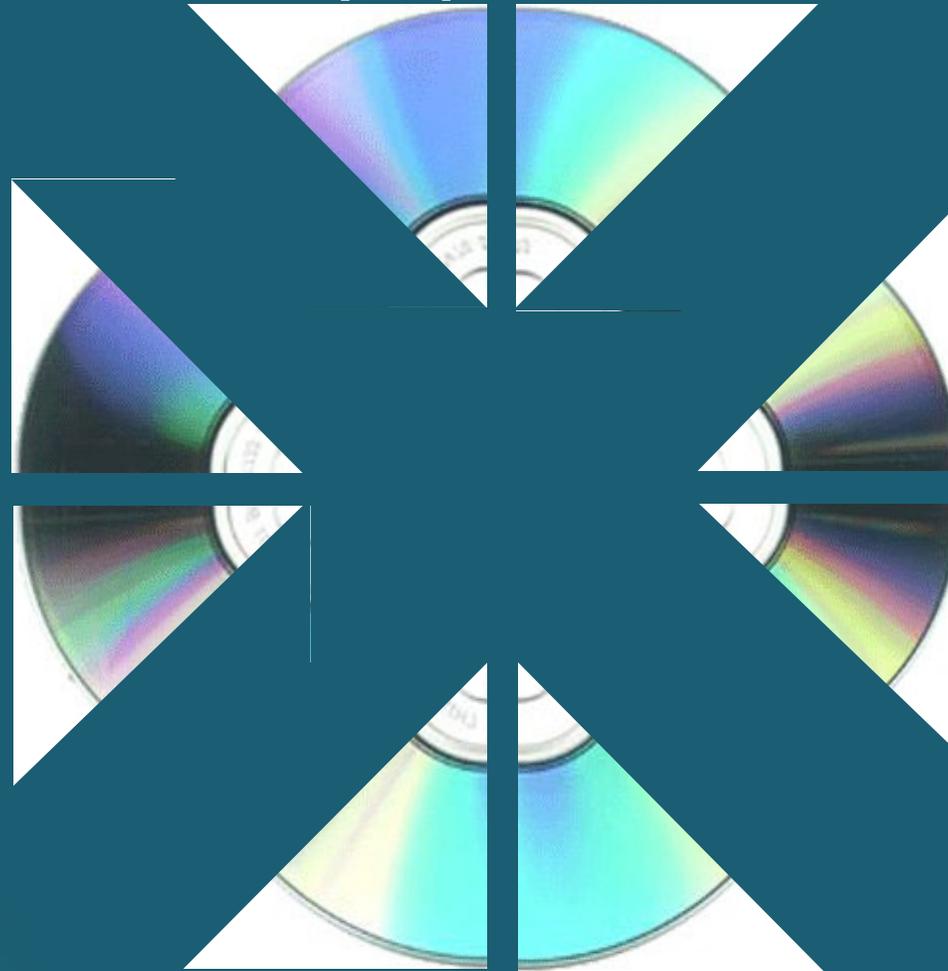
5. File Formats

3. Why Important

4. Storage

Preservation Challenges

- Technological Dependency
  - Records rely on specific hardware, software, etc. to be read

- Technical Obsolescence
  - Little incentive to support older technology over time
  - Hardware, software, platforms become obsolete by design

- Storage Media
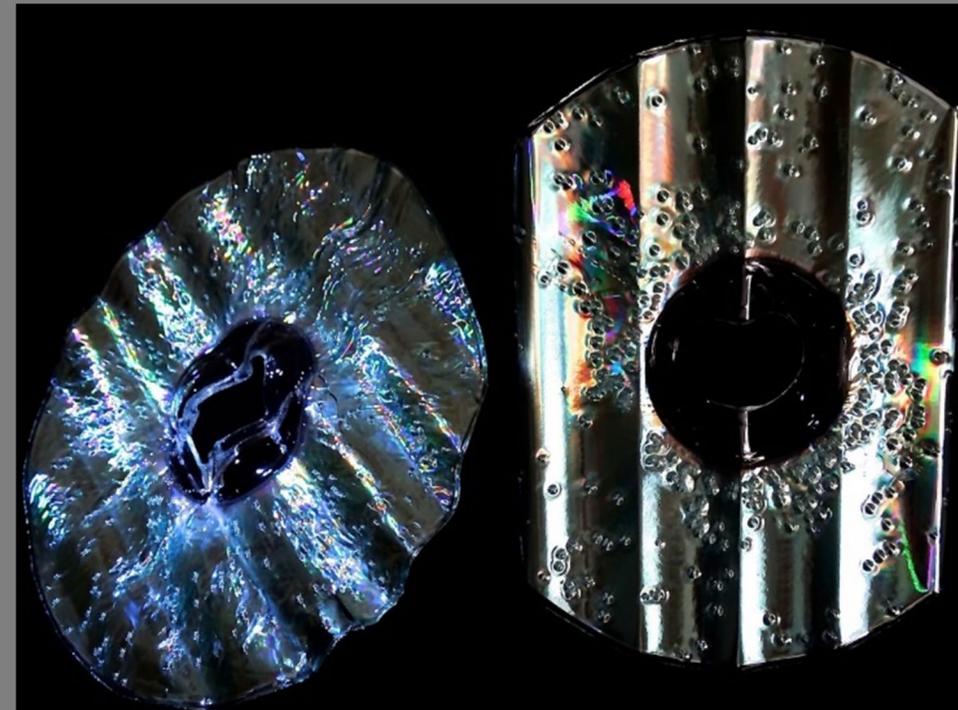 - All media deteriorates over time
  - Theft and loss

Fragility of digital information

Disk rot – gradual decay
of CDs and DVDs

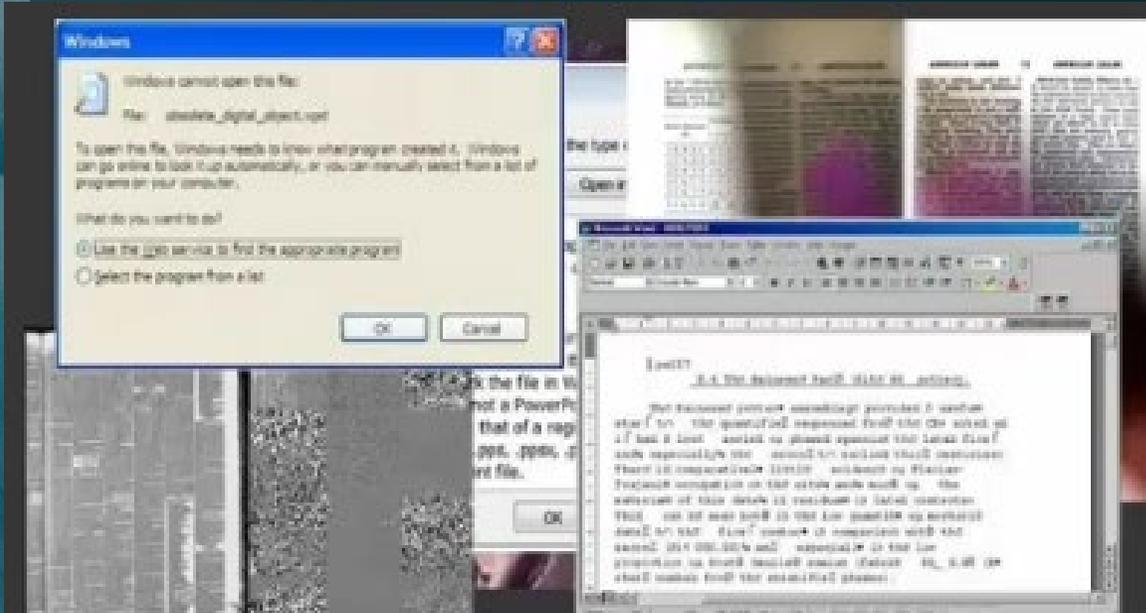...continued

Hard disk drives
(internal or external) 3-5 years
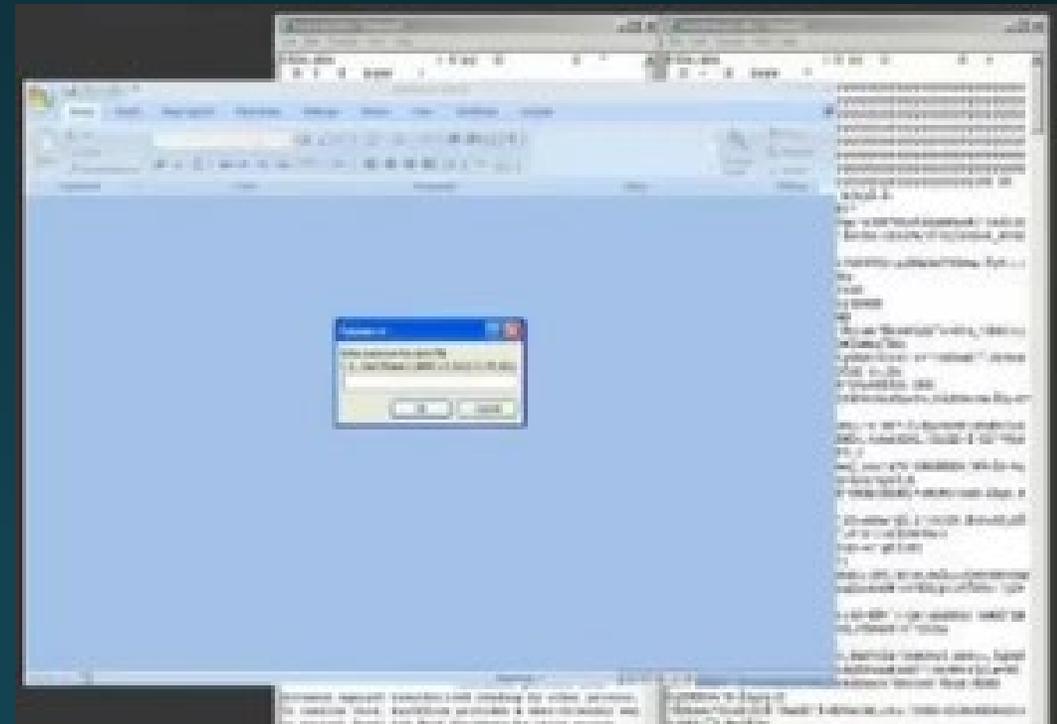
Flash media 10 years

…and

Unintelligibility of data

Encrypted, expired, or password-protected

...But, it's not just a technology issue

- Policies
- Staff
- Training
- Governance
- Sustainable funding

State Records Management Laws

Why is preservation important?

State Agency legal responsibility
Gov't Code Sec. 441.186(e)

(e) If the commission cannot accept custody of an archival state record, the record shall remain in the custody of the state agency and shall be preserved in accordance with this subchapter, rules adopted under this subchapter, and other terms on which the director and librarian and the agency head may agree.

Best storage option?
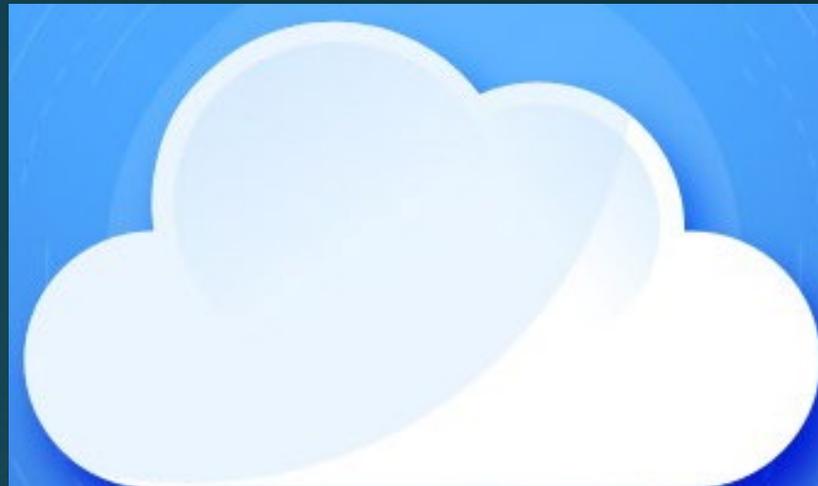
► A network server that gets backed up regularly
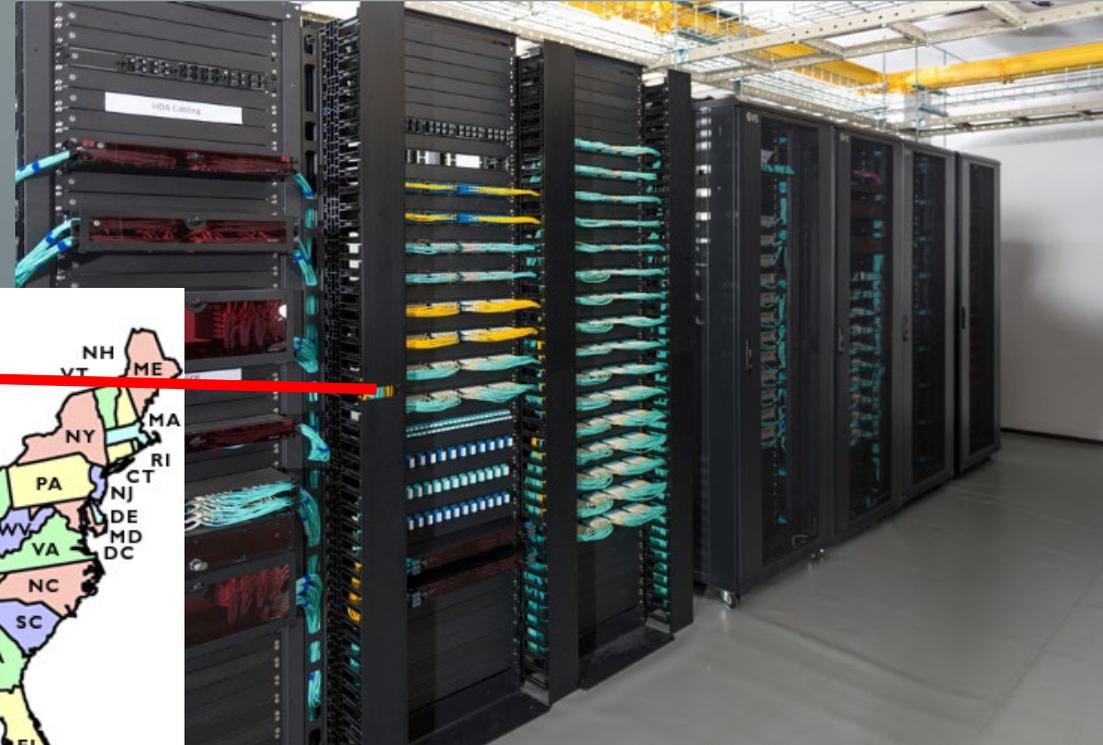
► More than one location

Storage Media

LOCKSS and the "3 copy rule"

► Data redundancy – 2-3 backup copies

► Combinations of cloud backup, external USB hard drive, flash drive

File Formats Matter

- ☞ The file format you choose will affect your long-term records management abilities

- ☞ Most require conversion

- ☞ Choose Preservation File Formats

- ☞ The most effective digital preservation techniques are pre-emptive

5? hard drive (CAV) MiniDisc (CLV) 8 floppy disk (CAV)

Iomega Zip (CAV)

audio Compact Disc (CLV*) Digital Versatile Disc (CLV*)

Blu-ray Disc (CLV)

12 vinyl record (CAV) LaserDisc (CLV/CAA/CAV)

File Formats: Text

☛ Recommended:

♦ PDF "Archival" (PDF/A) (.pdf)

♦ OpenDocument text (.odt)

♦ Extensible Markup Language (.xml)
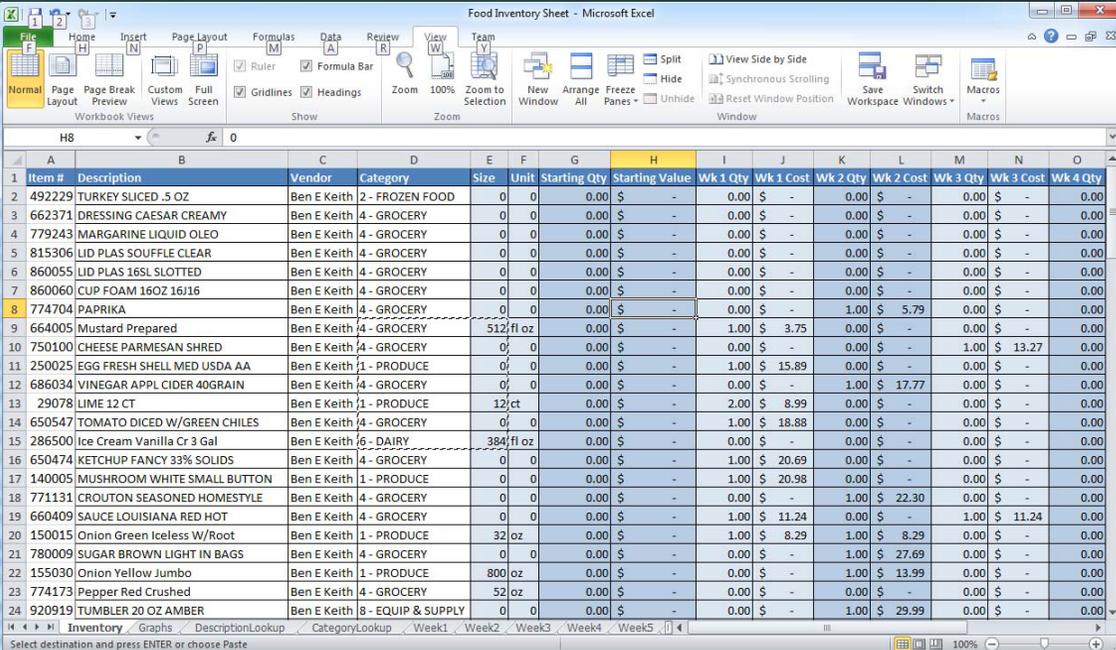
♦ Rich Text Format (.rtf)

♦ Text file (.txt)

ODT

PDF/A

File Formats: Spreadsheets

☛ Recommended:

♦ OpenDocument spreadsheet (.ods)

♦ Comma-separated Values (.csv or .txt)

File Formats: Images

☞ Recommended:

♦ JPEG 2000 (.jp2)

♦ Portable Network Graphics (.png)

PNG
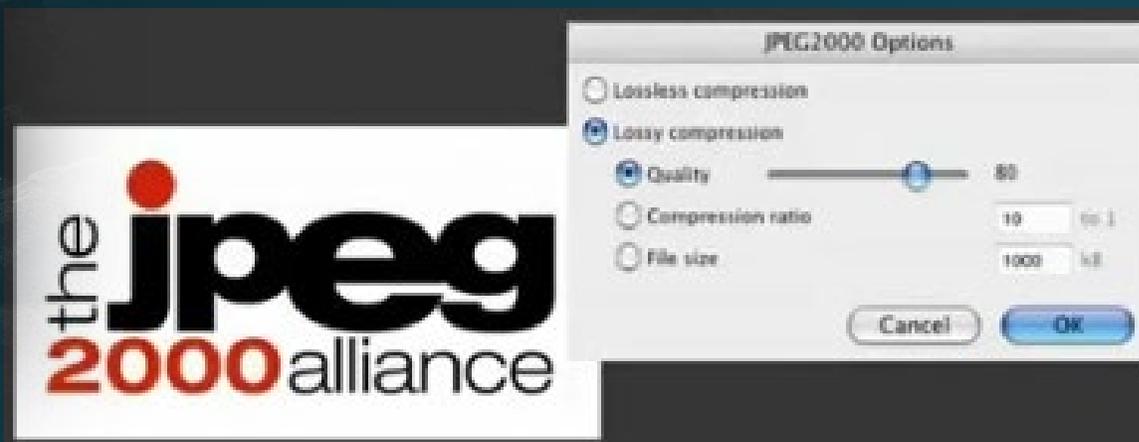
File Formats: Sound Recordings
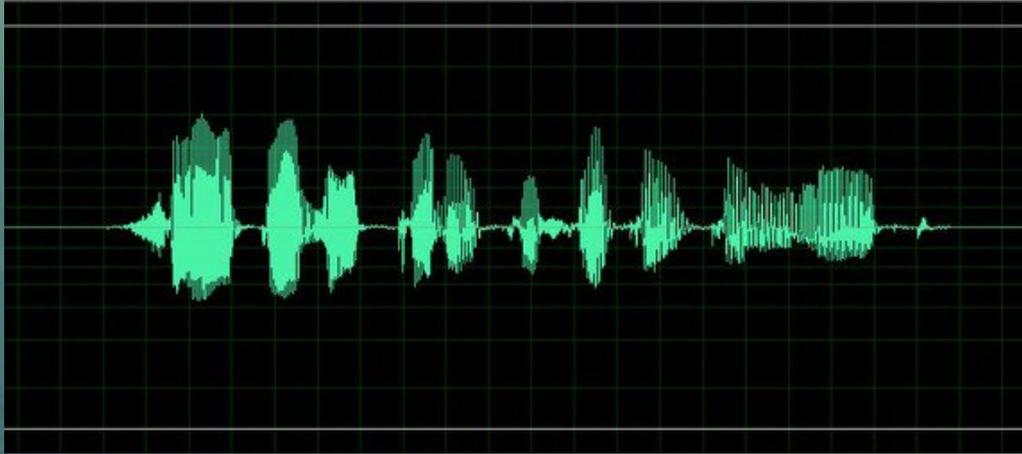
👉 Recommended:

♦ Wave/Broadcast Wave (BWF) (.wav)

♦ Free Lossless Audio Codec (.flac)

File Formats: Video Recordings

☛ Recommended:

♦ Matroska Media Container (.mkv)

☞ Recommended:

♦ MBOX (.mbox)

♦ …unless it doesn't need to be stored in the email account: TXT, PDF, print…

Preservation Goals

Availability

Authenticity

Functionality

Usability

Availability

- The record is kept for
  its full retention period

- The record can be located

More likely to

...be preserved

Less likely to be preserved
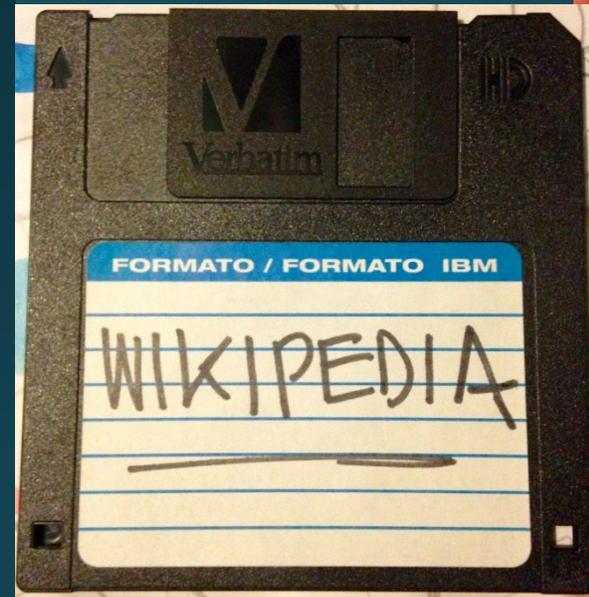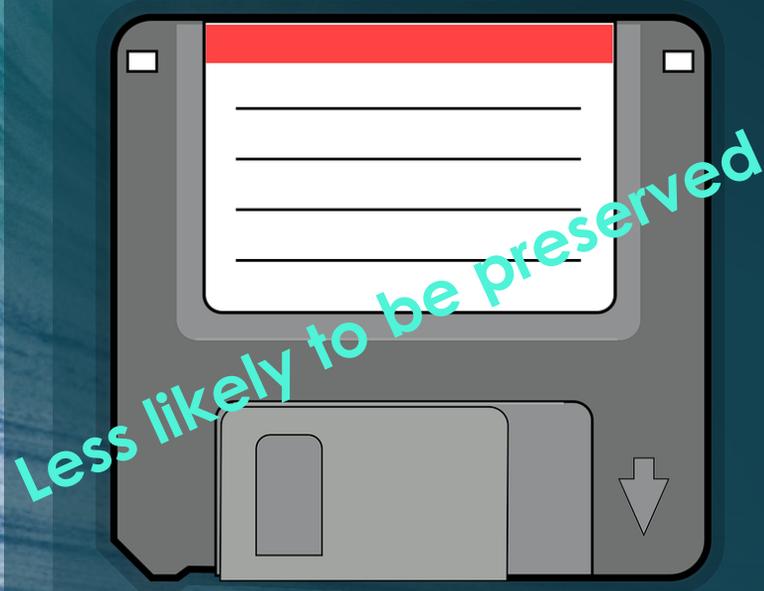
Usability

- ♦ The record can be accessed

- ♦ The record can be read

- ♦ The record is complete

Authenticity

The record is what it purports to be

Electronic Records

Standards and Procedures

(b) An electronic state record that is an archival record must be maintained by the agency through hardware and software migration and upgrades as authentic evidence of the state's business in accessible and searchable form, except as otherwise determined by the state archivist.

TEXAS STATE LIBRARY AND ARCHIVES COMMISSION

Functionality

Methods

1. Computer Museum

2. Emulation

3. Recopying

4. Media Migration

5. Normalization

6. Microfilm

7. Print to Paper

Computer Museum

Preserve the computing environment in which the records were created

- Linux Web Server
- Mac Plus
- Nimrod XT
- Microbee
- Mac LCII
- Ubuntu Box
- Mac G3
- Commodore 64
- Networked Scanner

Hardware

Software

Operating systems

Disk drives

Benefits

- Source document untouched: lower risk of data loss

- No up-front costs if simply retaining old hardware/software

- For obsolete, proprietary, homegrown software: might be best (only) option

Challenges

- Technical expertise needed

- Potentially costly to service old equipment / pay programmer

Emulation

Keep data in its original format

Microsoft Windows 3.1 running in DOSBox, on Windows XP
Image: Wikipedia

Run obsolete programs on existing platforms through special emulation programs
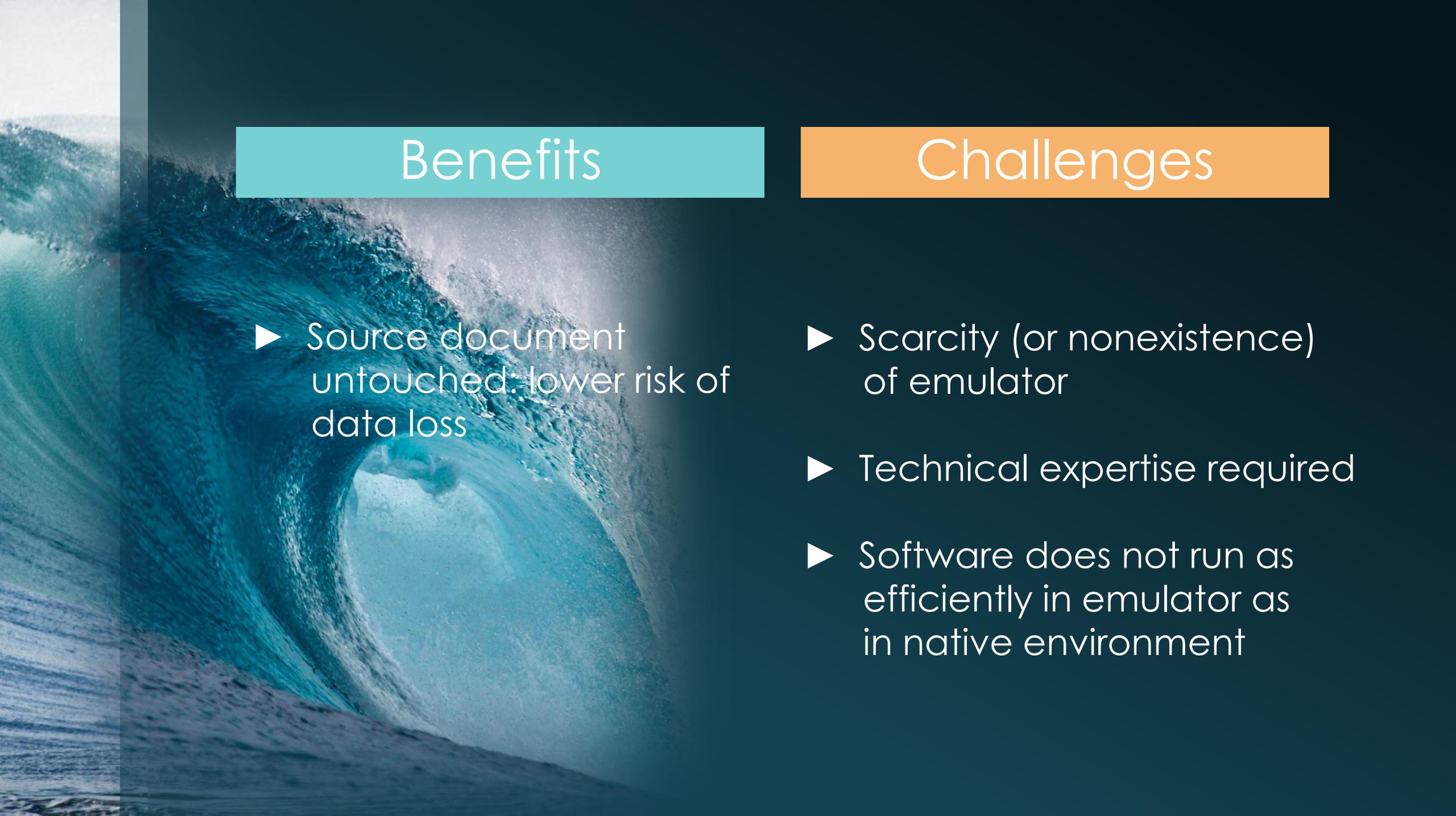
Benefits

► Source document untouched: lower risk of data loss

Challenges

► Scarcity (or nonexistence) of emulator

► Technical expertise required

► Software does not run as efficiently in emulator as in native environment

Recopying

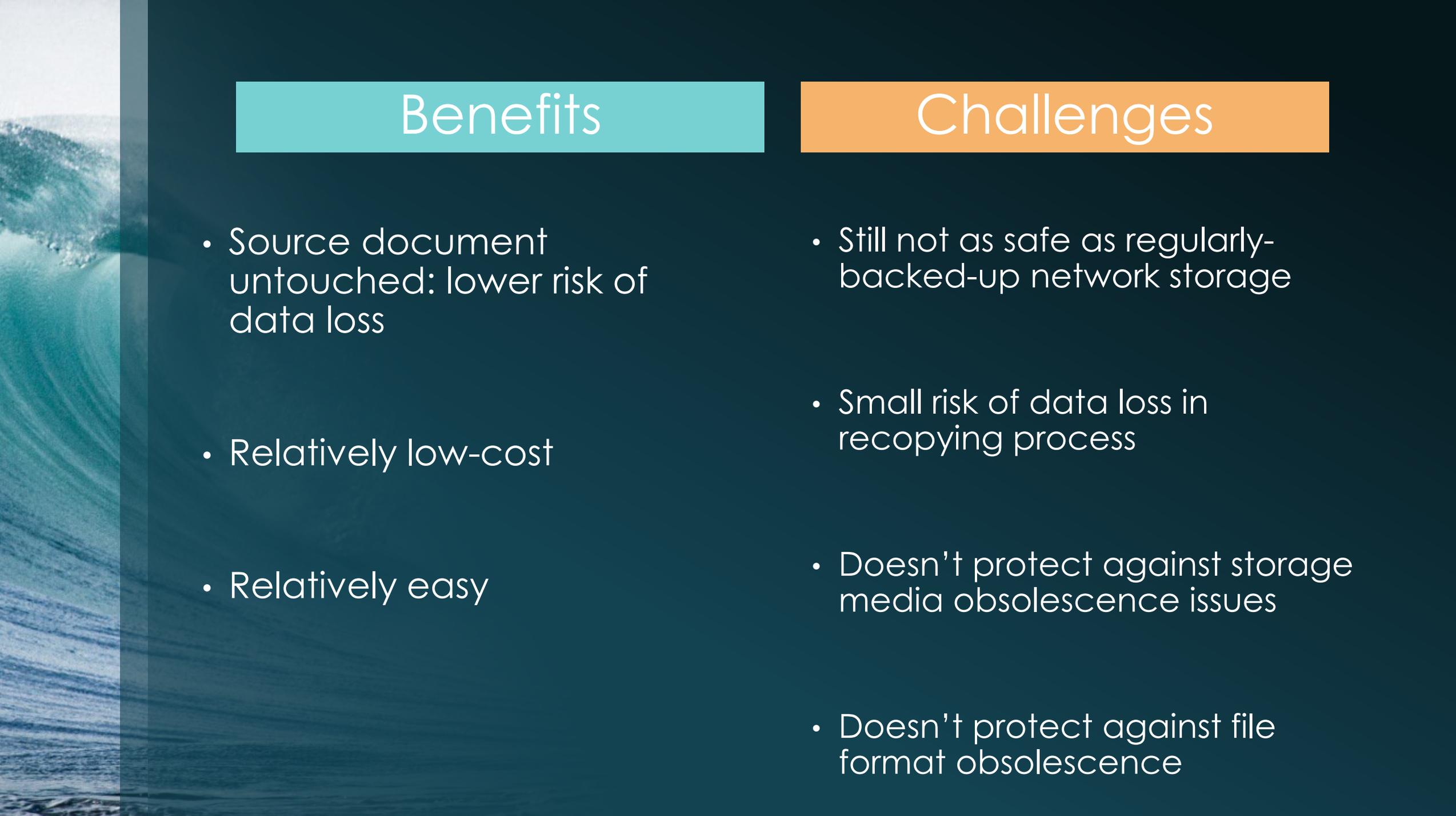Copy the data to new media periodically

Same medium, from (e.g.) an old CD to a new CD

Also called "Refreshing"

Benefits

- Source document untouched: lower risk of data loss

- Relatively low-cost

- Relatively easy

Challenges

- Still not as safe as regularly-backed-up network storage

- Small risk of data loss in recopying process

- Doesn't protect against storage media obsolescence issues

- Doesn't protect against file format obsolescence

Media Migration

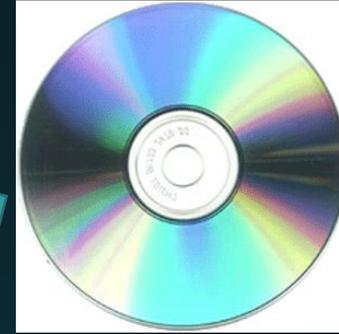Move data off of obsolete storage media

Benefits

- Source document untouched: lower risk of data loss

- Relatively low-cost
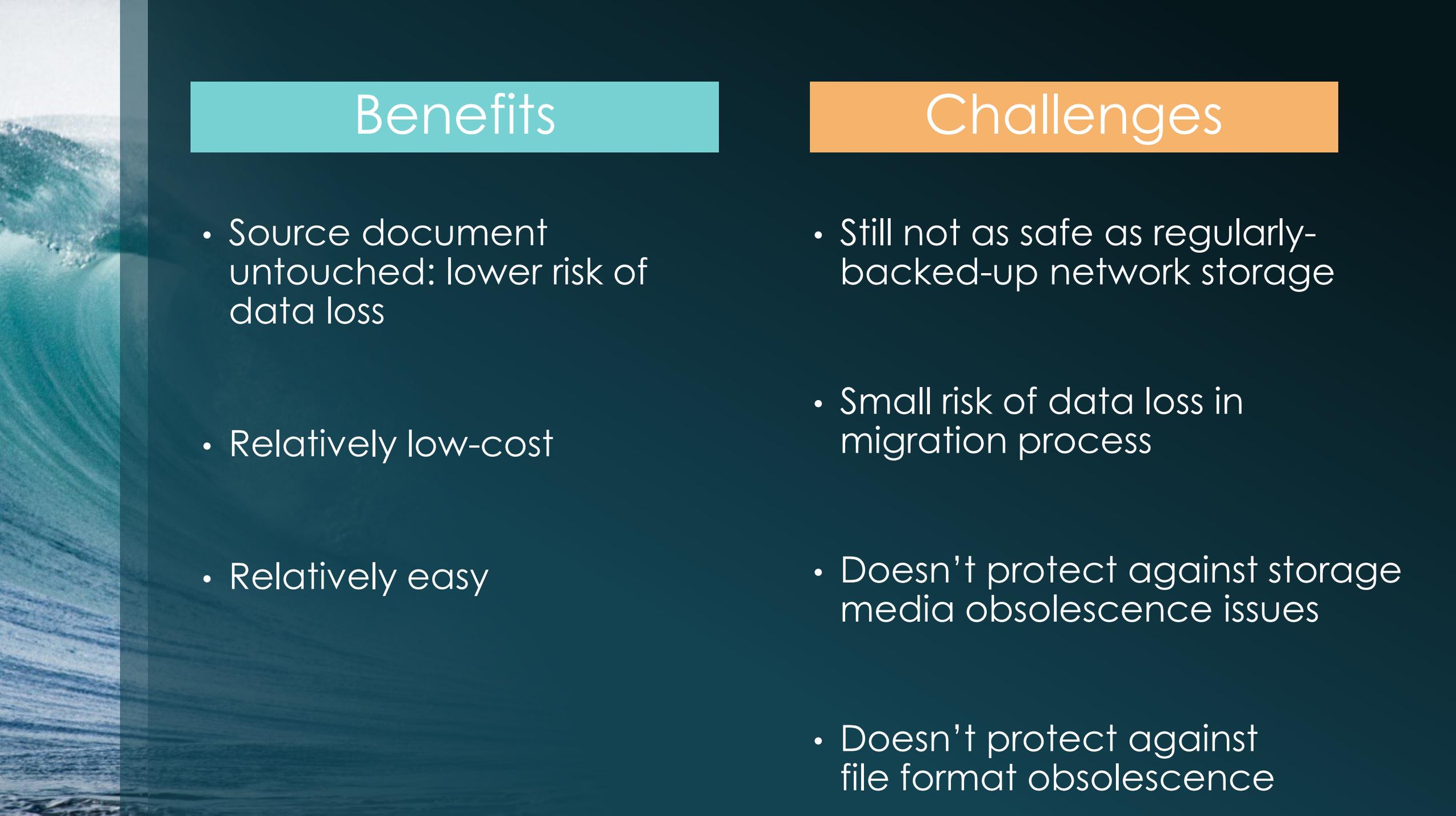
- Relatively easy

Challenges

- Still not as safe as regularly-backed-up network storage

- Small risk of data loss in migration process

- Doesn't protect against storage media obsolescence issues

- Doesn't protect against file format obsolescence

Normalization
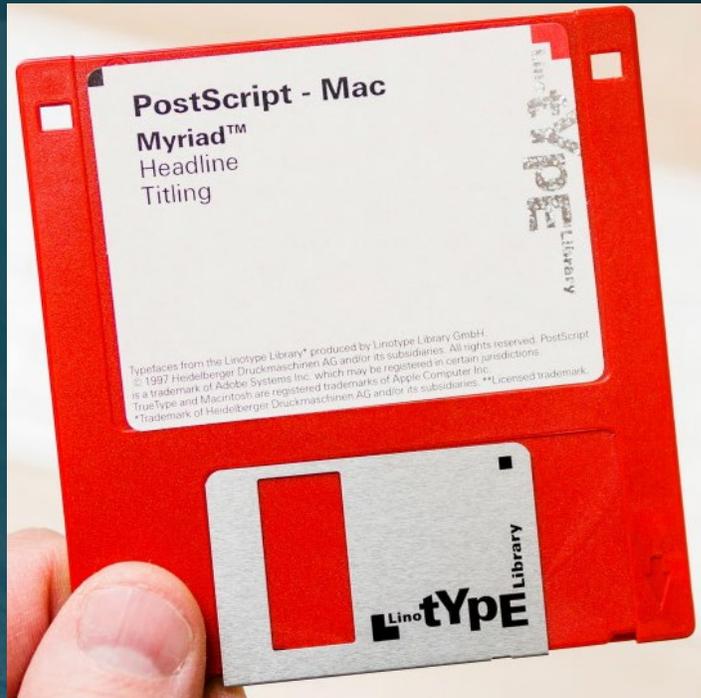
PDF/A: PDF Archival (.pdf)

OpenDocument Text File (.odt)

Plain Text File (.txt)

Convert data into Preservation (File) Formats

Microfilming

Create a microfilm copy

Microfilming

Standards and Procedures

2

Benefits

- Life expectancy of 500 years in proper storage conditions

- Creates unalterable, authoritative copy

- Protects against technology change

- Can be scanned back into digital format

- Space-saver

Challenges

- Slow retrieval time: not-too-frequently-used documents

- Can degrade if not stored properly

Printing to paper

Alkaline paper recommended

Benefits

- Life expectancy of 1,000+ years in proper storage conditions

- Protects against technology change

- Printing rarely requires technological expertise

- Can be scanned back into digital format

Challenges

- Takes up physical space

- Not as searchable as electronic documents

- Potential loss of metadata

- Potential loss of functionality

WRAP-UP

Electronic records are fragile

File formats are important

Bulletins 1, 2, 4

Preservation Goals

Thank you.

Please contact AR-EHS (*ar-ehs@wtamu.edu*),
or *806-651-2270* for more information.