MANAGING RESEARCH ASSETS

PRESERVING YOUR DIGITAL ASSETS

The Library of Congress offers guidelines for preserving digital material, including photographs, audio, video, email, digital records, and websites. See http://www.digitalpreservation.gov/you/

For a more technical and specific discussion of digital formats, see the Library of Congress's "Sustainability of Digital Formats": http://www.digitalpreservation.gov/formats/

In general, you should:

- **Identify**: Make an audit of what you have.
- **Decide** which of your assets you want to keep and which you don't need.
- **Organize** your assets: Give them descriptive filenames, organize them into a logical file structure, and write down your organizational scheme.
- **Make copies**. It's a good idea to have copies in a number of locations. Every few years, check your copies to see if you need to export them to a newer format.

DEVELOPING A DIGITAL RESEARCH WORKFLOW

There's no "right" research workflow. The practice that makes sense for you will depend on your own research habits and the kinds of material you work with. As you investigate tools, think about:

- **Capturing**. Do you do most of your research online, in an archive, or at the library? You'll need a tool (or tools) that's appropriate for the way you really work and easily captures the data you need in a format that's preservable — and preferably in a way that's organized.
- **Metadata**. Few things are more frustrating than locating just the information you need but not being able to determine its origin. That's why it's important to think about how you're capturing information *about* each asset you gather, like its source and its importance to your research.
- **Searching and retrieving.** None of this does you any good if you can't get your hands on the data you need when you need it. Metadata will help you find the right stuff, but you may also want to think about tools for OCR (optical character recognition) and for "fuzzy" searching.

You should also be thinking about whether and how you can **export your data**. That may seem boring now, but it won't when the tool you're using becomes obsolete!

TOOLS TO CONSIDER

Backup Tools
Yes, you need to be doing this. Why are you not doing this? Do it right now! Think about whether you want to store your backup on a hard drive or in the cloud — or both!

- Time Machine (Mac, already installed on your computer, automatically backs up your data to a hard drive at scheduled intervals)
- Windows Backup and Restore (Windows, already installed on your computer, backs up your data to a hard drive at scheduled intervals)
- Mozy (Mac and Windows, $5.99/month, backs up your data remotely at scheduled intervals)
- BackBlaze (Mac and Windows, $5/month, backs up your data remotely at scheduled intervals)
- SpiderOak (Mac and Windows, free or $100/year, backs up your data remotely)
- DropBox (Mac and Windows, free or $10–$20/month, backs up your data remotely)

Bibliographic Management
There are a lot of good options out there for saving, sorting, and citing your sources. The key point is that you really should be using some kind of bibliographic management system. You'll regret it if you don't.

- Zotero (Mac and Windows, free)
- EndNote (Mac and Windows, $249.95)
- Mendeley (Mac, Windows, and Linux, free)
- Sente (Mac, $89.95)
- Bookends (Mac, $69)

File Renaming and Organization
If, for example, you take a lot of photos in an archive, you probably come home with tons of files with totally unintelligible names. Several tools can help you organize these assets and give them human-readable names.

- NameDropper (Windows, $10, batch renamer that allows you to set patterns)
- Belvedere (Windows, free, allows you to set rules to rename and organize files)
- Hazel (Mac, $21.95, allows you to set rules to rename and organize files)
- Automator (Mac, already installed on your computer, allows you to perform many actions on your files)

Indexers and "Everything Buckets"
Depending on how you work, you may find it important to grab and tag things — from the Internet or from "real life" — quickly and easily. There are some very good tools for this. Be careful, though: It's not enough to grab something. You have to be able to find it again, too!

- EverNote (Windows, Mac, Android, and iPhone; free or $45/year; captures and tags Web pages, photos, and other documents)
- Yojimbo (Mac and iPhone, $38.99, capture and tag notes and documents)
- Voodoo Pad (Mac and iPhone, $39.96, capture and tag notes and documents)
- SOHO Notes (Mac and iPhone; $39.99; capture, tag, and organize notes and documents and create custom forms)
- DEVONthink (Mac and iPhone; $49.95; indexes your files, allows you to organize them and add notes and metadata, offers "fuzzy" searching)

Databases
An "everything bucket" is a database, of course, but sometimes you need a tool that structures your data, too. Structure is great, but you should also be honest with yourself about whether the tool will fit easily into your workflow.

- askSam (Windows, $149.95, designed to make database-creation quick and easy)
- Microsoft Access (Mac and Windows, $139.99)
- FileMaker Pro (Windows, $179.00)
- Bento (Mac, $49.00)

Optical Character Recognition (OCR)
Your sources become much more findable when your run OCR on them. Of course, depending on the kinds of sources you gather, OCR may be imperfect (or impossible).

- ABYY FineReader (Windows and Mac, $49.99 and $99.99, respectively)
- Adobe Acrobat Pro (Windows and Mac, $404.10)
- OCRopus (Mac and Linux, free)
- EverNote (see above; EverNote automatically performs OCR on your documents)
- DEVONthink (see above; DEVONthink automatically performs OCR on your documents)

