

Developing lockers at LELA

Not as good as archiving but better than nothing

Jeff Good, MPI EVA (good@eva.mpg.de)

The problem

- Lots of linguists have lots of materials that need to be kept safe but aren't yet “archival”
- For descriptive resources that take decades to produce this is an especially acute problem

A possible solution

- Lockers
 - Safe storage houses (backed up, not easily stolen/crushed/dropped)
 - Restricted access, excepting “disaster”
 - Deposit and withdrawal primarily under collector’s control

Problems

- How can we implement lockers in a way that makes them easy to use but gives us some archival “control”
- Allotting space on a disk is probably not enough—we want to keep some track of what goes on in the locker
- Ideally, we’d even want semi-automated metadata creation...

A beginning

- Subversion-based autoversioning
WebDAV mount points
- There are undoubtedly other similar
solutions

Subversion

- Subversion (aka 'svn'): A version control system designed for teams of software developers
- It allows different programmers to work on different pieces of a large project and keeps track of who made what changes
- It also helps makes sure two people don't try to alter the same lines of code at the same time

WebDAV

- An extension of the HTTP protocol to let a web address behave like a file server
- Supported by common operating systems
- From the user's perspective little different from accessing any external server from the desktop
- I've tried it with OWL-types—and they don't seem to have any problem with it

Autoversioning

- Backup with memory
- All changes ever made to a given set of files are kept track of and you can revert to earlier versions of them
- The Subversion system keeps only stores changes to files to help save disk space
- This can be presented as a service to depositors

System in use

- We haven't fully exploited this kind of system at LELA yet
- But we have tested it in one case, and it works quite well
- One of our depositors has been working on a grammar for many years and wanted a safe place to keep the latest version in
- For this, it's worked without a problem

Hooks

- A possibility we haven't really exploited is the use of "hooks" that you can employ to get your system to do things whenever a file is added/changed
- A typical case for software developers is for the system to send out e-mails whenever a change to the code is made
- I hope soon to try using hooks to initiate automatic conversions for deposited files

Hooks

- A specific scenario of interest to me
 - Someone copies a FileMaker database
 - This activates a “post-commit” hook which exports the database to a text format
 - This format is then moved into an “archival” section of the locker

Browsers

- Subversion can also be associated with browser-based tools for examining the repository
- This makes it easier to explore the history of the changes to the system

Installation

- On a Mac, at least, installing Subversion isn't too hard
- Setting up new lockers is easy if you know your way around the Unix terminal
- But the difficulties are mostly on the archive's end
- The OWL doesn't know about any of the magic "under the hood"

Conclusion

- Setting up basic lockers is quite easy
- This is because the hard work of tool building has already been done for software development projects
- The hardest part for us is developing scripts for the “hooks”
- It would be nice if we could find ways to sneak in metadata, too...

Summary



File copying



Autoversioning



Hooks

Archiving
Miracles