Topics: Other Data Types and Sources

Approach: Learn new tools, Practice old ones

Outline:

What Have We Seen?
0. The three-level model: browser, script, tools
1. C programs: a communicating set of functions
2. Unix programs: a communicating set of programs
3. Web Programs: a communicating, distributed set of programs
4. Modes of communication:
   call-return, standard input/output, disk files
5. One picture describes all these levels

But There is More
Not all data is text -- there are images, sound, movies
Not all data is local -- generated, remote
Today we look at new tools to work with these

1. Tools for Images: netpbm (and ImageMagick)
   A project: generate a nice web album from a set of images
   Operations on text - search, sort, edit, compare, count, ...
   Operations on Images - resize, invert, rotate, emboss, oilify
   Intro to netpbm
   Solutions: (a) put in a directory, (b) make an index,
             (c) make an index with thumbnails

2. Tools to get remote data: wget (and curl)
   Problem - the weather
   Solution - ask the BBC
   Technique - use wget to pull data, then sed to format it

3. Using regular expressions to process text
   grep, sed, awk, perl, ...
   The basics: ., $, [], [^], *
   search and replace using marked substrings

4. Combos I - processing remote images
   The math dept webcam is a remote source of data

5. Combos II - an album of remote images
   The hubble telescope can provide a lot of images
   wget, netpbm, regex/sed can help here

6. MBTA Commuter Rail
   We close with our first example -- building a new sched

7. Moral: tools, glue, combinations, skill, imagination