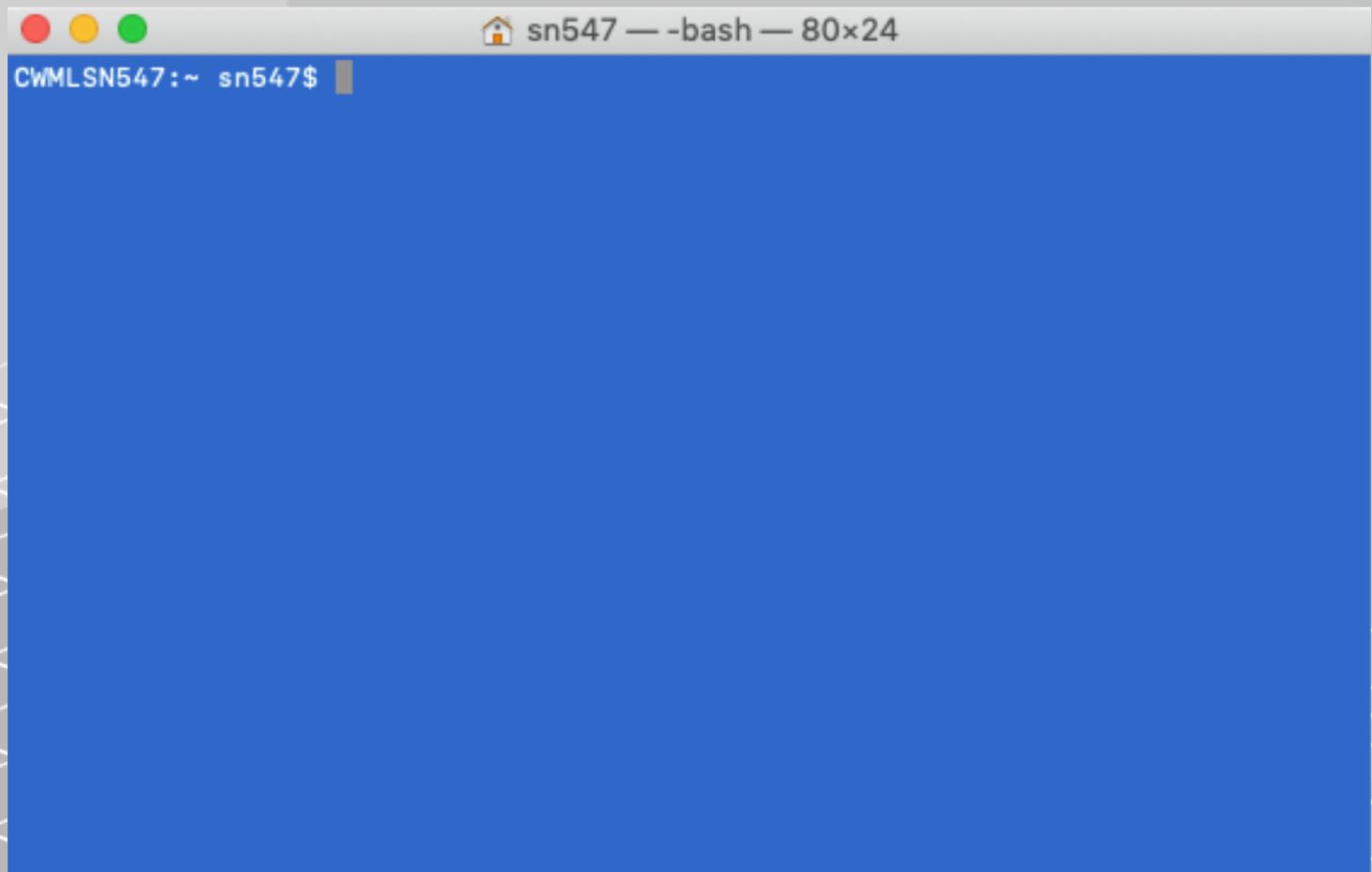


# Unix Shell: Working with Data

Yale UNIVERSITY LIBRARY  
*Harvey Cushing/John Hay Whitney Medical Library*

# What is the Shell?

- A way to interact with your computer without software
- A command line interface (CLI)

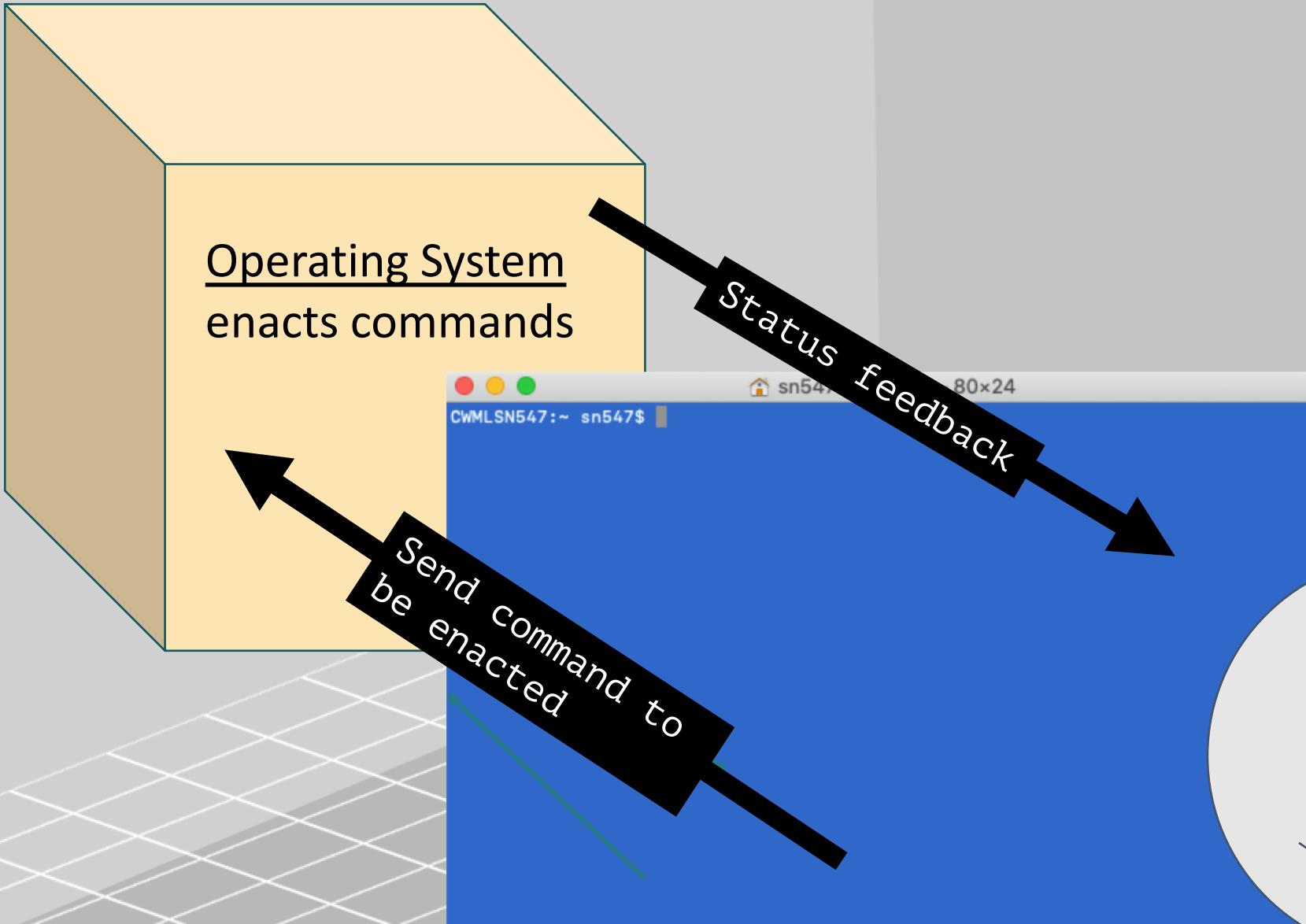


# Shell uses

- Obtain data from websites, APIs, databases, and spreadsheets
- Explore and analyze your file system
- Manage your data workflow
- Create reusable tools (i.e. scripts)
- Leverage programming languages across your data
- Better understand how your computer "thinks"

# Relevant terms

- Command line prompts: commands typed into the shell
- Terminal: the application where commands are typed
- Shell: where the command is sent to by the terminal to be interpreted
- Bash: a specific instance of the shell (Bourne again shell)
- Operating system: executes commands and connects hardware



User  
types commands

# Why are we using the Unix Shell

- High performance computing clusters (HPC) at Yale run on Linux
- 95% of the top 500 supercomputers are running Linux



# Unix command tools

- Find explanations for Unix commands
  - <https://explainshell.com/>
- Find bugs and mistakes in your shell scripts
  - <https://www.shellcheck.net/>
- Use the `man` command while working in the shell

# Continue learning

- Access LinkedIn Learning training videos through Yale
- The Yale Center for Research Computing provides training
- O'Reilly Text book: Data Science at the Command Line  
(Jeroen Janssens 2019)
- Ask Unix questions during Data Management office hour sessions

# Unix demo topics

- File navigation
- Creating new files and folders
- Moving files
- Copying files
- ReadMe files
- Finding data within files
- Merging and collecting data from files

Contact:  
medicaldata@yale.edu



Yale UNIVERSITY LIBRARY

*Harvey Cushing/John Hay Whitney Medical Library*