If you plan to process raw high throughput sequencing data yourself, you will need to learn shell.

Learning objectives review

✣ Recognize the importance of Shell for the analysis of high throughput sequencing data

✣ Install and locate the software/tools necessary for accessing the command line

✣ Describe the organization of the file directory and how to move through it

✣ Demonstrate different ways to manipulate files, such as moving, copying, and renaming

✣ Compare different commands for examining small or large files

✣ Create new files and make edits to existing files with nano

✣ Contrast shell scripts with direct commands

Now that you have completed Basic Shell …

The Foundation - Basic Shell
Feb 21st, 2024

Accelerate with Automation - Making your code work for you
March 20th

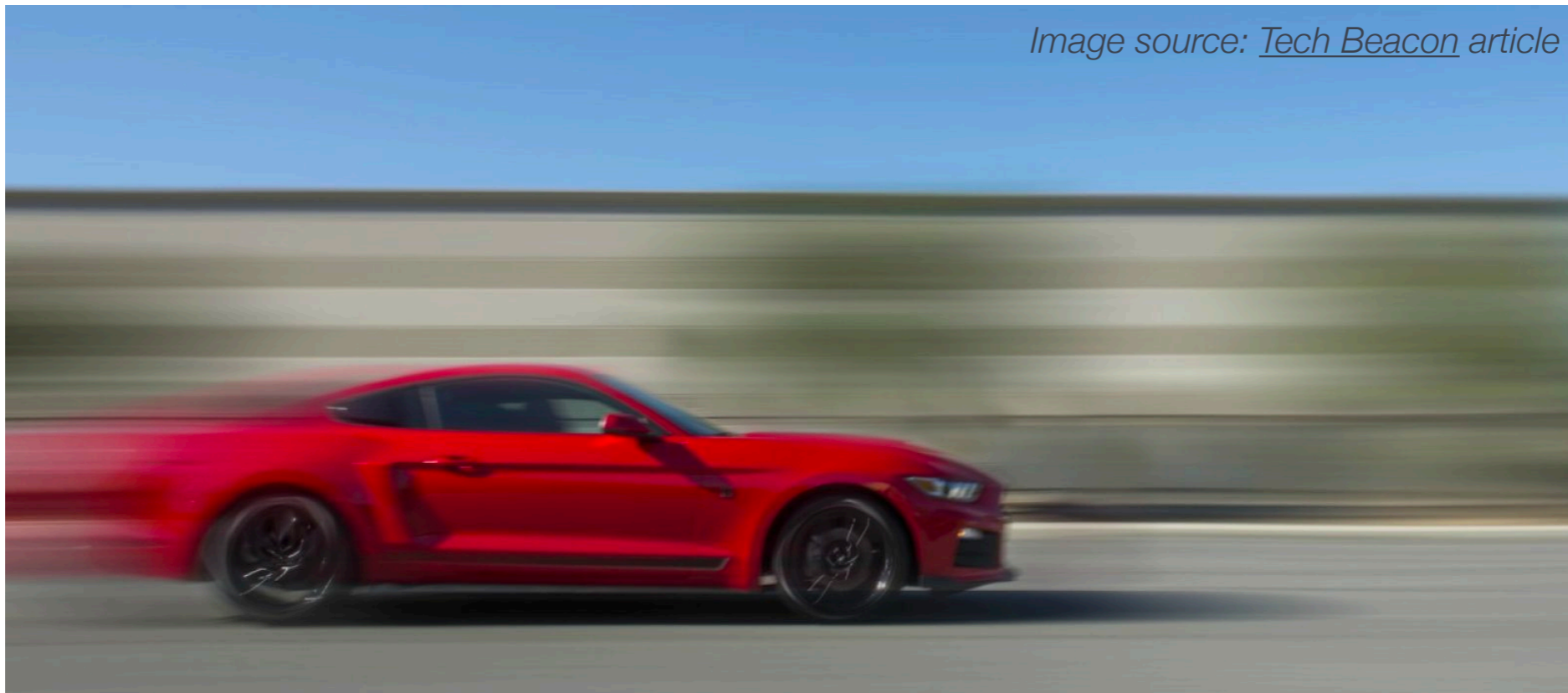# Needle in a Haystack - Finding and summarizing data from colossal files
April 17th

Shell Tips and Tricks on the O2 cluster
May 15th

**Accelerate with Automation -
Making your code work for you
March 20th**

- Run the same set of tasks for different input files

- Create flexible scripts that can be used for different data sets

  - Positional parameters to take in information to be used in the script

  - Bash variables to store information to be used later

Image source: Tech Beacon article

Needle in a Haystack - Finding and summarizing data from colossal files
April 17th

Use commands like grep, awk and sed to quickly sift through your data!

- Find and select specific information from large files

- Subset and create new files

- Easily summarize information from within a file

Image source: Cartoon illustration by John P. Weiss

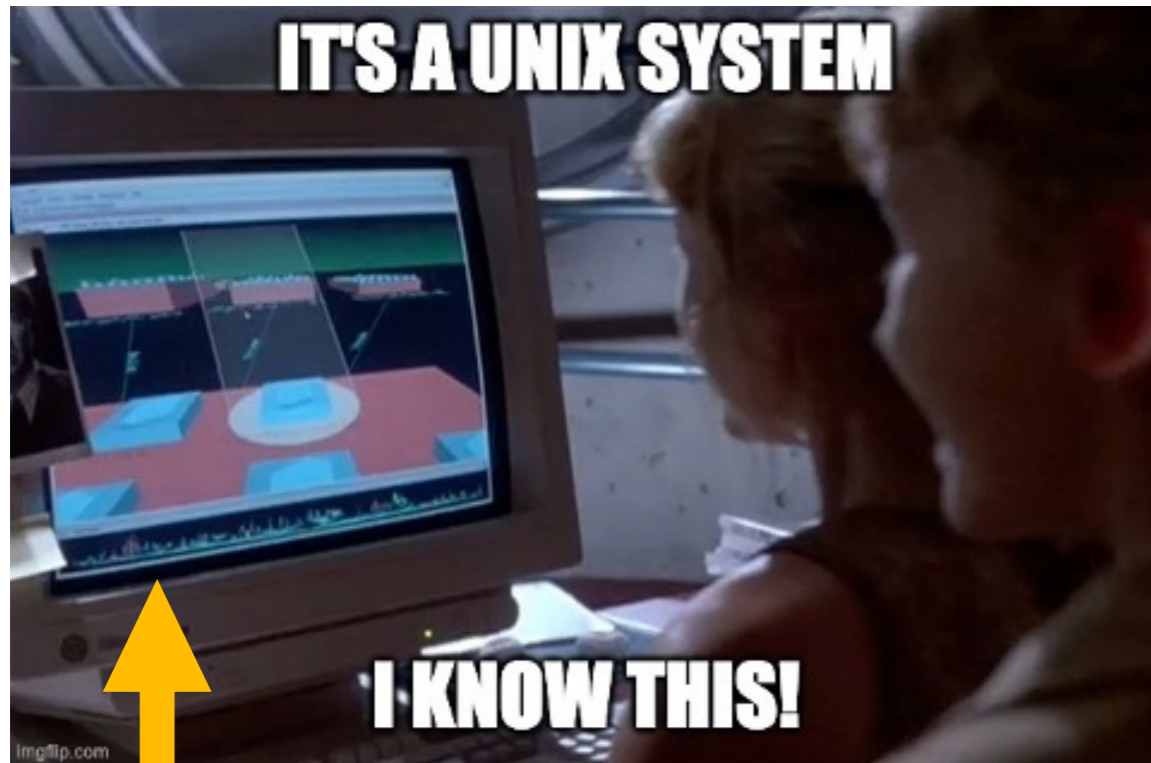**Shell Tips and Tricks on the O2 cluster**
May 15th

- Shortcuts and best practices for making your experience working on O2 much smoother!

- In collaboration with HMSRC we will teach you how to:

  - Copy large files to and from the cluster

  - Create an environment that loads up for each session on O2 (with specific software and tools loaded automatically)

  - Run lots of jobs at the same time and get things done even faster

Looking forward

Now you can save your friends from dinosaur attacks!

Real GUI UNIX system from the 1990s (Silicon Graphics 3D File System Navigator, fsn, for IRIX). **Could have locked the door a lot faster with a command line interface!**

Accelerate with Automation - Making your code work for you
March 20th

Needle in a Haystack - Finding and summarizing data from colossal files
April 17th

Shell Tips and Tricks on the O2 cluster
May 15th

We hope to see you soon!