

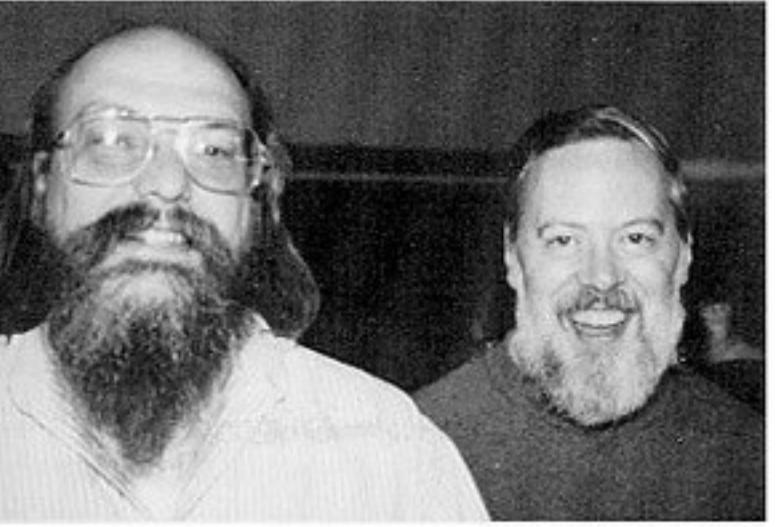
Introduction to Unix and Unix-like systems (Linux and MacOS)

Ramiro Logares, ICM-CSIC Barcelona



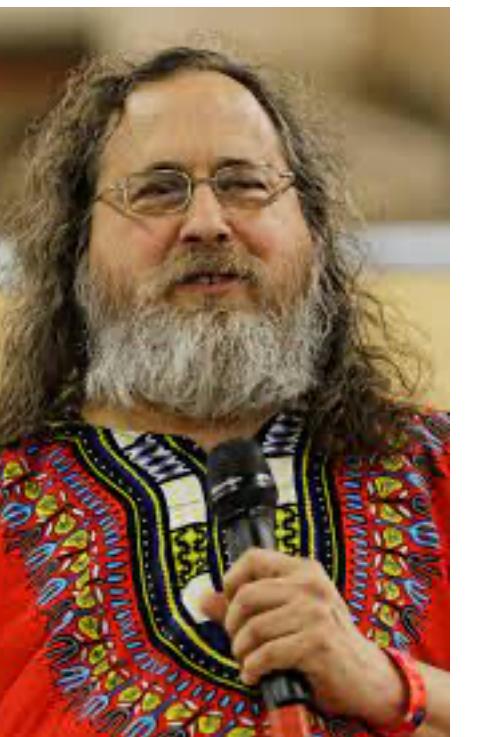
Unix history

- Started to be developed in the 60' at Bell Labs / AT&T for the PDP-7
- Coded mostly in C (also developed by AT&T)
- Modular philosophy (combine or “pipe” several small programs)
- Multi-tasking (resources are shared)
- Multiuser (can handle many users simultaneously)
- Initially, Unix was not open source
- The advent of Open Source
 - 1990s: GNU Open Source tools + kernel = Linux



- **Today Unix is widely used in:**

- Large supercomputers
- Computer clusters & servers
- Companies & organisations
- MacOS (BSD-like)
- Research (specially Bioinformatics)



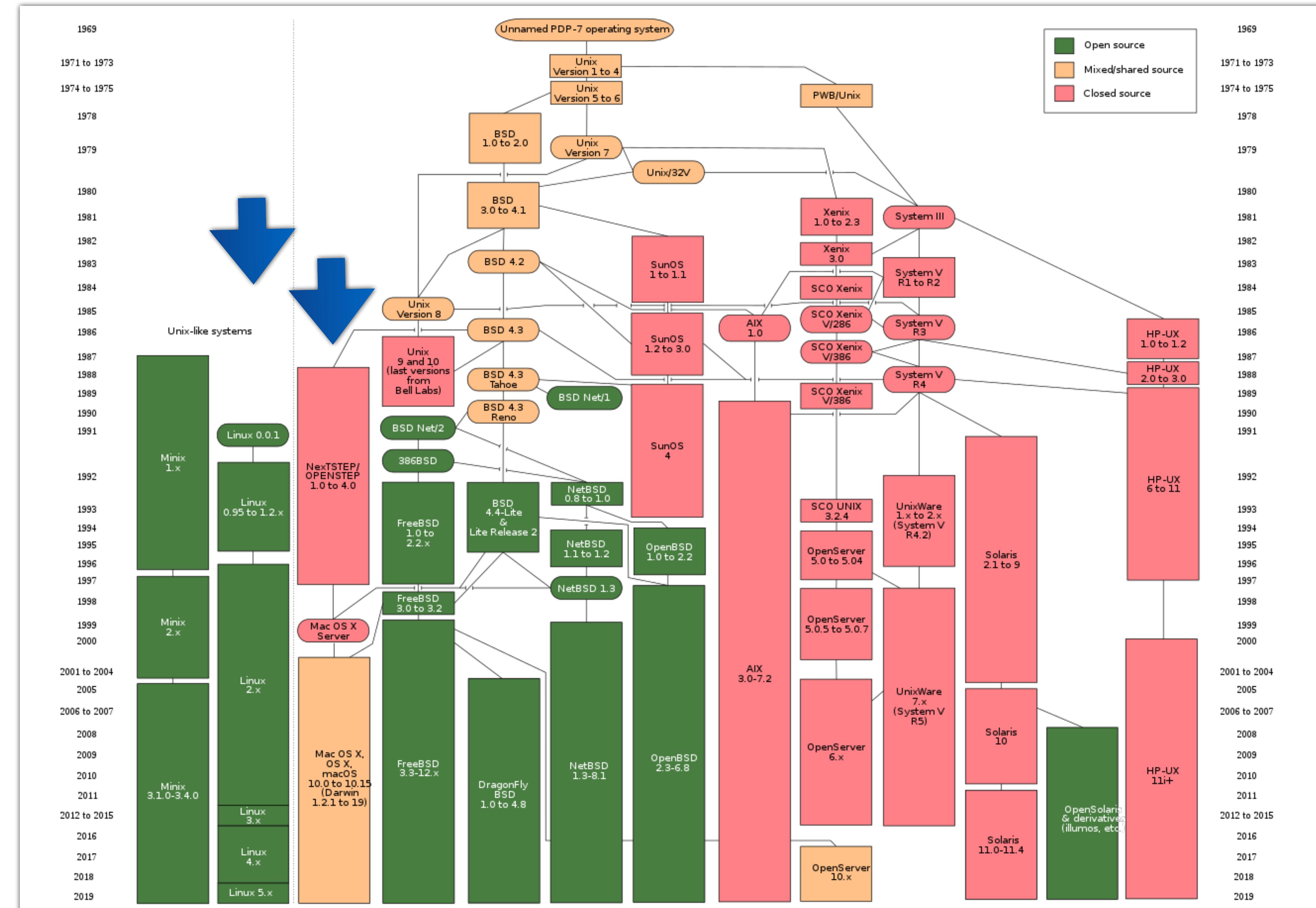
UNIX

stands for

Uniplexed Information and
Computing Service



Abbreviations.com



- **Advantages**

- Great stability
- Security
- Good control of users and processes
- Can handle large computing load
- Open Source FREE versions: Several options (Ubuntu, CentOS, Debian, etc.)
- Large community of developers
- Large amount of free software (specially for bioinfo)
- Excellent connectivity between computers

- **Disadvantages**

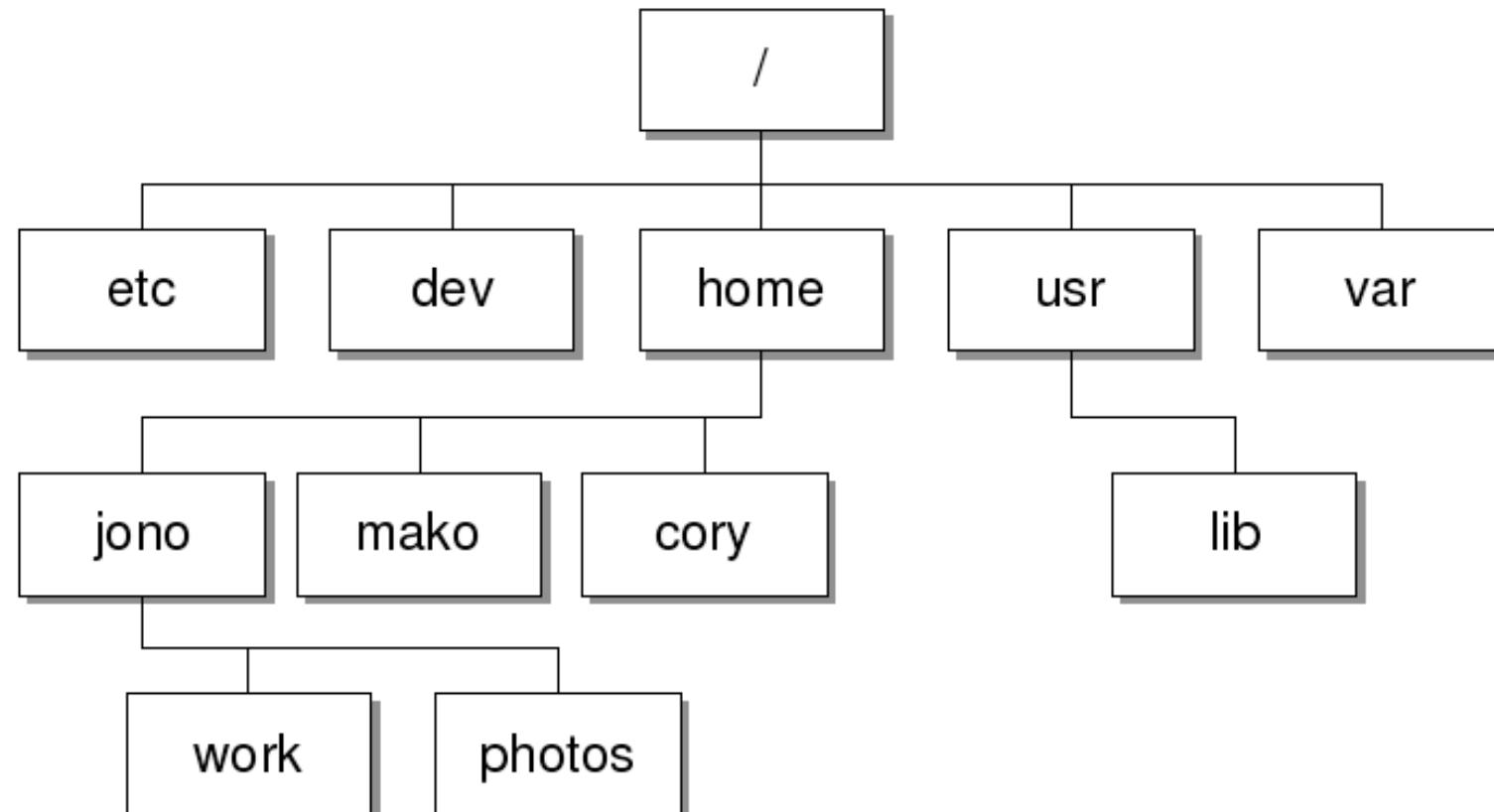
- Compatibility with popular tools
- Some hardware may not be supported
- Some applications may need more knowledge to use them

System

Basics

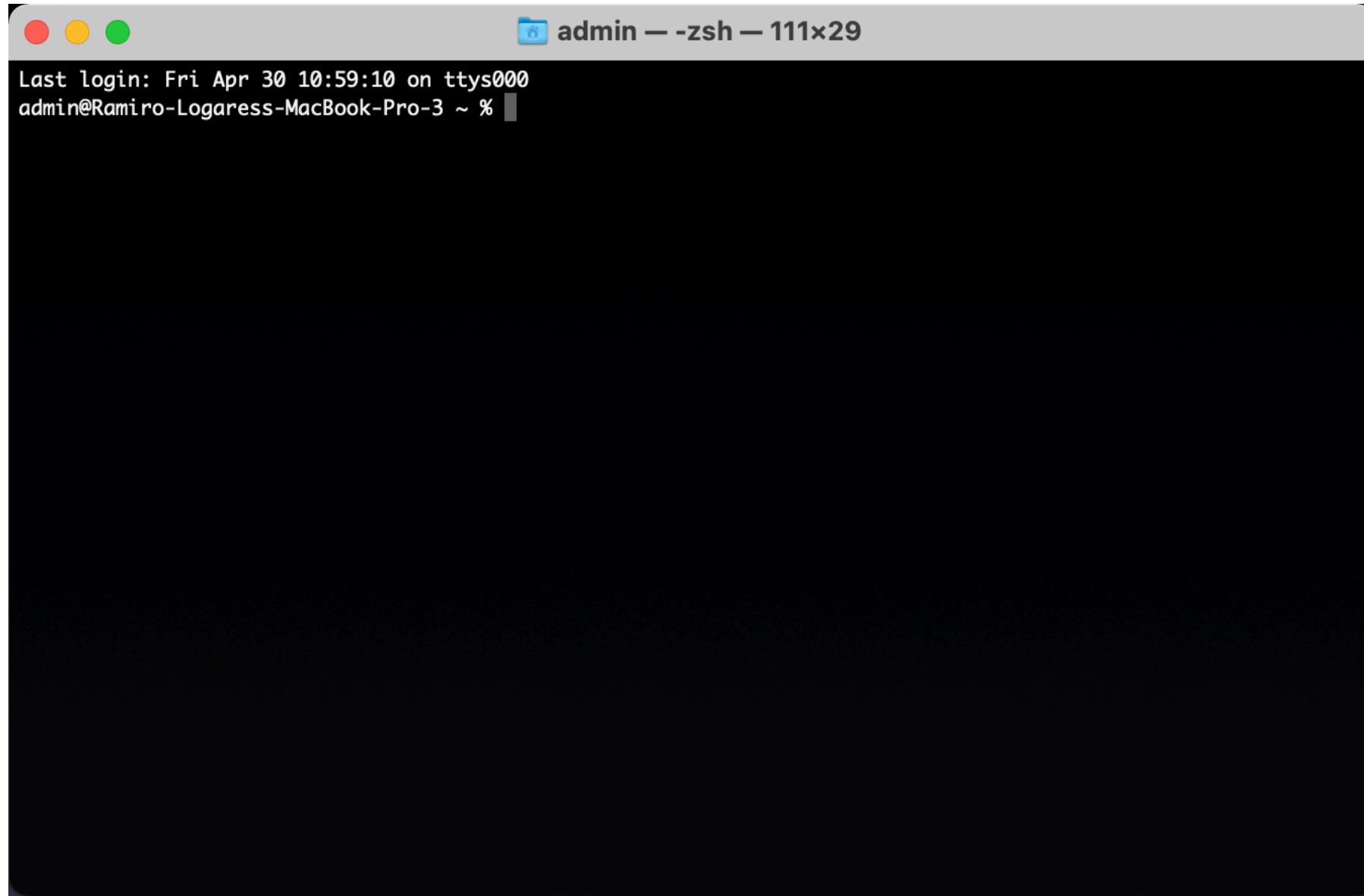
- User: performs several tasks, but has restricted access to specific functions (e.g. installing software)
- Has a “home” directory, where most operations are allowed (editing, generating files, deleting, etc.)
- Files belong to users and groups
- Root: super-user, can do everything, normally is the system administrator

Filesystem hierarchy



/ : root
/bin: or /usr/bin: files needed by Linux (binaries)
/boot: files needed for booting. Contains the kernel.
/etc: configuration files for Linux.
/dev: devices available (e.g. External Hard Drive)
/home: contain user homes
/lib: shared libraries
/usr: user applications
/var: variable data
/tmp: temporary files
/lost+found:files related to crashes
/opt: software and add-on packages (you may install software here)

Terminal



Mac

The homepage of the iTerm2 website. It features a large logo with a green dollar sign (\$) icon inside a dark square. The title "iTerm2" is prominently displayed in white. Below the title, a subtitle reads "iTerm2 is a terminal emulator for macOS that does amazing things." A navigation bar at the top includes links for Home, News, Features, FAQ, Documentation, Downloads, and a red "Donate" button. The main content area contains sections for "What is iTerm2?", "Why Do I Want It?", and "How Do I Use It?", each with descriptive text and links. A large "Download" button is centered at the bottom of the page.

Terminal available in:

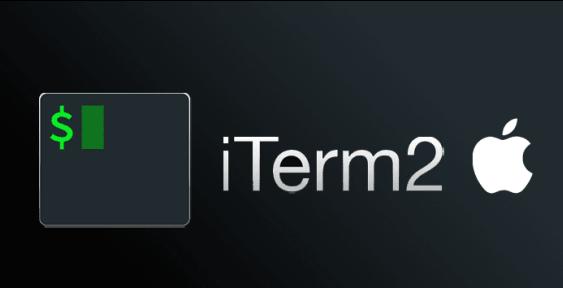
- all Unix versions
- MacOS
- Windows emulators

Windows

Check list of linux emulators:

<https://www.fosslinux.com/43264/best-linux-emulator-windows.htm>

```
[rlogares@marbits ~]$ si
PARTITION HOSTNAMES AVAIL STATE CPUS CPU_LOAD MEMORY FREE_MEM S:C:T REASON
main      c01       up   mix   48    3.67  773942  761589  2:12:2 none
main      c02       up   mix   48    29.44  773942  298384  2:12:2 none
main      c05       up   mix   48    24.75  257847  237107  2:12:2 none
main      c06       up   mix   48    27.04  257847  209366  2:12:2 none
main      c07       up   mix   48    24.51  257847  145737  2:12:2 none
main      c08       up   mix   48    7.90   257847  219942  2:12:2 none
main      c09       up   mix   48    8.10   257847  205758  2:12:2 none
main      c22       up   mix   24    7.02   80544   76375   2:6:2 none
main      c03       up   alloc  48    22.13  257847  185893  2:12:2 none
main      c04       up   alloc  48    39.53  257847  233318  2:12:2 none
main      c10      up   alloc  48    1.62   257847  229102  2:12:2 none
main      c13      up   alloc  24    1.16   112800  74372   2:6:2 none
main      c11      up   idle   24    0.01   80544   78076   2:6:2 none
main      c12      up   idle   24    0.01   80544   78103   2:6:2 none
main      c14      up   idle   24    0.01   80544   78179   2:6:2 none
main      c15      up   idle   24    0.01   80544   78176   2:6:2 none
main      c16      up   idle   24    0.01   80544   78195   2:6:2 none
main      c17      up   idle   24    0.01   80544   78195   2:6:2 none
main      c18      up   idle   24    0.01   80544   78199   2:6:2 none
main      c19      up   idle   24    0.01   64416   62251   2:6:2 none
main      c20      up   idle   24    0.01   62429   62224   2:6:2 none
main      c21      up   idle   24    0.01   80544   78101   2:6:2 none
```



```
[rlogares@marbits ~]$ logout
Connection to marbits.cmima.csic.es closed.
admin@Ramiro-Logaress-MacBook-Pro-3 ~ %
```

```
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 22 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 24 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 12 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 14 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 13 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides %
Session Contents Restored on 29 Apr 2021 at 22:49
Last login: Thu Apr 29 22:48:07 on console
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides %
Session Contents Restored on 30 Apr 2021 at 10:59
Last login: Fri Apr 30 10:58:54 on console
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides %
code.slide1.txt
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 18 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides % highlight -0 rtf code.slide1.txt --line-numbers --font-size 14 --font I
nconsolata --style solarized-dark -W -J 150 -j 3 --syntax r | pbcopy
admin@Ramiro-Logaress-MacBook-Pro-3 code.for.slides %
```

Some commands

Working with directories & files

ls = "list files/folders"
ls or ll

cd = "change directory (move around)"
cd (home) ; cd dir (go to dir); cd .. (one level up); cd ../../ (two levels up)

pwd = "see working directory"

cp = "copy files"
cp /route/to/file/to/copy/file /new/location/of/file
cp -r /dir/to/copy/dir /location/for/dir/copy # Copy dir and its contents

mv = "move or rename files"
mv /old/location/for/my/file /new/location/for/my/file
mv old_name new_name

rm = "delete a file or directory"
rm /file/to/remove
rm -rf /remove/directory/and/all/contents (use carefully!!)

Some commands

Working with directories & files

rmdir = "delete a directory"

mkdir = "make directory"

less/cat = "see what's inside a file"

head/tail = "see beginning / end of a file"

tail -n 300 file (see last 300 lines); head -n 300

cat = "concatenate files (also print files to the terminal)"

cat file1 file2 > new.cat.file

Permissions

chmod = "change permissions"

chmod a+x executable.file

chmod 755 file (owner:group:world permissions)

chown = "change ownership of a file or directory"

chown user:group file # Assign file to a new user and group

File permissions

```
shum@sol:~ $ ls -l
total 20
drwxr-xr-x 2 shum staff 4096 Jan 16 22:04 Mail
drwxr-xr-x 3 shum staff 4096 Jan 16 14:15 csc128
drwxr-xr-x 2 shum staff 4096 Jan 13 16:42 public
drwxr-xr-x 2 shum staff 4096 Jan 16 14:07 public_html
-rw-r--r-- 1 shum staff 628 Jan 15 20:04 verse
```

The diagram illustrates the structure of the `ls -l` command output. It shows a sample listing of files and then breaks down each entry into its components:

- file type**: The first column indicates the file type (e.g., directory, regular file).
- permissions**: The second column shows the permissions for the owner, group, and others. The first three characters represent the owner's permissions (read, write, execute), the next three represent the group's permissions, and the last three represent other users' permissions.
- number of hard links**: The third column shows the number of hard links to the file.
- user (owner) name**: The fourth column shows the user name who owns the file.
- group name**: The fifth column shows the group name associated with the file.
- size**: The sixth column shows the size of the file in bytes.
- date/time last modified**: The seventh column shows the date and time the file was last modified.
- filename**: The final column shows the name of the file.

Below the diagram, the permissions `rwx` are further broken down into their individual components:

- executable**: The 'x' component.
- writeable**: The 'w' component.
- readable**: The 'r' component.

Annotations also point to the "other (everyone) permissions" (red text) and "group permissions" (cyan text) sections of the permissions string.

`drwxrwxrwx`

d = Directory

r = Read

w = Write

x = Execute

`chmod 777`

`rwx|rwx|rwx`

Owner|Group|Others

7	<code>rwx</code>	<code>111</code>
6	<code>rw-</code>	<code>110</code>
5	<code>r-x</code>	<code>101</code>
4	<code>r--</code>	<code>100</code>
3	<code>-wx</code>	<code>011</code>
2	<code>-w-</code>	<code>010</code>
1	<code>--x</code>	<code>001</code>
0	<code>---</code>	<code>000</code>

Some commands

Processes & super user

top = "see running processes"

kill = "kill a process"

sudo = "request root access for a task (not always available)"

#Common Tools

emacs/vim/nano: "useful text editors"

grep = "search text for a pattern"

 grep pattern file # search pattern in file

 grep -c pattern file # count pattern in file

awk = "Powerful text editor"

 awk '{print \$1, \$5}' table # Prints fields 1 and 5 from table

sed = "search and replace patterns in files"

 sed "s/rat/cat/g" infile > outfile # Search "rat" and replace by "cat" in infile

 sed "s/rat//g" infile > outfile # Remove "rat" from infile

Some commands

```
# Processes & super user

# Combine commands with a pipe | (redirects the output of one command as input to another)

output | input
    ll | grep -c pattern (count filenames with pattern)
    ll | grep pattern > file
    cut: "cuts characters from a file"
    less file | cut -f 2 -d "-" # Bring the 2nd field from file, separated by "-"

sort = "sorts records"

uniq = "retrieves unique records"

# Often used together as sort | uniq

wget (get files or directories from remote computers)
    wget http://web.archive.com/file.txt

# Remote connectivity
    ssh = remote connection to a machine
    ssh user@machine (remote connection)
    scp file.to.copy user@machine:/path/to/location/to/copy # copy to a remote
computer
    scp user@machine:/path/to/file . # copy from a remote location to local computer
```

Some commands

#Manual

```
man = documentation about a command (also program -h)  
#Available for all core functions and programs
```

Execute a program/script
.executable

Add programs to your search path permanently

```
edit .bashrc # This is a file at your home with environmental variables  
emacs .bashrc  
export PATH=/path/to/file:$PATH # Add path to file to search path
```

Generate temporal variables

```
var1=/path/to/file # set up variable  
${var1} # calls variable
```

Some commands

Loops

Runs a function over several files

If you have: file1, file2, ..., filen

Then:

```
for i in $(ls); do cut -f 2 ${i} > ${i}.out; done
```

Cuts field 2 in each file and sends it to a corresponding output

Execute functions in bashscripts

emacs my.bash.script.sh

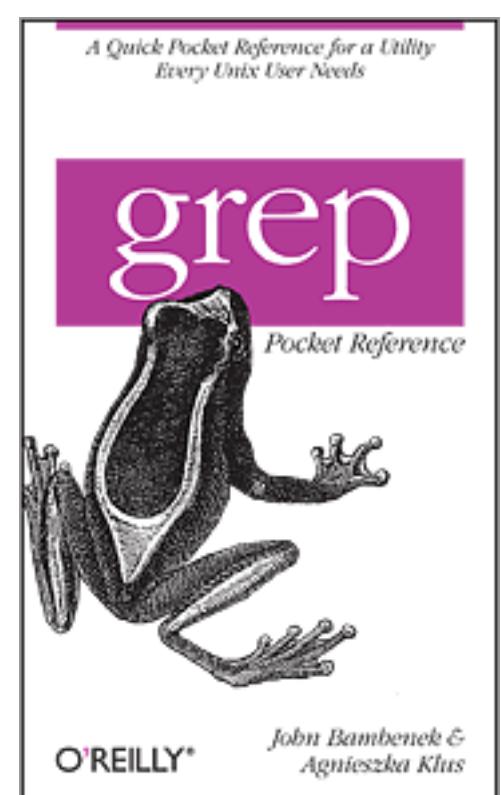
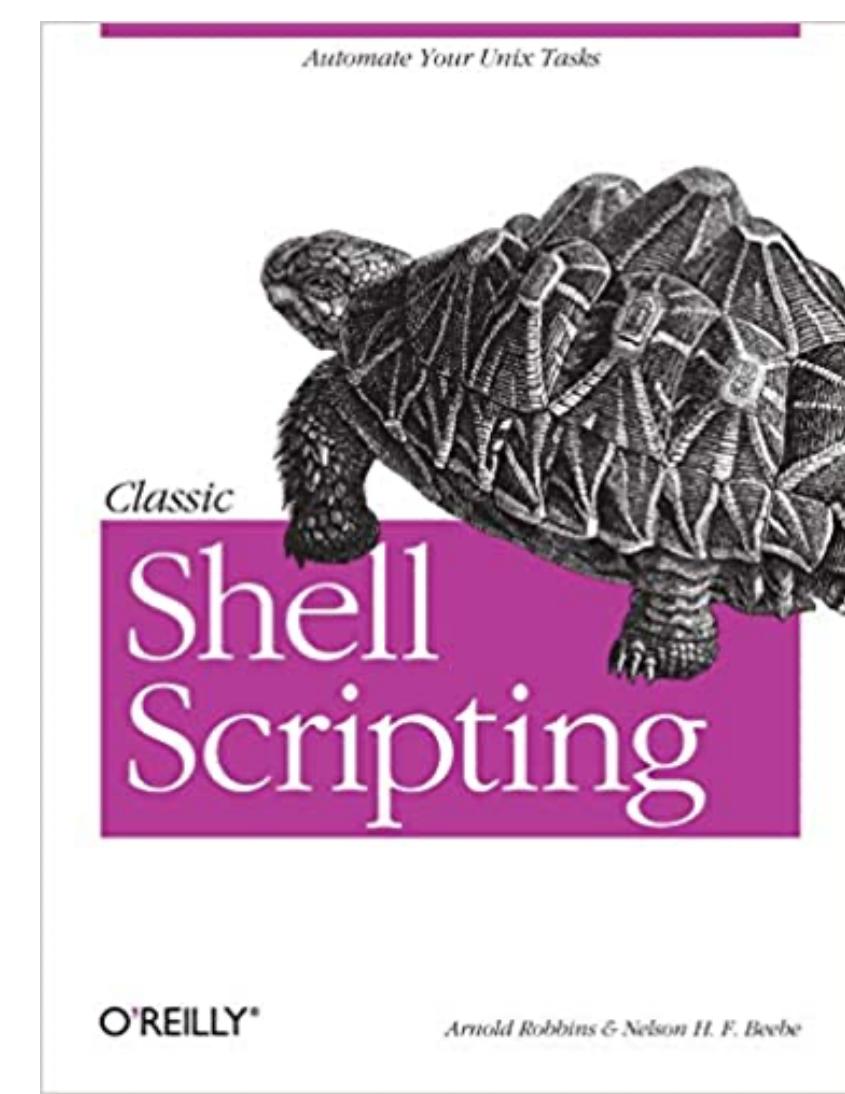
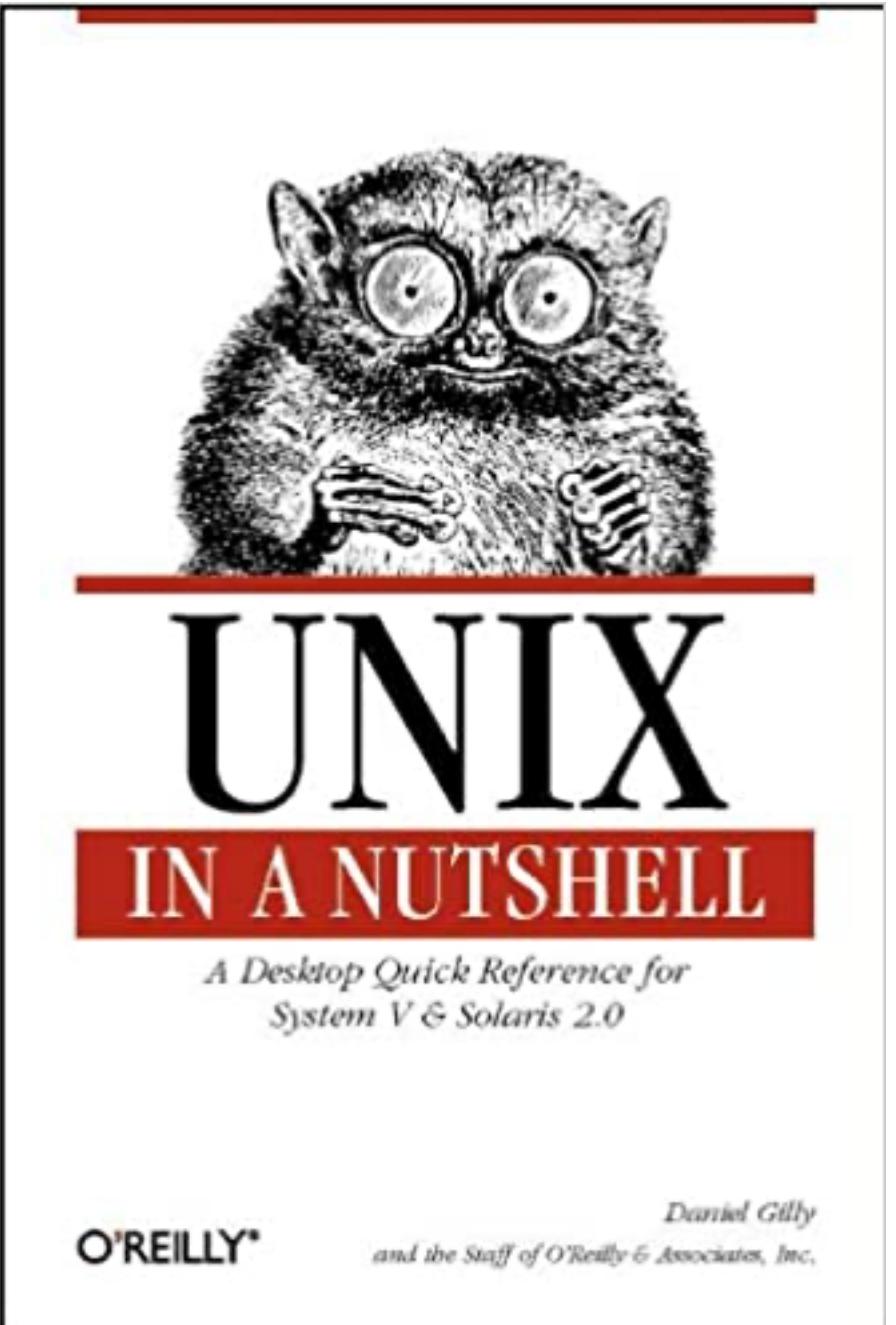
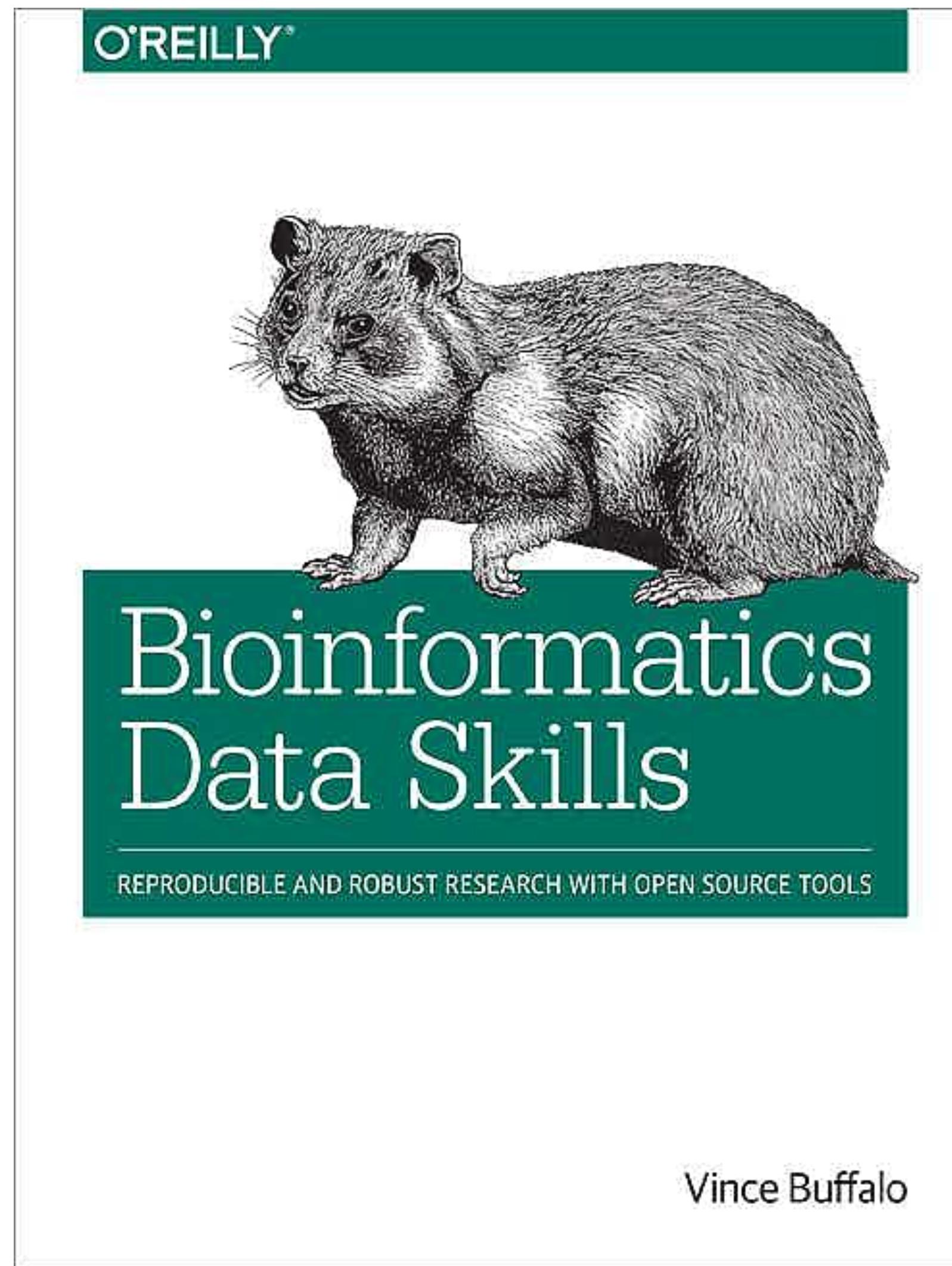
```
#!/bin/bash
for i in $(ls); do cut -f 2 ${i} > ${i}.out; done
```

```
chmod a+x my.bash.script.sh # make script executable
./my.bash.script.sh # execute script
```

File Commands	System Info
ls - directory listing ls -al - formatted listing with hidden files cd dir - change directory to <i>dir</i> cd ~ - change to home pwd - show current directory mkdir dir - create a directory <i>dir</i> rm file - delete <i>file</i> rm -r dir - delete directory <i>dir</i> rm -f file - force remove <i>file</i> rm -rf dir - force remove directory <i>dir</i> * cp file1 file2 - copy <i>file1</i> to <i>file2</i> cp -r dir1 dir2 - copy <i>dir1</i> to <i>dir2</i> ; create <i>dir2</i> if it doesn't exist mv file1 file2 - rename or move <i>file1</i> to <i>file2</i> if <i>file2</i> is an existing directory, moves <i>file1</i> into directory <i>file2</i> ln -s file link - create symbolic link <i>link</i> to <i>file</i> touch file - create or update <i>file</i> cat > file - places standard input into <i>file</i> more file - output the contents of <i>file</i> head file - output the first 10 lines of <i>file</i> tail file - output the last 10 lines of <i>file</i> tail -f file - output the contents of <i>file</i> as it grows, starting with the last 10 lines	date - show the current date and time cal - show this month's calendar uptime - show current uptime w - display who is online whoami - who you are logged in as finger user - display information about <i>user</i> uname -a - show kernel information cat /proc/cpuinfo - cpu information cat /proc/meminfo - memory information man command - show the manual for <i>command</i> df - show disk usage du - show directory space usage free - show memory and swap usage whereis app - show possible locations of <i>app</i> which app - show which <i>app</i> will be run by default
Process Management	Compression
ps - display your currently active processes top - display all running processes kill pid - kill process id <i>pid</i> killall proc - kill all processes named <i>proc</i> * bg - lists stopped or background jobs; resume a stopped job in the background fg - brings the most recent job to foreground fg n - brings job <i>n</i> to the foreground	tar cf file.tar files - create a tar named <i>file.tar</i> containing <i>files</i> tar xf file.tar - extract the files from <i>file.tar</i> tar czf file.tar.gz files - create a tar with Gzip compression tar xzf file.tar.gz - extract a tar using Gzip tar cjf file.tar.bz2 - create a tar with Bzip2 compression tar xjf file.tar.bz2 - extract a tar using Bzip2 gzip file - compresses <i>file</i> and renames it to <i>file.gz</i> gzip -d file.gz - decompresses <i>file.gz</i> back to <i>file</i>
File Permissions	Network
chmod octal file - change the permissions of <i>file</i> to <i>octal</i> , which can be found separately for user, group, and world by adding: <ul style="list-style-type: none"> ● 4 - read (r) ● 2 - write (w) ● 1 - execute (x) Examples: chmod 777 - read, write, execute for all chmod 755 - rwx for owner, rx for group and world For more options, see man chmod .	ping host - ping <i>host</i> and output results whois domain - get whois information for <i>domain</i> dig domain - get DNS information for <i>domain</i> dig -x host - reverse lookup <i>host</i> wget file - download <i>file</i> wget -c file - continue a stopped download
SSH	Installation
ssh user@host - connect to <i>host</i> as <i>user</i> ssh -p port user@host - connect to <i>host</i> on port <i>port</i> as <i>user</i> ssh-copy-id user@host - add your key to <i>host</i> for <i>user</i> to enable a keyed or passwordless login	Install from source: ./configure make make install dpkg -i pkg.deb - install a package (Debian) rpm -Uvh pkg.rpm - install a package (RPM)
Searching	Shortcuts
grep pattern files - search for <i>pattern</i> in <i>files</i> grep -r pattern dir - search recursively for <i>pattern</i> in <i>dir</i> command grep pattern - search for <i>pattern</i> in the output of <i>command</i> locate file - find all instances of <i>file</i>	Ctrl+C - halts the current command Ctrl+Z - stops the current command, resume with fg in the foreground or bg in the background Ctrl+D - log out of current session, similar to exit Ctrl+W - erases one word in the current line Ctrl+U - erases the whole line Ctrl+R - type to bring up a recent command !! - repeats the last command exit - log out of current session



- In bioinformatics, Unix is your friend
- There are plenty of free resources online
- Solutions to many problems are already indicated in one of the many forums
- Use ChatGPT, Gemini, Claude, Copilot, etc.
- You can easily work on local and remote computers and switch between them



- Even though most analyses that you'll see in this course are done in R, some steps are carried out in unix (e.g. initial sequence cleaning)
- Alternative analyses can be carried out using other tools, e.g. VSEARCH, that run in unix
- Normally, a substantial part of the bioinformatics work is done in clusters that run unix
- Moving between R and unix is normal (and easy)
- R runs smoothly in unix environments





Unix tutorial

1. Open the terminal, go to your "home" folder or another chosen location
2. Create a directory called "biocourse" (use cd)
3. Go into the biocourse directory, and create the file "norwegian.black.metal.bands" (use nano, emacs, vi) and write "windir, gorgoroth, emperor, abbath, urgehal"

If you have issues with the text editors, then run

```
echo "windir, gorgoroth, emperor, abbath, urgehal" > norwegian.black.metal.bands
```

What did we do in the line above?

4. Count how many words we have in the file norwegian.black.metal.bands (use wc -w)
5. Concatenate norwegian.black.metal.bands into a new file with repeated contents

```
cat norwegian.black.metal.bands norwegian.black.metal.bands > norwegian.black.metal.bands.x2
```
6. Open the new file norwegian.black.metal.bands.x2
7. Count again the number of words
8. See the unique words:

```
less norwegian.black.metal.bands.x2 | sort | uniq
```
9. Match the pattern "emperor" in the file norwegian.black.metal.bands.x2

```
grep --color "emperor" norwegian.black.metal.bands.x2
```

What do the results tell you?

```
# Unix tutorial
```

10. Paste the contents of norwegian.black.metal.bands.x2 next to each other
paste norwegian.black.metal.bands.x2 norwegian.black.metal.bands.x2 > norwegian.black.metal.bands.x4
See the contents and match the pattern "abbath"
11. Concatenate the contents of one file after the other
cat norwegian.black.metal.bands.x4 norwegian.black.metal.bands.x4 > norwegian.black.metal.bands.x8
12. Replace in the document the band "gorgoroth" by "norwegian.reggaeton"
sed "s/gorgoroth/norwegian.reggaeton/g" norwegian.black.metal.bands.x8 > norwegian.black.metal.bands.x8.v2
13. Let's change the order of bands with awk
less norwegian.black.metal.bands.x8.v2 | awk '{print\$2,\$1,\$5,\$3,\$10,\$6,\$9,\$8}' >norwegian.black.metal.bands.x8.v2.reordered
With "print" we indicate the order of the fields in the new file
14. We add the prefix bandname to each field
less norwegian.black.metal.bands.x8.v2.reordered | awk '{print "bandname", \$0}'
15. Have a look to the permissions of the files that were created:
ls -l norwegian.black.metal.bands*
What do these permissions mean?
16. Change the permissions of one of the created files
chmod 777 norwegian.black.metal.bands
What permissions did we give to this file? What can group and other users can do with it?
17. Remove the file
rm norwegian.black.metal.bands
18. Go one level up from the directory biocourse. Try to remove it with:
rmdir biocourse
What is happening? how would you solve this? NB: no need to actually remove this directory unless you want to do it.