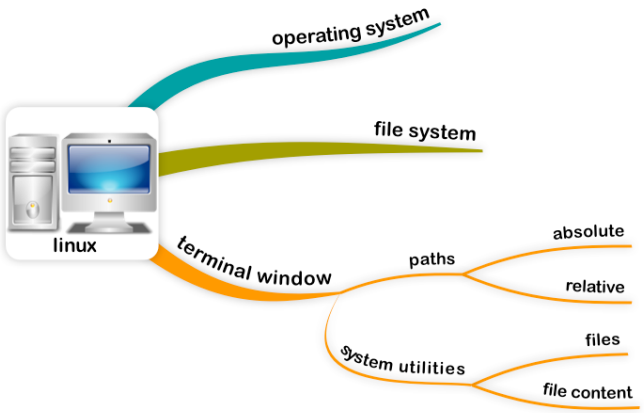
Bioinformatik
Stefanie Hartmann
Wintersemester 2019 / 2020, Universität Potsdam

The Linux operating system
Nov 01, 2019

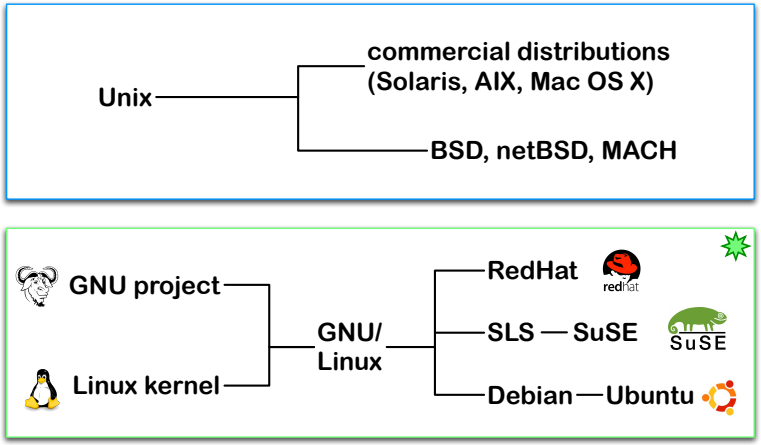
Today's topics



Unix / Linux

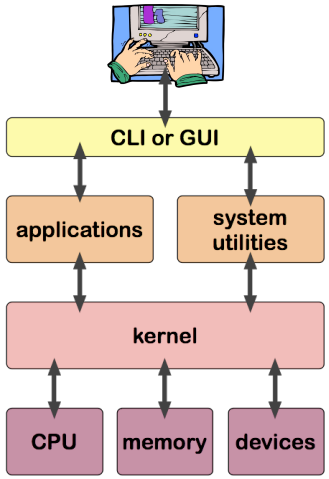
- multi-user, multi-tasking
- stability
- performance
- low cost
- available for most hardware platforms
- a standard file system
- the use of a large number of system utilities (that can also be used in sequence)
- many distributions exist

Unix / Linux



open source: source code for the kernel and application programs are freely available

Operating Systems



CLI
- command line interface

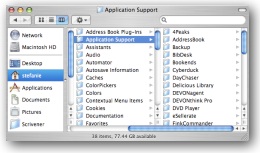
GUI
- graphical user interface

kernel
- central core of an OS

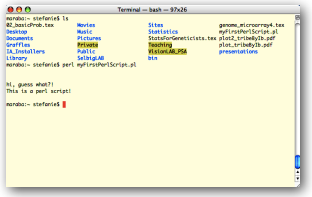
CPU
- central processing unit (the processor)

GUI vs. CLI

graphical user interfaces



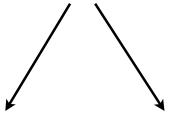
command line interfaces



- ➡ CLI software
- ➡ reproduce and automate analyses
- ➡ work on remote computers

The terminal (command line)

- a human-computer interface
- a text-only window
- copy and paste can be done using the mouse
- a “prompt” indicates that the computer is ready to accept typed input
- allows to type in a command that is then executed (run)



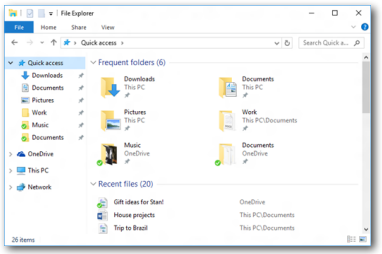
name of:

system utility `(cat, ls)`

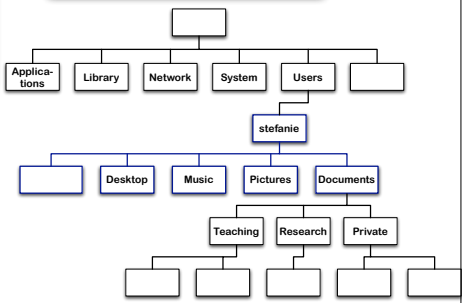
application program `(firefox, xed)`

File system

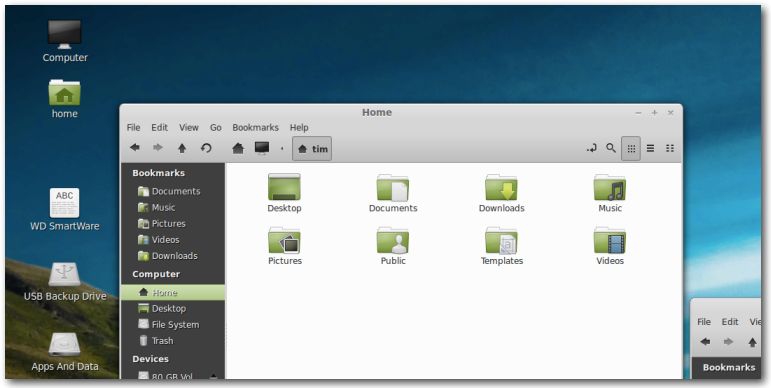
Windows OS



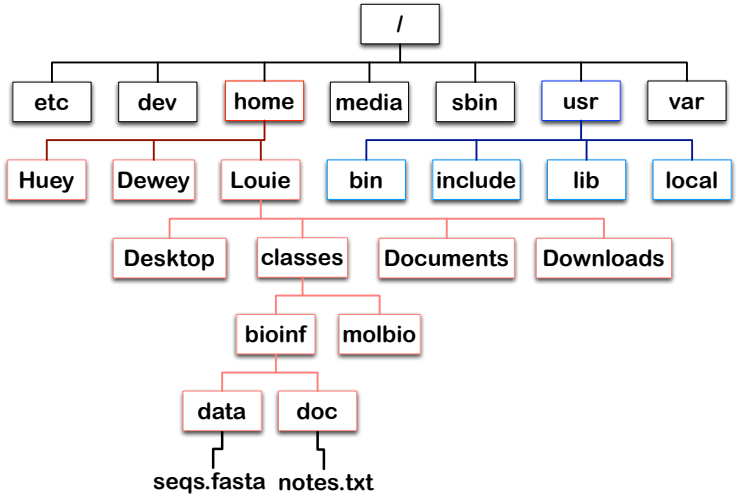
Mac OS



File system: Linux



File system: Linux



Working in the terminal

working with files
- create directories
- move directories and files
- copy directories and files
- delete directories and files
- traverse the file hierarchy to access directories and files

working with file content
- inspecting content
- searching through file content
- extracting file content

System utilities

small programs that
- each do a single task extremely well
- (some) can be used in sequence (“>”, “|”)

system utilities that operate on
- files and directories
  - `pwd, ls, cd, cp, rm, mv, mkdir`
- data (text) within files
  - `less, cat, head, tail`
  - `wc, sort, uniq, grep, ...`

file: plantSpecies.txt

```
Solanum lycopersicum (tomato)
Arabidopsis thaliana (mustard weed)
Helianthus annuus (sunflower)
Helianthus annuus (sunflower)
Beta vulgaris (sugar beet)
Arabidopsis thaliana (mustard weed)
Solanum tuberosum (potato)
Solanum lycopersicum (tomato)
Arabidopsis thaliana (mustard weed)
```


input

```
cat plantSpecies.txt
```

output

```
Solanum lycopersicum (tomato)
Arabidopsis thaliana (mustard weed)
Helianthus annuus (sunflower)
Helianthus annuus (sunflower)
Beta vulgaris (sugar beet)
Arabidopsis thaliana (mustard weed)
Solanum tuberosum (potato)
Solanum lycopersicum (tomato)
Arabidopsis thaliana (mustard weed)
```


input

```
head -n 5 plantSpecies.txt
```

output

```
Solanum lycopersicum (tomato)
Arabidopsis thaliana (mustard weed)
Helianthus annuus (sunflower)
Helianthus annuus (sunflower)
Beta vulgaris (sugar beet)
```


input

```
sort plantSpecies.txt
```

output

```
Arabidopsis thaliana (mustard weed)
Arabidopsis thaliana (mustard weed)
Arabidopsis thaliana (mustard weed)
Beta vulgaris (sugar beet)
Helianthus annuus (sunflower)
Helianthus annuus (sunflower)
Solanum lycopersicum (tomato)
Solanum lycopersicum (tomato)
Solanum tuberosum (potato)
```

input `sort plantSpecies.txt | uniq`

output

```
Arabidopsis thaliana (mustard weed)
Beta vulgaris (sugar beet)
Helianthus annuus (sunflower)
Solanum lycopersicum (tomato)
Solanum tuberosum (potato)
```


input `sort plantSpecies.txt | uniq -c`

output

```
3 Arabidopsis thaliana (mustard weed)
1 Beta vulgaris (sugar beet)
2 Helianthus annuus (sunflower)
2 Solanum lycopersicum (tomato)
1 Solanum tuberosum (potato)
```


input `sort plantSpecies.txt | uniq > myresult.txt`

output

Key terms and concepts

- linux OS, advantages
- terminal / command line interface
- file system (definition, principle)
- absolute vs. relative paths
- system utilities
  - pwd, ls, cd, cp, mv, rm, mkdir, ..
  - wc, less, cat, head, tail
  - sort, uniq
  - >, |