

# ITBC BI1 Python 1 en Basiskennis voor Bioinformatici

## Afvinkopdracht 1

---

## Afvinkopdracht 1

### Inhoudsopgave

- Achtergrond
- Opdracht 1: Oefenen met Linux
- Opdracht 2: Omgaan met Linux
  - Deel 1
  - Deel 2
  - Deel 3

## Achtergrond

In deze opdracht gaan we aan de slag met het computer besturingssysteem Linux. De meesten van jullie zijn waarschijnlijk tot nu toe omgegaan met Windows of MacOS om gebruikt te maken van een computer. Linux is een verzamelnaam voor een grote groep diverse besturingssystemen (distributies) die allemaal op dezelfde basis (kernel) draaien. Van een kernel bestaan een hoop verschillende versies, waar iedere Linuxdistributie op een andere manier gebruik van kan maken.

Het voordeel van al deze distributies is dat er voor iedere toepassing een special besturingssysteem is. Er is bijvoorbeeld een speciale Bio-informatica Linuxdistributie, BioLinux, welke weer gebaseert is op de distributie Ubuntu. Het nadeel is dat verschillende distributies soms net op een iets andere manier werken, dus dat er soms van je gevraagd wordt dat je wat meer ervaring met je besturingssysteem moet hebben dan wanneer je iets in Windows voor elkaar zou willen krijgen.

Gedurende dit blok maken wij gebruik van Ubuntu, een van de meest gebruikte distributies voor beginners, maar voel je vrij om je eigen keuze te maken als je wat meer vertrouwd bent geraakt met Linux.

## Opdracht 1: Oefenen met Linux

In de weektaak heb je de opdracht gekregen om Linux te installeren op je laptop. Het is natuurlijk ook handig om te weten hoe je om moet gaan met Linux. Ubuntu heeft zoals je ziet een grafische weergave met een desktop, net als Windows en MacOS, maar niet alle Linux distributies hebben dit. Veel Linux servers hebben bijvoorbeeld alleen maar een tekstuele weergave, die we de command line noemen. Hierbij kan je niet gebruik maken van een muis om te klikken, dus moet je weten hoe je door je besturingssysteem kunt navigeren zonder kliks. In deze opdracht gaan we gebruik maken van de opgaven in de Linux handleiding om te oefenen.

Volg de instructies in de handleiding en geef antwoord op de vragen: BIN-

owe1\_LinuxVoorBioInformatici.pdf Als je vragen hebt kan je uiteraard de hulp inschakelen van de (praktijk)docent, maar vraag gerust aan je buurman of -vrouw of zij je een handje verder kunnen helpen. Schrijf de antwoorden op in het antwoordmodel (BIN-owe1\_LvBI\_antwoordmodel.docx) en toon deze aan de docent.

## Opdracht 2: Omgaan met Linux

Nu je wat oefening gehad hebt met de verschillende Linux commando's kan je individueel aan de slag met de volgende opdracht. Gebruik het antwoordvel (BIN-OWE1\_Afv1\_antwoordmodel.docx) om je antwoorden op te schrijven en toon de resultaten aan de docent zodat deze je opdracht kan aftekenen.

### Deel 1

Log in op een Linux computer (je eigen of een van de desktops in lokaal 1.12) en beantwoord de volgende vragen. Iedere vraag kan beantwoord worden met behulp van een Linux commando.

1. Wat is je huidige locatie?
2. Toon de inhoud van je huidige locatie.
3. Maak een nieuwe directory aan en noem deze *afvink1*.
4. Wat zijn de rechten van de directory *afvink1*?
5. Zonder te verplaatsen, wat is de inhoud van *afvink1*?
6. Verplaats naar de nieuwe map.
7. Maak een bestand aan en noem deze *bestand.txt*.
8. Maak een kopie van het nieuwe bestand en noem deze *kopie.txt*.
9. Wat zijn de rechten van beide bestanden?
10. Maak *bestand.txt* alleen voor jou leesbaar.
11. Maak *kopie.txt* voor iedereen lees- en schrijfbaar.
12. Toon alle bestanden waarvan de tweede letter een o is.
13. Keer terug naar je home directory

### Deel 2

Nu gaan we op een remote server inloggen, onze eigen server cytosine.nl. Deze server draait ook op Ubuntu en verschilt uiteindelijk niet zoveel van een desktop, alleen heeft de server geen grafische weergave. Het leuke van een server is dat je bepaalde directories kunt open stellen zodat je er via een browser ook op kunt kijken, een website. Er zijn twee manieren waarop je verbinding kan maken met cytosine.nl, vanuit Linux of vanuit Windows. Je login gegevens zijn:

- username: owe1\_pg# (waarbij je # vervangt met je projectgroepnummer)
- wachtwoord: blaas1234
- hostname: cytosine.nl

Als je vanuit Linux verbinding maakt typ je op de commandline `ssh username@hostname`. Vanuit Windows inloggen doe je met behulp van PuTTY. Dit programma is gratis beschikbaar op [www.putty.org](http://www.putty.org). Geef ook hier weer de antwoorden in de vorm van een commando.

1. Log in op cytosine.nl
2. Controleer of er een directory *public\_html* is.
3. Maak een subdirectory aan met jouw naam onder *public\_html*.

4. Verplaats naar deze nieuwe subdirectory.
5. Maak een bestand *readme.html*.
  - Controleer met behulp van een browser dat dit bestand er ook is. Ga naar [http://cytosine.nl/~owe1\\_pg#/jouwnaam](http://cytosine.nl/~owe1_pg#/jouwnaam)
  - Vervang # met je projectgroepnummer en jouwnaam met de naam van de directory die je aangemaakt hebt.
  - Zie je het bestand *readme.html* staan?
6. Geef iedereen leesrechten op *readme.html*.
7. Toon de naam van de actuele directory.
8. Maak een bestand *private.txt*.
9. Zorg dat jij alleen leesrechten hebt op *private.txt*.
  - Controleer met behulp van de browser dat je inderdaad het bestand *private.txt* niet kunt uitlezen.
10. Wijzig de inhoud van *readme.html* naar de inhoud van het volgende bestand: *readme.html*
11. Log uit.

## Deel 3

Grep is een krachtig middel om te zoeken naar bijvoorbeeld tekstbestanden, of een woord of zinnen in een document. We gaan in een FASTA file zoeken naar patronen in DNA met behulp van Grep en regular expressions. Het is nu nog niet van belang dat je de inhoud van het FASTA bestand begrijpt of weet wat de verschillende patronen in DNA betekenen. Het doel van deze opdracht is dat je inzicht krijgt in hoe je met grep snel door een (groot) bestand heen kan zoeken naar zinnen met bepaalde woorden of woorddelen er in. Op een later moment komt DNA uitgebreider aan bod.

Een FASTA bestand is te herkennen aan de .fasta, .fsa of .fa file extensie en aan de inhoud. Het bestaat uit de combinatie van een beschrijving, de regel die met een > begint, en een sequentie, de regels met heel veel hoofdletters A, T, G, C en N. Het FASTA bestand voor deze opdracht is een samenstelling van DNA sequenties van de mens, bakkersgist, de muis en de alpaca.

Werk op een Linux computer. Download het FASTA bestand via <http://cytosine.nl/~esther/downloads> en sla het bestand op in de eerder gemaakte *afvink1* directory. Raadpleeg de grep manual als je niet precies weet hoe je grep moet gebruiken.

1. Benader de folder waar je het bestand in opgeslagen hebt.
2. Gebruik grep om regels te tonen uit het fastabestand waar het startcodon (ATG) in voorkomt.
3. Gebruik grep om regels te tonen uit het fastabestand waar een van de stopcodons (TAG/TGA/TAA) in voorkomen.
4. Gebruik grep om regels te tonen uit het fastabestand waar een TATA box in voorkomt, dit is een herhaling van het patroon TATA een of meerdere keren.
5. Gebruik grep om regels te tonen uit het fastabestand waar een header in in voorkomt, dus alle regels die met > beginnen.