

Mintaillesztés és keresés Linuxban

Linux alapok 2019

Varga Tibi

Linux grep parancs család

A grep parancs családjába tartozik, a grep, az egrep és az fgrep parancsok amelyet a minták keresésére használnak Linux alatt.

A grep parancs karakterláncokban képes mintákat illeszteni az egrep viszon már kiterjesztett reguláris kifejezésekkel is képes dolgozni ezért modernebb környezet ben ezt használjuk. Az egrep ugyanúgy működik, mint a grep -E (grep Extended regex).

Az **egrep** átvizsgál egy adott fájlt, sorról sorra, és kiírja azokat a sorokat, amelyek tartalmazzák a keresési karakterláncot / reguláris kifejezést.

Az egrep parancs szintaxisa:

egrep [opciók] 'MINTA' fájl(ok)

Az egrep minta illesztés

1. Egy adott karakterlánc keresése egy fájlban

Ez az egrep parancs leggyakoribb használata. megadjuk a keresni kívánt karakterláncot és a fájlnevet, amelyben meg akarjuk keresni a karakterláncot.

Az eredmény a keresett karakterláncot tartalmazó teljes sort.

Ebben a példában a „deb” szórészletre kerestem a megadott szövegfájlban. Láthatjuk, hogy az eredmény hogyan jeleníti meg a „deb” szót tartalmazó teljes sort:

```
root@kali:~# egrep deb minta.txt
debian9
debian
```

```
root@kali:~# cat minta.txt
alma
Alma
Debian
Debian9
debian9
debian
körte
mama
lala
lala 22
irisz
Agi
agika
almafa
laci
lali
ili
lili
bili
Bili
Lilla
peti
alom
ólom
```

Kapcsolók az egrep parancshoz

A minta illesztés kifejezetten érzékeny a kis és nagy betűkre ezért az -i kapcsolóvak figyelmen kívül hagyható – lásd a második parancsnál.

```
# egrep -i 'deb' minta.txt
```

```
root@kali:~# egrep deb minta.txt
debian9
debian
root@kali:~# egrep -i deb minta.txt
Debian
Debian9
debian9
debian
```

Kapcsolók:

- **# egrep -c deb minta.txt**

A **-c** kapcsoló az illeszkedő sorok tartalma helyett csak azok darabszáma jelenik meg.

```
root@kali:~# egrep -c deb minta.txt
2
root@kali:~# # kombinálva az -i vel
root@kali:~# egrep -i -c deb minta.txt
4
```

- **# egrep -v deb minta .txt**

A **-v** kapcsolóval azon sorait írja ki, amik nem tartalmazzák a minta karakterláncot.

```
root@kali:~# egrep -i -v deb minta.txt
alma
Alma
körte
mama
lala
lala 22
irisz
Agi
agika
almafa
laci
```

- **# egrep -e 'deb' -e 'alma' minta.txt**

Az -e minta : minta megadása; segítségével több minta is megadható, illetve akkor is jó, ha a minta a karakterrel kezdődik. Ez felfogható egy vagy opciónak is.

```
root@kali:~# egrep -e deb -e alma minta.txt
alma
debian9
debian
almafa
```

Hasonló eredményt hoz egy másik megoldás:

- **# egrep "minta1|minta2" fájl**

```
root@kali:~# egrep -i "alma|lili" minta.txt
alma
Alma
almafa
lili
```

Az n sor megjelenítése a keresési karakterlánc előtt, után vagy körül

Néha nagyon fontos az előtte, utána és körülötte meghatározott számú sort is megjelenítheti.

- -A $n : n$ db sor kiírása az illeszkedő sorok után.
- -B $n : n$ db sor kiírása az illeszkedő sorok előtt.
- -C $n : n$ db sor kiírása az illeszkedő sorok előtt és után.

- Karakterlánc keresése teljes szóként, és nem rész-karaktorsorozatként
- **# egrep -w [minta] fájl**
- Amikor normál esetben egy sztringet keresünk az egrep-en keresztül, az a karaktorsorozatot tartalmazó összes szót al-karaktorsorozatként kiírja, a **-w** kapcsolóval az mintát szóként kezeli.

```
root@kali:~# egrep alma minta.txt
alma
almafa
root@kali:~# egrep -w alma minta.txt
alma
```

egrep -n [minta] fájl

Kiírja a talált egyezés elé a sorszámot is

2. Egy adott karakterlánc keresése több fájlban

Az egrep paranccsal kereshetünk egy karakterláncot több, ugyanazon könyvtárban található fájl között. Csak egy kicsit pontosabban kell meghatározni a keresett fájlok „mintázatát”

Az eredmény az előfordulási fájl neve és a keresett karakterláncot tartalmazó teljes sort.

egrep [MINTA] fájl1 fájl2 vagy # egrep [minta] *.kiterjesztés

```
root@kali:~# egrep alma fruits.txt minta.txt
fruits.txt:alma
minta.txt:alma
minta.txt:almafa
root@kali:~# egrep alma *.txt
fruits.txt:alma
minta.txt:alma
minta.txt:almafa
szoveg2.txt:alma
szoveg3.txt:alma
```

Csak fájlnevek kiírása amelyek a minta karakterláncot tartalmazzál

- Előfordul, hogy csak egy adott karakterláncot tartalmazó fájlneveket akarunk beolvasni, nem pedig az azt tartalmazó sorokat kinyomtatni. Ez megtehető az egrep paranccsal, a -l (kisbetűs L) kapcsolóval.
- Szintaxis: **# egrep -l „minta” fájlnev**

```
root@kali:~# egrep -l alma *.txt
fruits.txt
minta.txt
szoveg2.txt
szoveg3.txt
```

3. Rekurzív keresés a karakterláncban a teljes könyvtárban és alkönyvtárakban

Ha egy könyvtárból és annak alkönyvtáraiból az összes fájlban szeretnénk karakterláncot keresni, akkor ezt megtehetjük a -r kapcsoló használatával (rekurzív keresés).

egrep -r [MINTA] *

```
root@kali:~# egrep -r lali *  
gyak1/minta:lali  
minta.txt:lali
```

A reguláris kifejezés

A reguláris kifejezés (rövidítve regex vagy) egy karakterlánc, amely meghatározza a keresési mintát . Általában az ilyen minták által használt karakterlánc-keresési algoritmusok, a keresés és csere vagy beviteli ellenőrzés.

A minták megadásánál sorozatok, logikai értékek, számlálás, kezdő és vég pontokat adhatunk meg és ezeket a mintákba illesztett méta karakterekkel valósítjuk meg:

**A szokásos méta karakterek a: {} [] () ^ \$. | * + ? **

A reguláris kifejezés

Általános karakter "C" : Bármely közönséges karakter illeszkedik saját magára.

Speciális karakterek:

- ? - pontosan egy karaktert helyettesít: pl. ?Ima lehet alma vagy álma is így.
- * - bármennyi (akár 0) karaktert helyettesít: pl. *gép lehet gép, mosógép, számítógép, stb.
- [...] - a [] között felsorolt karakterekből pontosan egyet helyettesít: pl. [aó]lom lehet ólom vagy alom, fajl[0123456789] pedig fajl0, fajl1, ... fajl9.
- [^...] - pontosan egy karaktert helyettesít, mely nem lehet a [] között felsorolt karakterek egyike sem.

- `^` : A mintát a sor elejére igazítja. Csak azok a sorok illeszkednek rá, amelyek a `^` jel utáni reguláris kifejezésre illeszkedő szóval kezdődnek.
- `$` : Ugyanazt csinálja, mint az előző, annyi különbséggel, hogy a mintát a sor végére igazítja.
- `.` : Az új sor kivételével minden karakter illeszkedik rá.
- `\c` : Kikapcsolja a `c` speciális karakter speciális (méta karakter) jelentését. Akkor használatos, ha történetesen épp speciális karaktereket szeretnénk keresni.
- `[n-n]` : A megadott tartományon belül eső karakterek bármelyike illeszkedik rá.

Minta illesztés sor elejére és végére

Soreleji minta horgony

- # egrep "^[minta]" fájl

Pl.: minden „i” vel kezdődő sor kiíratása

```
root@kali:~# egrep -i "^[i]" minta.txt
irisz
ili
```

Sorvégi minta horgony

egrep "[minta\$]" fájl

Pl.: minden „i” vel végződő sor kiíratása

```
root@kali:~# egrep -i "[i]$" minta.txt
Agi
laci
lali
ili
lili
bili
Bili
peti
Jani
```


Több minta együttes teljesülése, az ,és' kapcsolat a minták között

Ilyenkor az egrep parancs kimentének átirányítását végezzük a második minta feltételt tartalmazó egrep parancs bemenetére.

- **# egrep [minta1] fájl | egrep [minta2]**

Pl.: Minden („i”-re végződő és „l” –vel kezdődő sor kiírás a minta.txt fájlból (az – i kapcsoló miatt figyelmen kívül marad a kis és nagy betű)

```
root@kali:~# egrep -i "[i]$" minta.txt | egrep -i "^[l]"
laci
lali
lili
lali
Légrádi
```

A pontok „.”

A pontok darabonként egy bármilyen karaktert jelenthetnek de semmit nem, csak más karakterek között használjuk.

- **# egrep [a..a] fájl**

Példa: ha egy olyan mintát keresünk amiben van kettő „a” betű és közöttük valamilyen egy darab karakter

```
root@kali:~# egrep -i "a.a" minta.txt
mama
lala
lala 22
almafa
.valami
apa
ara
```

Példa: ha egy olyan mintát keresünk amiben van kettő „a” betű és közöttük valamilyen két darab karakter

```
root@kali:~# egrep -i "a..a" minta.txt
alma
Alma
almafa
banya
sanya
barna
arra
```

Speciális karakterek kikapcsolása

- Előfordul hogy a keresendő mintában valamelyik speciális (méta) karakter szerepel. Ilyenkor a " \" –sel (vissza-per -- backslash) jelezzük az értelmezőnek hogy a méta karakter csak sima sztring.

Példa: ha olyan sorokat
keresünk amelyekben
kérdőjel szerepel

```
root@kali:~# egrep -i "\?" minta.txt  
létezik?  
érted?
```

Példa: ha olyan sorokat
keresünk amelyekben
pont jel szerepel

```
root@kali:~# egrep -i "\." minta.txt  
.valami  
.ez egy pont a szövegben
```

Karakter részek [] közötti viselkedése a mintában

Bármely „a” vagy „b” a minta előtt illeszkedik

- # egrep "[ab]minta" fájl

```
root@kali:~# egrep -i "[tf]arka" minta.txt  
tarka  
farka
```

Bármely „a”-tól „z”-ig a minta előtt illeszkedik

- # egrep "[a-z]minta" fájl

```
root@kali:~# egrep -i "[a-t]nya" minta.txt  
banya  
sanya
```

Példák

- Az össze s a betűt tartalmazó fájl listázása:

```
tibi@server:~/szures$ ls -1 | egrep a
abba
agi
ali
apa
baab
bab
baba
gabi
lali
mami
pap
papa
papi
```

Példa

Az összes betűvel kezdődő majd az összes **a** betűvel végződő fájl listázása :

```
tibi@server:~/szures$ ls -1 | egrep '^a'
abba
agi
ali
apa
tibi@server:~/szures$ ls -1 | egrep 'a$'
abba
apa
baba
papa
```

Példa

Az összes **a** betűvel kezdődő és **a** betűvel végződő fájl listázása :

```
tibi@server:~/szures$ ls -1 | egrep '^a' | egrep 'a$'  
abba  
apa
```

Az olyan fájlok listázása melyek **a** betűvel kezdődnek **i** betűvel végződnek és közöttük **egy** olyan kis betű van mely **a-tól l-ig** helyezkedik el az ABC-ben:

```
tibi@server:~/szures$ ls -1 | egrep '^a[a-l]i$'  
agi  
ali
```

Példa

Az összes **a** betűvel kezdődő és **a** betűvel végződő és közöttük **egy** bármilyen karakter van - fájlok listázása :

```
tibi@server:~/szures$ ls -1 | egrep '^a.a$'
```

apa

Az összes **a** betűvel kezdődő és **a** betűvel végződő és közöttük **kettő** bármilyen karakter van - fájlok listázása :

```
tibi@server:~/szures$ ls -1 | egrep 'a..a$'
```

abba

A . (pont) karakter látja el a helyettesítést.

Példa

Az összes **a** vagy **b** betűvel kezdődő fájl listázása :

```
tibi@server:~/szures$ ls -1 | egrep '^[ab]'
```

abba
agi
ali
apa
baab
bab
baba

Példa

Az egrep parancs a **-c** kapcsolóval megszámolja hogy a könyvtárban **hány** olyan fájl van amelyik **b** betűvel kezdődik.

```
tibi@server:~/szures$ ls -1 | egrep -c '^b'  
3
```

Az egrep parancs a **-c** kapcsolóval megszámolja hogy a elso.log fájlban **hány olyan sor** van amelyik **t** betűvel kezdődik.

```
tibi@server:~/Letöltések$ egrep -c '^t' elso.log  
5
```

Példa

- Szöveg vagy szöveg részlet keresése fájlban **egrep** „szöveg” file.kit

```
tibi@server:~/Letöltések$ egrep "hat" elso.log
```

```
hatodik
```

```
tibi@server:~/Letöltések$ cat elso.log | egrep "het"
```

```
hetedik
```

```
tibi@server:~/Letöltések$ egrep -i "tiz" elso.log
```

```
tizedik
```

```
tizenegyedik
```

```
tizenkettedik
```

```
tizenharmadik
```

```
tizennegyedik
```

Példa

Egyszerre több minta keresés az `-e` kapcsolóval:

- `egrep -e „hat” -e „het” elso.log`

```
tibi@server:~/Letöltések$ egrep -e "hat" -e "het" elso.log
hatodik
hetedik
```

Kivonó keresés `-v` kapcsolóval:

- `egrep -v „ik” elso.log`

```
tibi@server:~/Letöltések$ egrep -v "ik" elso.log
elso
utolso
```

Listázás példák

- Az összes **.txt** végződésű fájl kilistázása:

```
tibi@server:~$ ls -l *.txt
-rw-rw-r-- 1 tibi tibi  0 máj  22  2019 akarmi.txt
lrwxrwxrwx 1 tibi tibi 12 okt  23 20:19 feladat2.txt -> feladat1.txt
-rw-rw-r-- 1 tibi tibi 22 okt  23 20:32 feladat4.txt
-rw-r--r-- 1 tibi tibi  0 nov  12 20:25 hiba.txt
-rw-r--r-- 1 tibi tibi 371 jan  19 09:24 lista_nezet.txt
-rw-r--r-- 1 tibi tibi  0 jan  19 09:09 szia.txt
```

- Az összes **pdf** kiterjesztésű fájl másolása /home/tibi/Dokumentumok könyvtárba majd az összes pdf listázása

```
tibi@server:~/Dokumentumok$ cp *.pdf /home/tibi/Képek
```

```
tibi@server:~/Dokumentumok$ ls -l /home/tibi/Képek/*.pdf
-rw-r--r-- 1 tibi tibi 668695 jan  27 16:22 /home/tibi/Képek/alapok.pdf
-rw-r--r-- 1 tibi tibi 856657 jan  27 16:22 '/home/tibi/Képek/könyvtár .pdf'
-rw-r--r-- 1 tibi tibi 182958 jan  27 16:22 /home/tibi/Képek/környezeti.pdf
```

Példa

Az összes **[a á o]** kezdetű és **lom** –al végződő fájl nevű fájl listázása majd törlése a Dokumentumok mappából:

```
tibi@server:~/Dokumentumok$ ls -l [aoá]lom
-rw-r--r-- 1 tibi tibi 0 jan 27 16:56 alom
-rw-r--r-- 1 tibi tibi 0 jan 27 16:54 álom
-rw-r--r-- 1 tibi tibi 0 jan 27 16:54 olom
tibi@server:~/Dokumentumok$ rm [aoá]lom; ls -l
összesen 10272
-rw-r--r-- 1 tibi tibi 668695 szept 14 07:13 alapok.pdf
-rw-r--r-- 1 tibi tibi 1720320 nov 18 04:17 all.tar.gz
drwxr-xr-x 2 tibi tibi 4096 jan 14 08:47 gyak
-rw-r--r-- 1 tibi tibi 856657 okt 8 18:24 'könyvtár .pdf'
-rw-r--r-- 1 tibi tibi 182958 nov 15 17:31 környezeti.pdf
-rw-r--r-- 1 tibi tibi 1585189 nov 18 04:03 mind.tar.gz
-rw-r--r-- 1 tibi tibi 1720320 nov 18 03:40 taroltbdf.tar
-rw-r--r-- 1 tibi tibi 1557292 nov 18 03:40 taroltbdf.tar.gz
-rw-r--r-- 1 tibi tibi 618495 nov 18 04:00 tomar.tar.bz2
-rw-r--r-- 1 tibi tibi 1585189 nov 18 04:02 v
```

Példa

Az összes **számot** tartalmazó és **.test** kiterjesztésű fájlnev listázása:

```
tibi@server:~/Dokumentumok$ ls -l *[0-9]*.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 hat66.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 nyolc88ny.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 tiz10t.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 va12gg.test
```

Az összes **számot** nem tartalmazó és **.test** kiterjesztésű fájlnev listázása

```
tibi@server:~/Dokumentumok$ ls -l *.test
-rw-r--r-- 1 tibi tibi 0 jan 27 18:04 hat66ht.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 héthét.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 kilenck.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 nyolc88ny.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 tiz10t.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 tizenegy.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 va12gg.test
tibi@server:~/Dokumentumok$ ls -l *[^0-9]??*.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 héthét.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 kilenck.test
-rw-r--r-- 1 tibi tibi 0 jan 27 17:09 tizenegy.test
```

Fájlok neveinek illeszkedése, az ls és egrep parancsok kapcsolása

Tegyük fel, hogy még mindig az **ls** kimenetével szeretnénk dolgozni.

- Az alap könyvtártartalom:

```
tibi@server:~/szures$ ls -1
abba
agi
ali
apa
baab
bab
baba
flori
gabi
ili
irisz
lali
lo
mami
pap
papa
papi
peti
szilvi
szivi
```


Keresés Linux könyvtárakban

find [útvonal...] [kifejezés]

- Fájlokat keres egy könyvtárstruktúrában.
- A find parancs rengeteg kapcsolóval rendelkezik, emellett operátorokat is használhatunk vele.
- Részletesebben lásd man find.

Csak néhány egyszerűbb kapcsoló használata :

-name (sztringet vár)

-type (**d**: dir, **f**: fájl, **l**: link)

-size (lehet +/- majd a méret, attól függően, hogy annál nagyobb vagy kisebb méretűeket keresünk)

- Egyszerre több keresési feltételt is megadhatunk, ehhez azonban zárójeleket kell használnunk, amelyeket meg kell védenünk attól, hogy a shell saját belátása szerint értelmezze őket.
- A következő példa megkeresi az aktuális könyvtárban és az abból nyíló alkönyvtárakban található .c-re és .o-ra végződő nevű fájlokat.
- **# find . \(-name '/*.c' -o -name '/*.o' \) -print**

Példa

- Keressük meg a **.test** kiterjesztésű fájlokat:

find . -name "*.test"

```
tibi@server:~/Dokumentumok$ find . -name "*.test"
./hat66ht.test
./nyolc88ny.test
./tiz10t.test
./héthét.test
./kilenck.test
./va12gg.test
./tizenegy.test
```

- Hány ilyen fájl van?

find . -name "*.test" | wc -l

```
tibi@server:~/Dokumentumok$ find . -name "*.test" | wc -l
7
```

Példa

Az összes **.log** kiterjesztésű fájl keresése a **/tibi** könyvtárban (alkönyvtárokkal): **find /tibi -name "*.log"**

```
tibi@server:/home$ find ./tibi -name "*.log"
find: './tibi/.dbus': Engedély megtagadva
find: './tibi/.cache/dconf': Engedély megtagadva
./tibi/.cache/mozilla/firefox/1pssa1nh.default/cache2/index.log
./tibi/mod.log
./tibi/Képek/log.log
./tibi/.local/share/gvfs-metadata/smb-share:server=server.local,share=diak-a76
ea8.log
./tibi/.local/share/gvfs-metadata/root-aeccf055.log
./tibi/.local/share/gvfs-metadata/smb-server:server=server.local-130ac476.log
./tibi/.local/share/gvfs-metadata/admin:-078ba641.log
./tibi/.local/share/gvfs-metadata/home-1ad8f86e.log
./tibi/.local/share/gvfs-metadata/smb-network:-cbef1ed3.log
./tibi/.local/share/gvfs-metadata/uuid-AB7B-E8F5-3f21ee70.log
./tibi/.local/share/gvfs-metadata/trash:-2c205278.log
./tibi/.local/share/xorg/Xorg.0.log
./tibi/Letöltések/matrix.log
./tibi/Letöltések/elso.log
```

Példa

- Összes mappa keresése: **find . -type d**
- Hány mappa van a /Képek könyvtárban? **find . -type d | wc -l**

```
tibi@server:~/Képek$ find . -type d
.
./szürke
./színes
tibi@server:~/Képek$ find . -type d | wc -l
3
```

- Összes fájl keresése: **find . -type f**

```
tibi@server:~/Képek$ find . -type f
./Képernyőkép erről: 2019-09-16 02-55-37 - 1.png
./Képernyőkép erről: 2019-09-16 02-55-30.png
./környezeti.pdf
./Képernyőkép erről: 2019-09-16 02-55-37.png
./Képernyőkép erről: 2019-09-16 02-57-06.png
./Képernyőkép erről: 2019-09-16 02-55-31.png
./ls
./log.log
```

Fájlok keresése mérettartományban

- Az összes, 50 MB-nál nagyobb és 100 MB-nál kevesebb fájl megtalálásához.

```
$ find / -size + 50M -size -100M
```

Az üres fájlokat és könyvtárakat keresése

A következő parancs a `find` parancs "üres" opcióját használja, amely az összes üres fájlt megtalálja.

```
# find / tmp-type f -empty
```

Az összes üres könyvtárat a "d" gombbal írhatja be.

```
$ find ~ /-type d -empty
```

Fájlok és könyvtárak keresése a módosítási dátum és idő alapján

Az összes fájl, amelyet 50 nappal ezelőtt módosítottak.

```
# find / -mtime 50
```

Keresse meg az összes fájlt, amelyet az elmúlt 50 napban elértek.

```
# find / -atime 50
```

Keresse meg az összes fájlt, amelyet 50–100 nappal ezelőtt módosítottak.

```
# find / -mtime +50 -mtime -100
```

Keressen az elmúlt egy órában módosított fájlokat.

```
$ find / home / bob -cmin -60
```

Keresés és csere

Nagyon gyakori művelet, hogy egy szövegben valamilyen szót szeretnénk egy másikra cserélni. Ezt a leggyorsabban a **sed** nevű paranccsal hajthatjuk végre, az alábbi szintaxis szerint:

- **#sed ,y/mit/mire/g' hol > hova**

A fenti parancs végignézi a "hol" fájlt, kicseréli benne az összes "mit" szót "mire"-re és az eredményt a "hova" nevű fájlba (átírányítás jel és fájl név megadása nélkül a képernyőre) írja.

A sed egy nagytudású szó szerkesztő, de sajnos szinte lehetetlen kezelni, ezért csak a legelvetemültebb buherátoroknak javasolom, hogy parancsait elsajátítsák. A mindennapi életben elég, ha a fenti példát megjegyezzük, valamint azt, hogy "hol" és a "hova,,, de fájlként SOHA ne adjuk meg ugyanazt a nevet!!

Példa: ékezetes karakterek lecserélése szövegben

```
root@kali:~# cat csere.txt
```

éva
ági
délen
árpás
ézsiás
Édua
ágota
Ágota
Állat
Bélám

```
root@kali:~# sed -e 'y/éáÁÉ/eaAÉ/' csere.txt > uj.txt
```

```
root@kali:~# cat uj.txt
```

eva
agi
delen
arpas
ezsias
Édua
agota
Agota
Allat
Belam