



6. cvičení

Zpracování textu programem awk.





- awk = **A**ho–**W**einberger–**K**ernighan
 - vzory
 - akce
 - pole
 - funkce
 - rozšíření awk (nawk)





- Vypis prvních 10 řádků (head)

```
cat -n /etc/passwd | awk 'NR<=10'
```

```
cat -n /etc/passwd | awk 'NR==1,NR==10'
```

- Vypis posledních 10 řádků (tail)

```
awk NR\>=$( (`wc -l </etc/passwd` - 10) ) /etc/passwd
```

- Vypis řádků obsahujících RE (grep)

```
awk /root/ /etc/group
```

- Vypis řádků neobsahujících RE (grep -v)

```
awk '!/root/' /etc/group
```



- Výběr položek z `ypcat passwd` podle jmen

```
ypcat passwd \  
| awk '/Adela/ /Barbora/ /Celeste/ /Dana/ /Eliska/'
```

```
ypcat passwd \  
| awk -f /home/courses/Y36U0S/common/cv06/names.awk
```

- Z výstupu `ypcat passwd` výběr položek s UID 1000-9999

```
ypcat passwd | awk -F: '$3>=1000 && $3<=9999'
```

- Zjištění počtu uživatelů z výstupu `ypcat passwd`

```
ypcat passwd \  
| awk 'END { print "Total users: " NR }'
```



- Výběr a změna pořadí jednotlivých položek /etc/passwd

```
awk 'BEGIN { FS=":"; OFS=":" }  
      { print $3,$1,$5 }' /etc/passwd
```

- Nalezení uživatele s nejdelším jménem

```
yycat passwd | awk -F: '  
  {  
    if (length($5) > max) {  
      max=length($5); user=$1; name=$5  
    }  
  }  
  END {  
    print user ":" name " (" max ")"  
  }'
```



- Formátovaný výpis uživatelů Korn shellu

```
ypcat passwd | awk -F: '  
  BEGIN {  
    print "Uzivatele Korn shellu:"  
    for (d=0; d<=60; d++) printf "-"; printf "\n"  
  }  
  $7 ~ /ksh/ {  
    if (length($1)<=7) tab="\t\t"; else tab="\t"  
    print $1 tab $5  
    ksh++  
  }  
  END {  
    for (d=0; d<=60; d++) printf "-"; printf "\n"  
    print "Celkem " ksh " z " NR " uzivatelu"  
  }'
```



- Vypis uživatelů ve skupině x36alg

```
ypcat group | nawk -F: '{
  /^x36alg/ {
    items=split($4,a,",")
    for (i=1; i<=items; i++) print a[i]
  }'
```





- Frekvenční slovník

```
man man | awk '
  BEGIN { FS="[a-zA-Z]" }

  { for (i=1;i<=NF;i++) words[tolower($i)]++ }

  END {
    # prázdné položky nepočítáme
    delete words[""]
    for (i in words) {
      sum++
      print words[i],i
    }
    print "V textu bylo " sum " různ"ch slov"
  }' | sort -n
```




awk – uživatelské funkce +

ls / | nawk ' • Náhodné pořadí vstupních řádků

```
function swap(x,y) {
    tmp=item[x]
    item[x]=item[y]
    item[y]=tmp
}

{ item[NR]=$0 }

END {
    srand();
    for (i=1; i<=NR; i++) {
        RAND=int(NR*rand())+1
        swap(i,RAND)
    }
    for (i=1; i<=NR; i++) { print item[i] }
}
```



Příprava na příští cvičení

- Vyjděte z minulé přípravy...
- Ke každému řádku souboru seznam.csv přidejte položku obsahující uživatelské jméno uvedeného studenta.
- Vytvořte HTML stránku obsahující tabulku se všemi údaji obsaženými v souboru seznam.csv (pokud možno obecně, nezávisle na počtu řádků i položek v souboru).

