

UNIXi/LINUXi kasutusjuhend keeleteadlastele

Sisselogimine

Masinasse adalberg.ut.ee sisselogimiseks kasuta programmi SSH Secure Shell Client, Putty või WinSCP.

Puttyt kasutades:

Puttyt saab kasutada otse ülikooli it-osakonna veebilehelt, ei nõua eraldi installimist:

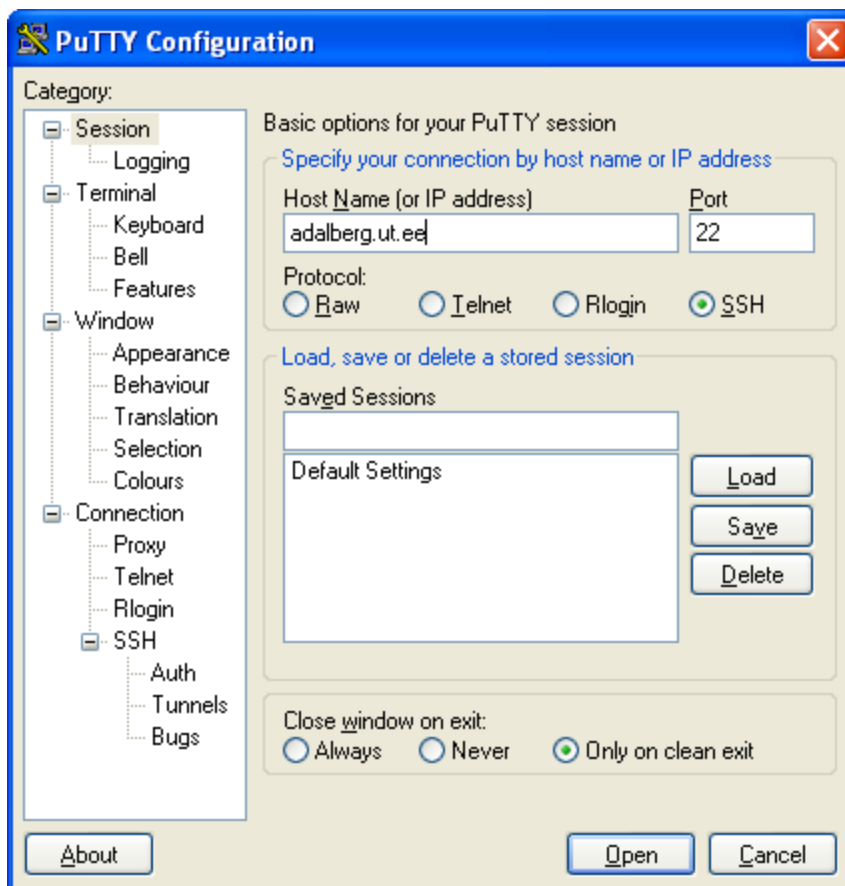
<http://www.ut.ee/4324>

WinSCP kasutusjuhend:

<https://wiki.ut.ee/display/AA/Failid>

Sisesta masina nimi (host name): adalberg.ut.ee

vali protokolliks SSH



Seejärel küsitakse kasutajanime ja parooli. Need on samad, mis sul ülikooli võrgus üldiselt (nt ÕIsi logides).

SSH-d kasutades:

ka programmi SSH Secure Shell Client leiate ülikooli IT-osakonna veebilehelt. See tuleb installida oma masinasse. Kasutamiseks:

- vajuta Quick Connect

- sisesta masina nimi (host name): adalberg.ut.ee

- oma kasutajanimi (User name), see on kasutajanimi ülikooli arvutivõrgus (sama, mis nt ÕISI sisenemisel)

- vajuta Connect.

Seejärel küsitakse sult parooli (sama, mis ÕISI sisenemisel).

Kui lõpetate töö ja lahkute arvuti juurest, tuleb end kindlasti ka serverist välja logida. Selleks kirjutage käsureale `exit`.

Kodeeringust

Täpitähtede jms nägemine ja sisestamine sõltub masinas kasutatavast kodeeringust. Tänapäeval on kõige tavalisemaks kodeeringuks, mida masinad kasutavad, UTF-8, see on kasutusel ka Tartu ülikooli serverites. Puttyga sisselogimise järel näetegi õpetust, kuidas masinasse sisselogimise järel oma Putty terminaliaken seadistada nii, et see näitaks serveri pilti korrektselt. Oletuslikult kasutab Putty kodeeringut ISO-8859-1, mis on varasem standard. Ka ssh ei ole seadistatud UTF-8 peale, nii et täpitähtede sisestamise ja nägemise puhul võib olla probleeme.

UTF-8 on ilmselt ka korpuste tulevik, ent praegusel hetkel on vähemalt murdekorpuses kasutatud varasem kodeeringustandard ISO-885915. Selleks, et server, kasutajaprogramm Putty või SSH ja korpused omavahel kenasti läbi saaks, on mõttekas kasutada seega ISO standardit. Selleks tuleks sisselogimise järel anda masinale käsk
`setenv LC_ALL et_EE.iso885915`

Selle käsu sooritamise järel kasutab ka server kodeeringut ISO-8859-15 ja peaks korrektselt aru saama sisestatud täpitähtedest ja korpuses kasutatud täpitähtedest.

Seadistus kehtib kuni seansi lõpuni, st kuni väljalogimiseni.

Töö failidega

Unixi/Linuxi shelli kasutades unusta ära, et arvuti küljes on hiir, sellega pole siin midagi teha! Käske saab anda ainult neid kirjutades. Selleks peab neid aga teadma. Siin dokumendis on toodud käsud, mida on vaja lihtsaks tööks tekstifailidega.

Rohkem infot iga käsu kohta:

käsk `--h` (näiteks `cd --h`)

Põhjalikum juhend: man käsk (nt `man uniq`)

Unixi kohta on internetis väga palju kasutusjuhendeid, näiteks
http://www.msg.ucsf.edu/local/programs/ono/manuals/unix_for_beginners.html
Enamik neist siiski pole mõeldud päris lingvistile, seepärast olen siia koondanud just keeleteadlastele vajaminevate käskude juhendeid.

Liikumine kataloogides, kataloogi sisu vaatamine

Unixis /Linuxis paiknevad failid kataloogides nagu muudiski operatsioonisüsteemides. Kataloogipuud saavad alguse nõ juurikast. Näiteks paikneb murdekorpuse väike õppekorpus kataloogipuus järgmiselt:

```
/home/mustikas/k/kriztel5/korpuste_kursus
```

/ kõige ees tähendab, et kataloogipuud jälgitakse nõ juurest alates (absoluutne tee) murakas on kataloog, milles on väga suur hulk ülikooli kasutajate kodukatalooge kriztel5 on Kristel Uihoaedi isiklik kodukataloog, normaalselt (kui ma pole lubanud teisiti) ei ole teistel sellele ligipääsu.

korpuste_kursus on Kristeli kodukataloogis olev kataloog, millele on selle kursuse jaoks antud ligipääsuõiguse ka teistele inimestele.

Seda, kus kataloogis sa parasjagu paikned, näeb käsuga pwd.
Kataloogide vahel liikumiseks on käsk cd.

See, kus kataloogis parasjagu oled, ilmneb ka käsurealt (näites punasega märgitud). Järgnevas näites ollakse kriztel5 kodukataloogi alamkataloogi kirjkorpus alamkataloogis 1930_ilu_ttxt.

```
[125] kriztel5@adalberg:~/kirjkorpus/1930_ilu_ttxt>
```

Siia taha kirjutatakse käsud.

Harjuta:

logi end sisse masinasse adalberg.ut.ee
selgita välja, kus paikneb sinu kataloog (käsk pwd; väga oluline edaspidi teada!)

mkdir nimi	kataloogi <i>nimi</i> loomine
cd nimi	kataloogi vahetamine (<i>change directory</i>), liigud kataloogi <i>nimi</i>
cd ..	liigud kataloogipuus ühe sammu võrra tagasi
cd	liigud oma kodukataloogi
pwd	näitab, kus kataloogis oled
ls	sirvib faile (lühike versioon)
ls -l	pikk versioon, annab järgmist infot: õigused omanik grupp-suurus muutmisaeg nimi

Vihje:

Failinimede puhul on võimalik osa tekstist asendada *-ga. * tähistab üht või mitut sümbolit, mis pole määratletud, st märgib, et failinimes võib selles osas olla mis iganes. See on eriti kasulik, kui ühe käsuga tahetakse manipuleerida mitme failiga. Kui tahetakse liigutada, kopeerida, kustutada ainult üht faili, siis on mõistlikum kogu failinimi välja kirjutada.

Näiteks käsuga `cp *.txt murre`
kirjutatakse kõik .txt-lõpulised failid kataloogi murre.

`mv KOD* murre` liigutatakse kõik failid, mille nimi algab KOD-iga, kataloogi murre.

Murdekorpuses on failinimed üldjuhul organiseeritud nii, et failinime alguses on suurte tähtedega lühend kihelkonnast, millest see tekst pärit on.

Järgmisel ekraanipildil on näidatud, real [141] kuidas kõigepealt liigutakse kataloogi murdekorpus alamkataloogi Idamurre; [142] sealt sammu võrra tagasi (kataloog murdekorpus); [143] sama kataloogi alamkataloogi Vorumurre; [144] vaadatakse kataloogi Vorumurre sisu; [145] liigutakse oma kodukataloogi.

```
[141] liina_l@adalberg:~> cd /home/liina_l/murdekorpust/Idamurre/
[142] liina_l@adalberg:~murdekorpust/Idamurre> cd ..
[143] liina_l@adalberg:~liina_l/murdekorpust> cd Vorumurre/
[144] liina_l@adalberg:~murdekorpust/Vorumurre> ls
~$R_Emmi_Nurm_synt.txt*      VAS_Ann_Kommer_synt.txt*
HAR_Ella_Uibu_synt.txt*     VAS_Emilie_Saarniit_synt.txt*
HAR_Emmi_Nurm_synt.txt*     VAS_Harald_Holter_synt.txt*
HAR_Hella_ja_Julius_Kokk_synt.txt* VAS_Juula_Pilt_synt.txt*
HAR_Minna_Hanimagi_synt.txt* VAS_Juuli_Pilt_synt.txt*
PLV_August_Katai_synt.txt*  VAS_Leena_Pilt_synt.txt*
q.txt                      VAS_Mari_Kartsepp_synt.txt*
RAP_Loti_Ritsmaa_synt.txt*  VAS_Miili_Virm_synt.txt*
URV_Kaarli_ja_Liisa_Myrk_synt.txt* VAS_Miili_Virm_syntl.txt*
VAS_Aino_Runthal_synt.txt*  VAS_Paul_Ots_synt.txt*
[145] liina_l@adalberg:~murdekorpust/Vorumurre> cd
[146] liina_l@adalberg:~>
```

Connected to adalberg.ut.ee SSH2 - aes128-cbc - hmac-md5 - none 80x24

Järgmises näites on kasutatud pikemat failide loendamise käsku `ls -l`, mis annab lisaks kataloogis olevate failide ja alamkataloogide nimedele ka infot õiguste, omaniku, grupi, faili suuruse ja viimase muutmisaja kohta.

Näide:

```
[103] kriztel5@adalberg:~/korpuste_kursus/murdekorporus> ls -l
```

```
total 64
```

õigused	omaja	grupp	faili suurus	viimane muutmisaeg	faili/kataloogi nimi
drwxr-xr--	2	liina_1	users	512 dets 3 2007	bin/
drwxr-xr--	2	liina_1	users	512 dets 3 2007	Idamurre/
drwxr-xr--	2	liina_1	users	1536 dets 5 2007	Saartemurre/
drwxr-xr--	2	liina_1	users	1024 dets 5 2007	Vorumurre/

Kataloogis Idamurre:

-rwxr--r--	1	liina_1	users	5855 dets 3 2007	KOD_Anna_Lindvere_synt.txt
-rwxr--r--	1	liina_1	users	6210 dets 3 2007	KOD_Jakob_Luka_synt.txt
-rwxr--r--	1	liina_1	users	20039 dets 3 2007	KOD_Miili_Lepp_synt.txt
-rwxr--r--	1	liina_1	users	21166 dets 3 2007	KOD_Miili_Mardijarv_synt.txt

Õigused

Õigused faili vaadata, kirjutada ja käivitada on kirjeldatud kolme rühma kohta: 1) faili omaja (näites kriztel5), grupp (näites users), ülejäänud masina kasutajad. Õigused on kirjeldatud 10-kohalise koodina.

Selles 1. koht näitab, kas tegemist on kataloogiga või mitte:

`drwxr-xr--` 2 kriztel5 users 512 dets 3 2007 bin/ on kataloog

`-rwxr--r--` 1 kriztel5 users 5855 dets 3 2007 KOD_Anna_Lindvere_synt.txt ei ole kataloog, vaid on fail

Järgmised kolm kohta kirjeldavad faili omaniku õigusi:

r – faili lugemisõigus (*read*)

w – faili kirjutamisõigus (*write*)

x – faili käivitamisõigus (*execute*)

Järgmised kolm kohta kirjeldavad grupi õigusi, samuti r, w, x. Et teada saada, mis gruppi sa kuulud, küsi oma gruppe nii:

groups sinu_kasutajanimi

Näide:

```
[105] kriztel5@adalberg:~/korpuste_kursus/murdekorporus> groups kriztel5
users
```

See tähendab, et kasutaja kriztel5 kuulub gruppi users. (Sinna kuuluvad kõik kasutajad, osal kasutajatel on muidugi rohkem gruppe.)

Harjuta:

liigu oma kodukataloogi ja vaata, mis selle sees on: ls

tee sinna uus kataloog, mille nimi on murre: mkdir murre

Liigu oma kodukataloogist õppekorpusesse (siina antakse ette tee n-ö failipuu päris juurelt alates):

```
cd /home/mustikas/k/kriztel5/korpuste_kursus/murdekorpus
```

liigu murdekorpuse sees Idamurdesse: cd Idamurre

liigu sealt tagasi: cd ..

liigu Saarte murde kataloogi: cd Saartemurre

vaata selle sisu: ls

kopeeri sealt kõik xml-laiendiga failid oma kodukataloogi alamkataloogi murre:

```
cp *.xml /home/sinu_kasutajanimi/murre/
```

liigu oma kodukataloogi alamkataloogi murre ja vaata, kas failid on seal kenasti olemas, vaata ka failide õigusi:

```
cd
```

```
cd murre
```

```
ls -l
```

Vihje: käsuread on sageli pikad ja tüütud iga kord kirjutada, eriti kui näiteks käsurida läheb vussi ja tuleb otsast alata. Siis oleks mugavam võtta eelmine käsurida ja seda modifitseerida. Selleks mõned võtted:

nool üles: eelmine käsk, mille sisestasid/sooritasid

Cntrl-u käsurea kustutamine (vajuta Cntrl-klahv ja u-klahv korraga alla!)

Cntrl-a käsurea algusesse

Failide kopeerimine, nime muutmine, kustutamine

cp fail1 fail2 faili kopeerimine failist1 faili2

cp fail /home/kasutajanimi

faili kopeerimine oma kodukataloogi (kasutajanimi = sinu kasutajanimi)

mv fail1 fail2 failinime muutmine, faili ümbertõstmine (failist1 faili2)

rm fail faili kustutamine

Harjuta:

- Loo oma kodukataloogis kataloog murdekorpust. `mkdir murdekorpust`
- Kopeeri sinna üks fail Idamurdest (selleks liigu `cd`-ga eelnevalt kataloogi `/home/mustikas/k/kriztel5/korpuste_kursus/murdekorpust`)
`cp KOD_Miili_Lepp_synt.xml/home/sinu_kodukataloog/murdekorpust`
- liigu sinna kataloogi: `cd /home/sinu_kodukataloog/murdekorpust`
(või 2 käsku:
`cd`
`cd murdekorpust`)
- nimeta see fail ümber failiks `proov.txt`
`mv KOD_Miili_Lepp_synt.xml`
- kustuta see fail
`rm proov.txt`

Vihje: pikkade failinimede ümberkirjutamine iga kord on väga tüütu ja lihtne on teha näpuvigu. Selle tegevuse hõlbustamiseks on TAB-klahv (kust tavaliselt saab taanet): sisesta failinime (või ka kataloogi) esimesed tähed ja vajuta TAB-i, siis pakutakse sulle failinime/kataloogi nime lõppu.

Tekstide sirvimine

Kõige lihtsam tekstieditor `pico`, mõnes serveris on kasutusel ka analoogiline `nano`.

Kui fail on olemas:

`pico` failinimi faili loomine või toimetamiseks avamine
(või: `nano` failinimi)

Kui faili pole, luuakse sama käsuga uus fail.

Programmist `pico` või `nano` väljumine: `Cntr-x`, akna allservas ka muud valikuvariandid. Kui failis on tehtud muutusi, küsitakse väljumisel, kas salvestada.

Faile saab avada lugemiseks/vaatamiseks/uute käskude rakendamiseks ka teisiti, näiteks:

<code>cat fail</code>	avab terve faili (suunab standardväljundisse, st ekraanile)
<code>tail fail</code>	avab faili lõpu (oletuslikult 10 rida)
<code>tail -25 fail</code>	avab viimased 25 rida failist
<code>head fail</code>	avab faili alguse (oletuslikult 10 rida)
<code>head -25 fail</code>	avab esimesed 25 rida failist
<code>more fail</code>	võimaldab faili lehekülghaaval lehitseda (saab lehitseda ainult edaspidi) siin edasiliikumine tühikuklahviga
<code>less fail</code>	võimaldab faili lehekülghaaval lehitseda (saab lehitseda ainult edaspidi) siin edasiliikumine tühikuklahviga katkestamine <code>Cntrl-c</code>

Kui soovid avada/sirvida korraga mitut faili, siis tuleb kasutada failinimede ühisosa (näiteks seda, et failid algavad või lõpevad ühtmoodi). * tähistab mistahes arvu sümboleid:

<code>cat KOD_*</code>	avab kõik failid, mille alguses <code>KOD_</code>
<code>cat *.txt</code>	avab kõik <code>.txt</code> -laiendiga failid

| "toru" võimaldab järjestikku sooritada mitut käsku, nt
cat KOD_* | more avab selles kataloogis kõik failid, mille alguses on KOD_ ning laseb neid
ekraanil lehitseda
cat KOD_* | less
"Toru" on väga vajalik edaspidi pikemate otsiridade vms koostamisel.

Kui soovid teada, mitu sõna on avatud failis (st standardväljundisse = ekraanile suunatud failis),
saab seda teada käsuga wc:

wc
word count, loeb ridu, sõnu ja sümboleid

Näide:

Et teada saada, kui palju materjali on idamurde alal Kodavere tekstides, liigun kõigepealt idamurde tekste sisaldavasse kataloogi, avan Kodavere failid ning seejärel loen wc-ga sõnad üle.

```
cd murdekorpust/Idamurre
```

```
cat KOD* | wc
```

Tulemuseks saan kolm arvu, esimene neist näitab ridu, teine sõnade arvu, kolmas sümbole (tähe märkide) arvu.

```
[108] kriztel5@adalberg:~/murdekorpust/Idamurre> cat KOD* | wc  
379 8579 53270
```

Failide salvestamine

Kui faili on kuidagi modifitseeritud, võib selle suunata standatdväljundi (ekraani) asemel ka uude faili. Näiteks on võimalik kõik Kodavere failid korraka avada ning suunata nad ühte faili näiteks mõnes muus kataloogis. Suunamiseks kasutatakse noolt, failinimi ja asukoht tuleb ette anda.

Näide

```
cd murdekorpust/Idamurre
```

```
cat KOD* > /home/mustikas/k/kriztel5/korpust_kursus /murdekorpust/Idamurre/kodavere.txt
```

Nüüd, liikudes vastavasse kataloogi, saate seda faili sirvida või editeerida picot või nanot kasutades.

Masinast väljumine

Sisesta käsureale exit