Ekilex ja Sõnaveeb rakenduste paigaldusjuhend TripleDev, Martin Laubre

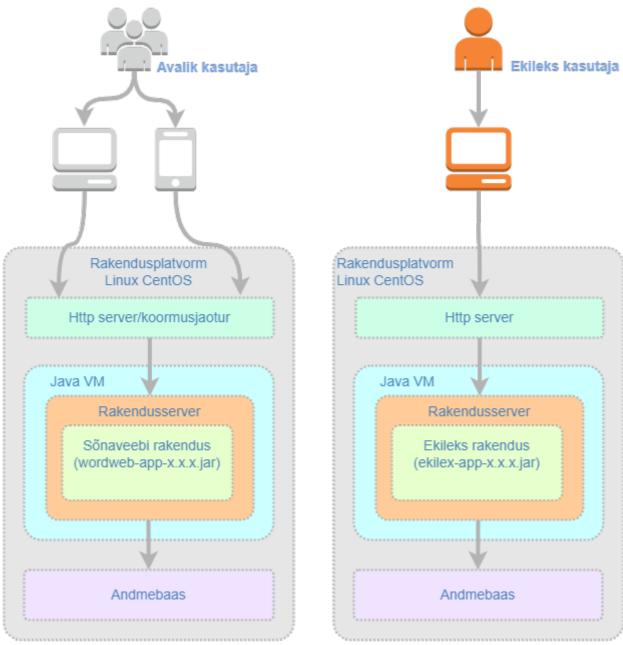
Sisukord

Sissejuhatus	3
Tarkvara paigaldamine	4
Baas-tarkvara	4
Java	4
Apache HTTP	4
Apache Maven	4
Postgres	5
Apache Tomcat	5
Git	5
Rakendustarkvara	6
Tarkvaraprojekt Ekilex	6
Veebirakendus Ekilex	7
Lähtekood	7
Konfigureerimine	7
Andmebaasi struktuuride loomine	8
Ehitamine	9
Käivitamine	9
Kasutajate juurdepääs	11
Veebirakendus Sõnaveeb (Wordweb)	11
Lähtekood	11
Konfigureerimine	11
Andmebaasi struktuuride loomine	12
Ehitamine	13
Käivitamine	13
Sõnakogude laadurid	14
Lähtekood	14
Konfigureerimine	14
Ehitamine	15
Käivitamine	15
Sõnakogude laadimine	15
Ekilex	15
Sõnaveeb	16
Mängude laadimine	16
Ekilex	16
Sõnaveeb	17

Sissejuhatus

Järgnev juhend kirjeldab kahe autonoomse rakenduse, Ekilex ja Sõnaveeb ning nendega seotud baas-tarkvara paigaldamist.

Soovitav arhitektuur on ilmestatud järgneva joonisega:



Joonis 1. Arhitektuur

Kuna Ekilexil ja Sõnaveebil on erinevad kasutajad ning sellest tingitud erinevad süsteemi jõudluse eeldused, siis peavad rakendused asuma füüsiliselt erinevates serverites. Kummagi rakenduse jaoks tuleb teha eraldi täielik paigaldus baas-tarkvarast alates.

Sõnaveebi reaalsetest päringumahtudest tingitult on võimalik lisada jõudlust läbi andmebaasi ja rakenduse füüsilise lahutamise ning andmebaasi ja/või rakenduse klaster-konfiguratsiooni, kuid

seda siis kui selleks reaalne vajadus tekib. Esialgu käesolev juhend sellist konfiguratsiooni ei käsitle.

Samuti ei käsitle käesolev juhend vabavaralise baas-tarkvaraga seotud spetsiifilisi määranguid ja häälestusi. Vastavad juhendid individuaalselt parima konfiguratsiooni koostamiseks on veebis vabalt leitavad.

Tarkvara paigaldamine

Baas-tarkvara

Java

Paigalda Java rakenduste virtuaalmasin Java JDK ver 8. Operatsioonisüsteemile sobiv versioon leia siit:

http://www.oracle.com/technetwork/java/javase/downloads/index.html

http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html

Apache HTTP

Rakendusserver on avaliku veebiliikluse eest varjatud Apache HTTP serveriga, mis klasterkonfiguratsiooni puhul käitub kui tarkvaraline koormusjaotur. Operatsioonisüsteemile sobiv versioon leia siit:

https://httpd.apache.org/download.cgi

Kindlasti paigalda ja konfigureeri HTTP serveris SSL sertifikaat, et rakenduse veebiliiklus käiks üle https protokolli. Täienda hiljem HTTP serveri määranguid reaalsete rakenduse viidetega kui need ükskord paigaldatud saavad. Sulge igasugune muu juurdepääs rakendusserveritele. Rakenduse viite näidis:

```
<Location /wordweb>
   ProxyPass ajp://localhost:5577/wordweb
   ProxyPassReverse ajp://localhost:5577/wordweb
</Location>
```

Apache Maven

Paigalda tarkvara ehitamise ja sõltuvuste halduse raamistik Apache Maven.

https://maven.apache.org/download.cgi

http://www-eu.apache.org/dist/maven/maven-3/3.5.3/binaries/

Käitu vastavalt juhendile:

https://maven.apache.org/install.html

Oluline on, et Maveni määrangud viitaksid korrektsele Java JDK-le.

Postgres

Paigalda andmebaasi tarkvara Postgres ver 9.6 Operatsioonisüsteemile sobiv versioon leia siit:

https://www.postgresql.org/download/

Ehkki Postgres andmebaasi vaikemäärangud on piisava jõudluse tagamiseks enamasti sobivad, siis saab jõudlust oluliselt tõsta andmebaasi kasutamise iseloomu ja reaalset raudvara arvestades. Hea abivahend määrangute häälestamiseks asub siin:

https://pgtune.leopard.in.ua/

Loodud andmebaasi serveritesse loo vastava rakenduse andmebaas ja süsteemne kasutaja.

Rakendus Ekilex:

Andmebaas: ekilex Kasutaja: ekilex

Skeem: public (vaikimisi)

Kodeering (*encoding*): UTF8

Tähestik (collation): et EE.UTF-8

Rakendus Sõnaveeb:

Andmebaas: wordweb Kasutaja: wordweb

Skeem: public (vaikimisi)

Kodeering (encoding): UTF8

Tähestik (collation): et EE.UTF-8

Apache Tomcat

Rakendustarkvara rakendusserveriks on Apache Tomcat (http://tomcat.apache.org/). Antud juhul Tomcat serverit siiski eraldi paigaldama ei pea, sest rakendused Ekilex ja Sõnaveeb kasutavad Spring Boot Tomcat pistakut, mille vahendusel käivitatakse rakendus nö virtuaalses Tomcat ümbrises (wrapper container).

Git

Paigalda tarkvara versioneerimise ja säilitamise tehnoloogia Git.

Operatsioonisüsteemile sobiv versioon leia siit:

https://git-scm.com/downloads

Git kasutamise eelduseks on kasutajakonto olemasolu GitHub-s:

https://github.com/

Rakendustarkvara

Ekilex ja Sõnaveeb rakenduste lähtekoodi struktuur ning tarkvara ehitamise ja käivitamise mehhanism on samasugused. Seepärast on Ekilex rakenduse juhend detailsem ning Sõnaveebi juhend väldib korduvusi.

Tarkvaraprojekt Ekilex

Ekilex tarkvaraprojekti lähtekood asub siin:

https://github.com/tripledev/ekilex.git

Kui Git klientrakendus on konfigureeritud kasutama SSH privaatvõtit, siis alternatiivselt on lähtekoodi asukohaks:

git@github.com:tripledev/ekilex.git

Lae lähtekood kohalikus failisüsteemis kindlasse asukohta kuhu on mugav navigeerida ja tarkvara ehitamine käivitada. Ära kasuta tühikuid sisaldavate nimedega katalooge! Selle juhendi ulatuses on tinglikult selliseks kataloogiks

/apps/source>

Järgnev korraldus laeb alla kogu Ekilex tarkvara projekti, mille alamosadeks on Ekilex ja Sõnaveebi moodulid.

/source>git clone https://github.com/tripledev/ekilex.git

Navigeeri allalaetud tarkvara projekti juur-kataloogi ekilex

Allalaetud lähtekood esitab vaikimisi "master" arendusharu seisu, mis pole ettenähtud toodangu režiimisi rakendamiseks. Toodangu jaoks on kokkulepitud stabiliseeritud versioon, mis asub mingis kindlas harus (*branch*) või *tag*-s. Sobivasse harusse ümberlülitumine käib käsuga

/ekilex>git checkout x.x.x

kus x.x.x on vastava haru nimi, mis tõenäoliselt on tarkvara versiooni number, näiteks 1.0.0 Kõikide eksisteerivate harude nimekirja saab käsuga:

/ekilex>git branch

Ekilex ja Sõnaveeb rakenduste käivitamise eelduseks on keskkonna-spetsiifiliste määrangutega lähtekoodist ehitatud artefaktid

Veebirakendus Ekilex

Lähtekood

Ekilex rakenduse lähtekood asub tarkvara projekti alamkataloogis

/ekilex/ekilex-app>

Konfigureerimine

Peale lähtekoodi allalaadimist ja õigesse harusse ümberlülitumist, konfigureeri toodangukeskkonna-spetsiifiliste määrangute fail

/ekilex/ekilex-app/src/main/resources/application-prod.properties

Rakenduvad kõik määrangud application.properties ja valitud profiili (prod) laiendiga application-prod.properties faili peale kokku kusjuures profiili laiendiga fail on kõrgema prioriteediga. Seepärast ei ole vaja application-prod.properties failis kirjeldada määranguid, mis application.properties failis juba sobival kujul eksisteerivad. Samas saab application-prod.properties määrangutega kirjutada üle sobimatud määrangud application.properties failis. Olulised on järgmised parameetrid.

Rakendusserver:

```
server.port
server.servlet.context-path
server.servlet.session.timeout
```

tomcat.ajp.port
tomcat.ajp.enabled

AJP määrangud reguleerivad rakendusserveri ja avalikku veebi eksponeeriva HTTP serveri vahelist kommunikatsiooni. AJP port peab kindlasti erinema HTTP pordist. AJP protokolli kasutamine rakendusserveri ja HTTP serveri vahel on tungivalt soovitav. Rakendusserveri määrangud peavad olema sellised, et need oleks võimalik avaliku veebi eest varjata.

Andmebaasi ühenduse kirjeldus:

```
spring.datasource.url
spring.datasource.username
```

spring.datasource.password

Vastav loogiline andmebaas tuleb eelnevalt füüsilisse andmebaasi luua. Võimalik, et loogilise andmebaasi nimi ja kasutaja juba sobivadki. Samuti, kui andmebaas ja rakendus asuvad samas füüsilises serveris, võiks ka localhost url sobida.

Failiressursid – pildid, hääldused, jms. Sisu kohta küsi täpsemalt arendajatelt.

file.repository.path

Kasutajate registreerimisega seotud määrangud:

```
ekilex.app.url
email.sending.enabled
email.from.address
email.from.name
spring.mail.host
spring.mail.port
spring.mail.username
spring.mail.password
spring.mail.properties.mail.smtp.auth
spring.mail.properties.mail.smtp.starttls.enable
```

Logimise detailsus:

```
logging.level.root
logging.level.eki.ekilex
logging.level.eki.common
logging.level.org.jooq
```

Vaikimisi on rakenduse enda logid debug detailsusega, muud sõltuvused pääsevad läbi ainult warn taseme detailsusega. Alustuseks võib just selline logi detailsus olla isegi hea mõte.

Logifail:

```
logging.path
logging.file
```

Failinime (logging.file) võiks muuta ainult juhul kui sellest rakendusest on plaan käivitada mitu erinevat instantsi. Reaalselt rikastatakse failinime veel kuupäevaga. Oluline on määrata logifailile asukoha absoluutne rada (logging.path) kuhu eri päevade logifailid kogunema hakkavad. Kui rada ei eksisteeri, see luuakse kui rakendust käivitaval kasutajal on selleks operatsioonisüsteemi mõistes vajalikud õigused.

Andmebaasi struktuuride loomine

Kõik andmebaasi struktuuride loomise ja baasandmetega täitmise skriptid asuvad kataloogis

/ekilex/ekilex-app/src/main/resources/sql/

Oluline! Kui sõnakogude laadimiseks kasutada sõnakogude mass-laadijat (*UltimaLoader*), siis pole nende skriptide käivitamine vajalik – laadur teeb seda ise.

Ehitamine

Peale rakenduse konfigureerimist tuleb rakenduse lähtekoodist ehitada rakenduslik artefakt. Alustuseks tuleb ehitada kogu tarkvara projekt, mis loob ja laeb alla kohalikku Maven repositooriumisse nii selle tarkvaraprojekti moodulite artefaktid kui kõik nendele moodulitele vajalikud sõltuvused kesksetest Maven repositooriumitest. Esmakordne ehitamine võib kesta minuteid.

/ekilex>mvn clean install -D skipTests

Ehitamine õnnestus kui protsessi lõpuks on kuvatud teade: "BUILD SUCCESS" Ehitamise ebaõnnestumisest annab märku teade: "BUILD FAILURE" Ebaõnnestunud ehitamise põhjus on kirjeldatud samas protsessi logis.

Edaspidi, kui on vaja muuta ainult vastava rakenduse määranguid, võib ehitamist käivitada selle rakenduse juures

/ekilex/ekilex-app>mvn clean install -D skipTests

Võti skipTests on vajalik kui kohalikult pole seadistatud automaattestide andmebaas. Igalpool mujal kui toodanguserveris, kus toimub lähtekoodi muutmine võiks automaattestide andmebaasi siiski konfigureerida ning ehitamine käivitada ilma -D skipTests määranguta, sest siis on eduka ehitamise protsessi eelduseks ka positiivse tulemusega automaattestid.

Käivitamine

Edukalt ehitatud rakenduse käivitamisel tuleb valida korrektne profiil (prod). Et kontrollida rakenduse konfigureerimise ja ehitamise edukust, saab seda testimiseks käivitada kahel viisil:

Maven pistaku vahendusel:

/ekilex/ekilex-app>mvn spring-boot:run -D spring-boot.run.profiles=prod

Otse Java rakendusena ehituse tulemi pealt:

/ekilex/ekilex-app/target>java -jar ekilex-app.jar --spring.profiles.active=prod

Kui rakendus käivitub normaalselt, on võimalik hakata jälgima rakenduse logi määratletud logifailis.

Selline rakenduse käivitusviis sobib siiski ainult ehitamise ja konfigureerimise edukuse testimiseks. Reaalselt on tungivalt soovitav käivitada rakendus systemd teenusena.

Loo sobiv Linux kasutaja, kellena rakendust hakatakse käivitama. See ei tohiks olla root kasutaja! Käesolevas juhendis on selleks kasutajaks ekilex

Kopeeri rakenduse ehituse tulemi kataloogist rakenduse artefakt ekilex-app.jar sobivasse kataloogi. Selle juhendi ulatuses on tinglikult selliseks kataloogiks

/apps/deploy/ekilex>

Loo fail

/etc/systemd/system/ekilex.service

Sisusta fail järgmisega:

[Unit]

Description=EKILEX application After=syslog.target

[Service]

User=ekilex

ExecStart=/apps/deploy/ekilex/ekilex-app.jar

WorkingDirectory=/apps/deploy/ekilex

Environment="JAVA HOME=/opt/jdk1.8.0 144"

SuccessExitStatus=143

[Install]

WantedBy=multi-user.target

Loo rakenduse artefakti naabrusesse täiendavate määrangute fail, mis kannab sama nime kuid on erineva faililaiendiga:

/apps/deploy/ekilex/ekilex-app.conf

Sisusta fail järgmisega:

JAVA_OPTS=-Xmx4096M RUN_ARGS=--spring.profiles.active=prod

Teenus on valmis kasutamiseks

/>systemctl start ekilex
/>systemctl restart ekilex

/>systemctl stop ekilex

Kasutajate juurdepääs

Ekilexi kasutamine on kaitstud autoriseerimisega. Kasutajad saavad ennast ise registreerida. Registreerimise protseduur sisaldab kinnituskirja saatmist kasutaja määratud e-posti aadressile. Kiri sisaldab linki, millega kasutaja oma registreerimist kinnitab. Kirja link on seotud ekilex rakenduse määrangutes kirjeldatud rakenduse aadressi.

Endiselt on lähtekoodi hulka lisatud seniste registreeritud kasutajate loomise skripti fail eki-users.sql

Veebirakendus Sõnaveeb (Wordweb)

Lähtekood

Sõnaveebi rakenduse lähtekood asub tarkvara projekti alamkataloogis

/ekilex/wordweb-app>

Konfigureerimine

Konfigureeri toodangukeskkonna-spetsiifiliste määrangute fail

/ekilex/wordweb-app/src/main/resources/application-prod.properties

Rakendusserver:

```
server.port
server.servlet.context-path
server.servlet.session.timeout
tomcat.ajp.port
tomcat.ajp.enabled
```

Rakendusserveri tööstatistika pistaku Actuator juurdepääsu kirjeldus:

```
spring.security.user.name
spring.security.user.password
spring.security.user.roles
```

Andmebaasi ühenduse kirjeldus:

```
spring.datasource.url
spring.datasource.username
spring.datasource.password
```

Failiressursid:

file.repository.path

Välisliidesed – kõnesüntesaator, kõnetuvastus, eesti ja vene korpused.

```
speech.synthesizer.service.url speech.recognition.service.url corpora.service.est.url corpora.service.est.corpname corpora.service.rus.url corpora.service.rus.corpname corpora.service.rus.corpname corpora.service.rus.username corpora.service.rus.api.key
```

Kasutajate tagasiside link ekilex suunas:

wordweb.feedback.service.url

Logimise detailsus:

```
logging.level.root
logging.level.eki.ekilex
logging.level.eki.common
logging.level.org.jooq
```

Logifail:

logging.path
logging.file

Andmebaasi struktuuride loomine

Sõnaveebi andmebaas koosneb peamiselt materialiseeritud vaadetest (*materialized views*), mis koostatakse üle andmebaasi side-ühenduse (*dblink*) Ekilexi andmete pealt. Kõik andmebaasi struktuuride loomise ja andmetega täitmise skriptid asuvad kataloogis

```
/ekilex/wordweb-app/src/main/resources/sql/
```

create_mviews.sql – kõikide materialiseeritud vaadete loomine ja sisustamine sõnakogude andmetega ja klassifikaatoritega

create tables.sql – kõikide tabelite loomine ja sisustamine mängude andmetega

Esmakordsel konfigureerimisel tuleb Sõnaveebi andmebaasis aktiveerida andmebaasi laiend dblink. See käib andmebaasi terminalis käsuga:

```
CREATE EXTENSION dblink;
```

Vajalikku side-ühendust kahe andmebaasi vahel saab testida käsuga:

SELECT dblink_connect('host=localhost user=ekilex password=<parool> dbname=ekilex');

Modifitseeri selle päringu parameetreid vastavalt reaalsele Ekilex konfiguratsioonile.

Edukas ühendus kahe andmebaasi vahel tagastab tulemuseks "OK"

Ehitamine

Kogu tarkvaraprojekti ehitamine:

/ekilex>mvn clean install -D skipTests

Sõnaveebi rakenduse ehitamine:

/ekilex/wordweb-app>mvn clean install -D skipTests

Käivitamine

Maven pistaku vahendusel:

/ekilex/wordweb-app>mvn spring-boot:run -D spring-boot.run.profiles=prod

Otse Java rakendusena ehituse tulemi pealt:

/ekilex/wordweb-app/target>java -jar wordweb-app.jar --spring.profiles.active=prod

Käivitamine teenusena. Vajalikud eeldused.

Kopeeri rakenduse artefakt wordweb-app.jar kataloogi:

/apps/deploy/wordweb>

Loo fail

/etc/systemd/system/wordweb.service

Sisusta fail järgmisega:

[Unit]

Description=WORDWEB application

After=syslog.target

[Service]

User=wordweb

ExecStart=/apps/deploy/wordweb/wordweb-app.jar

WorkingDirectory=/apps/deploy/wordweb

Environment="JAVA HOME=/opt/jdk1.8.0 144"

SuccessExitStatus=143

```
[Install]
WantedBy=multi-user.target
```

Loo rakenduse artefakti naabrusesse fail:

/apps/deploy/wordweb/wordweb-app.conf

Sisusta fail järgmisega:

```
JAVA_OPTS=-Xmx4096M
RUN ARGS=--spring.profiles.active=prod
```

Teenus on valmis kasutamiseks

```
/>systemctl start wordweb
/>systemctl restart wordweb
/>systemctl stop wordweb
```

Sõnakogude laadurid

Hetkel on võimalik sõnakogusid Ekilexi laadida ainult EKI oma XML-vormingus failidest.

Peale Ekilexi sõnakogude laadimist on võimalik need edasi laadida Sõnaveebi.

Ekilexi saab sõnakogusid laadida nii ükshaaval kui korraga.

Oluline on teada, et laadimine toimub käsurearakendustega, mille tehnoloogiline koosseis on erinev veebirakendustest ning seetõttu on erinev ka konfigureerimine ja käivitamine.

Lähtekood

Laadurite käsurearakenduste lähtekood asub tarkvara projekti alamkataloogis

```
/ekilex/ekilex-etl>
```

Konfigureerimine

Konfigureeri toodangukeskkonna-spetsiifiliste määrangute faili

```
/ekilex/ekilex-etl/envresources/prod/ekilex-etl.properties
```

Ekilex rakenduse andmebaasi ühenduse kirjeldus:

```
db.ekilex.url
db.ekilex.usr
db.ekilex.psw
```

Termeki andmebaasi ühenduse kirjeldus:

```
db.termeki.url
db.termeki.usr
db.termeki.psw
```

Termeki failide teenuse kirjeldus ja kohalik failide laadimise sihtkataloog:

```
termeki.file.service.url
file.repository.path=/apps/data/files/
```

Kui on plaanis kasutada universaalset kõigi sõnakogude mass-laadijat, siis tuleb konfigureerida ka järgmine fail:

```
/ekilex/ekilex-etl/envresources/prod/ultima-loader.properties
```

Logimise määrangud asuvad:

/ekilex/ekilex-etl/envresources/prod/logback.xml

Ehitamine

Kogu tarkvaraprojekti ehitamine:

```
/ekilex>mvn clean install -D skipTests
```

Sõltuvalt millist laadurit on soov käivitada, võib ehitamine olla erinev. Täpsemalt sellest peatükis "Sõnakogude laadimine"

Käivitamine

Erinevate laadurite käivitamisest räägib peatükk "Sõnakogude laadimine"

Sõnakogude laadimine

Ekilex

Hetkel käsitleb juhend ainult universaalset kõigi sõnakogude mass-laadijat (*UltimaLoader*), mis käivitamisel täielikult tühjendab määratud andmebaasi ning laeb sinna järjest kõik määratud sõnakogud. Ettevaatust!

Peale määrangufailide konfigureerimist, enne käivitamist, tuleb vastav rakendus ehitada.

```
/ekilex/ekilex-etl>mvn clean install -D skipTests -P prodsrvall
```

Käivitamine:

```
/ekilex/ekilex-etl>mvn exec:java -P prodsrvall
```

Laadimise edenemist ja õnnestumist saab jälgida määratud logifailist. Sõnakogud on laetud, kui

logi lõppeb teatega: "DONE LOADING DATASETS!!"

Sõnaveeb

Sõnakogude Sõnaveebi laadimine ning hiljem värskendamine toimub üle andmbaaside sideühenduse (DB link) andmebaasi vaadete (view) vahendusel. Andmebaaside side-ühendamisest on juttu Sõnaveebi rakenduse konfigureerimise juures.

Sõnakogude algne laadimine toimub andmebaasi skriptiga, mis asub tarkvaraprojekti lähtekoodis:

```
/ekilex/wordweb-app/fileresources/sql/create_mviews.sql
```

See skript kustutab eksisteerivad, loob uued materialiseeritud vaated (*materialized views*) vastavate Ekilex andmebaasi vaadete pealt ning indekseerib need vaated.

Modifitseeri skriptis andmebaasi side-ühenduse parameetreid.

Ära unusta lõpetada skript käsuga:

commit;

Edaspidi tuleb materialiseeritud vaadete loomise skripti create_mviews.sql käivitada alati kui Ekilexi või Sõnaveebi andmemudelisse tehakse muudatusi.

Kui aga Ekilexis muutuvad ainult andmed, siis piisab materialiseeritud vaadete värskendamisest:

```
refresh materialized view mview_ww_word;
refresh materialized view mview_ww_as_word;
refresh materialized view mview_ww_form;
refresh materialized view mview_ww_form;
refresh materialized view mview_ww_lexeme;
refresh materialized view mview_ww_collocation;
refresh materialized view mview_ww_classifier;
refresh materialized view mview_ww_dataset;
refresh materialized view mview_ww_dataset;
refresh materialized view mview_ww_lexeme_relation;
refresh materialized view mview_ww_lexeme_relation;
refresh materialized view mview_ww meaning relation;
```

Mängude laadimine

Ekilex

Sõnaveebi mängude alusandmed koostatakse kombineeritult sõnakogude andmetest ja täiendava(te)st andmefaili(de)st automaatselt koos sõnakogude laadimisega mass-laadija (*UltimaLoader*) poolt. Andmefailide asukoht on kirjeldatud ultima-loader.properties failis.

Sõnaveeb

Mängude algne laadimine toimub andmebaasi skriptiga, mis asub tarkvaraprojekti lähtekoodis:

/ekilex/wordweb-app/fileresources/sql/create_tables.sql

See skript kustutab eksisteerivad, loob uued tabelid kõigi mängude jaoks ning sisustab üle andmebaasi side-ühenduse andmetega Ekilexist. Ühtlasi luuakse tabelid mängude tulemuste salvestamiseks. Seetõttu olla ettevaatlik selle skripti korduval kasutamisel kui mängude tulemuste tabelid sisaldavad juba reaalsete kasutajate mängude tulemusi.

Juhendi lõpp

(Dokument täieneb jooksvalt. Jälgi dokumendi kuupäeva!)