

Kidetietokanta – Toteutusdokumentti

Vesa Ylhäinen & Jussi Tiira

Johdanto

Tämän järjestelmän tarkoitus on antaa tutkijoille helppo pääsy jääkiteistä kerättyyn tietoon. Lisäksi se mahdollistaa Hannakaisa Lindqvistin kehittämän IC-PCA kideluokittimen helpon testauksen. Järjestelmä on tehty käytettäväksi Helsingin Yliopiston fysiikan laitoksen ylailmakehän jääkide tutkimuksessa.

Toimintaymäristö

Tieto on tallennettu posgresSQL tietokantaan ja siihen pääsee suurilta osin php:llä generoitujen html-sivujen kautta. Järjestelmä on tällä hetkellä asennettu laitoksen Users palvelimelle.

Rajaukset

Järjestelmän toteutuksessa keskitymme järjestelmän välttämättömään toiminnallisuuteen. Ulkoasu on funktionaalinen ja minimalistinen. Järjestelmä on tarkoitettu tutkijoiden käyttöön joten tarjoamme datan mahdollisimman täydellisenä ja muokattavassa muodossa. Esimerkiksi graafisia esityksiä datasta järjestelmämme ei tarjoa. Oletamme että tutkijat haluavat kuitenkin kuvantaa dataa itse.

Suunnitellun lisäksi toteutimme yksinkertaisen statistiikan käyttäjän käsin luokittelemasta kidedatasta.

Kehitysmahdollisuuksia ovat muunmuassa automaattinen käyttäjien hallinta, csv-latauksen ohella tekstimuotoinen datan esikatselu datasettiä koskevan statistiikan laajentaminen.

Toteutusympäristö

Työ toteutetaan käyttäen Netbeans IDE:a ja testaamalla sitä työn edetessä suoraan Users-palvelimella.

Testaus

Ohjelma on testattu käsin. Seuraavaa testauksessa löytyneitä bugeja ei ole korjattu:

- Manual classification sivu ei muista site-valintaa.

Ohjelmiston yleisrakenne

Ohjelmisto koostuu seuraavista sivuista.

login.html

Sivu jonka käyttäjä näkee ensimmäisenä. Sisältää sisäänkirjautumislomakkeen joka oikeilla tunnuksilla johtaa csv-generator.php sivulle. Väärät tunnukset tuottavat virheilmoituksen. Lisäksi mahdollistaa käyttöoikeuksien pyynnön sähköpostilomakkeella.

csv.generator.php

Ojelmiston pääsivu joka sisältää valintakentät datasetin rajaamiseen, reset painikkeen lomakkeen tyhjentämiseksi sekä submit painikkeen jota painamalla käyttäjälle tarjoutuu mahdollisuus ladata valitsemansa datasetti csv tiedostona. Lisäksi sisältää navigointipalkin josta pääsee manual classification sivulle sekä kirjautumaan ulos. Uloskirjautuminen palauttaa käyttäjän login.php sivulle.

manual_classification.php

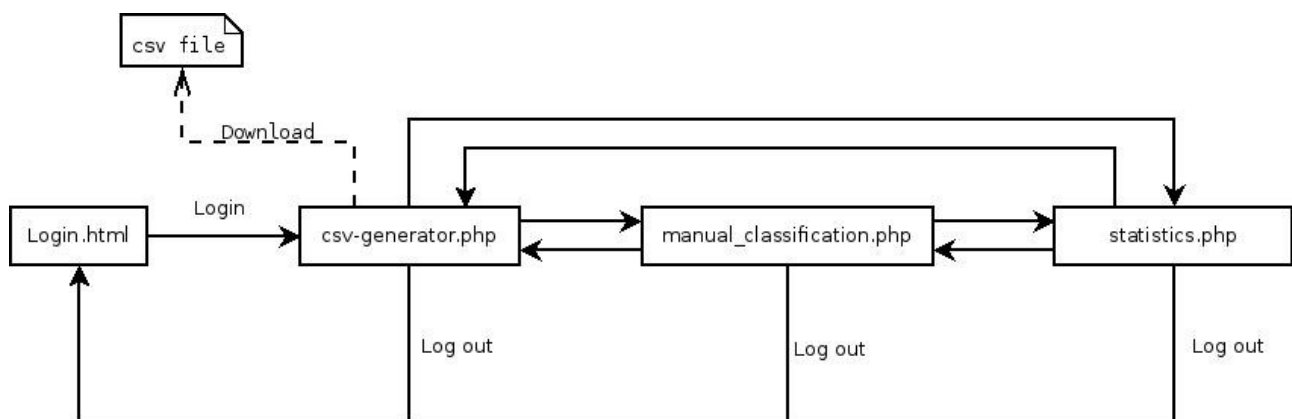
Näyttää kiteiden kuvia käyttäjälle. Sisältää dataseting rajausvalinnat sekä valikon josta käyttäjä voi luokitella annetun kiteen. Lisäksi sisältää navigointipalkin josta pääsee csv-generator sivulle sekä kirjautumaan ulos.

statistics.php

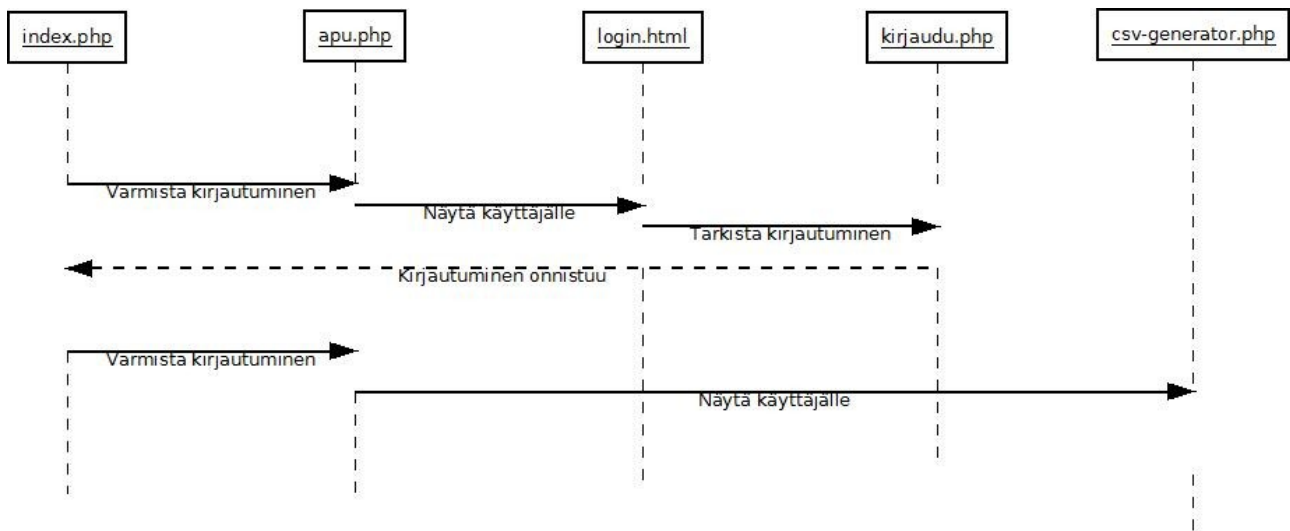
Yksinkertaista статистиikkaa käyttäjän käsin luokittelemista kiteistä.

Järjestelmän toiminta

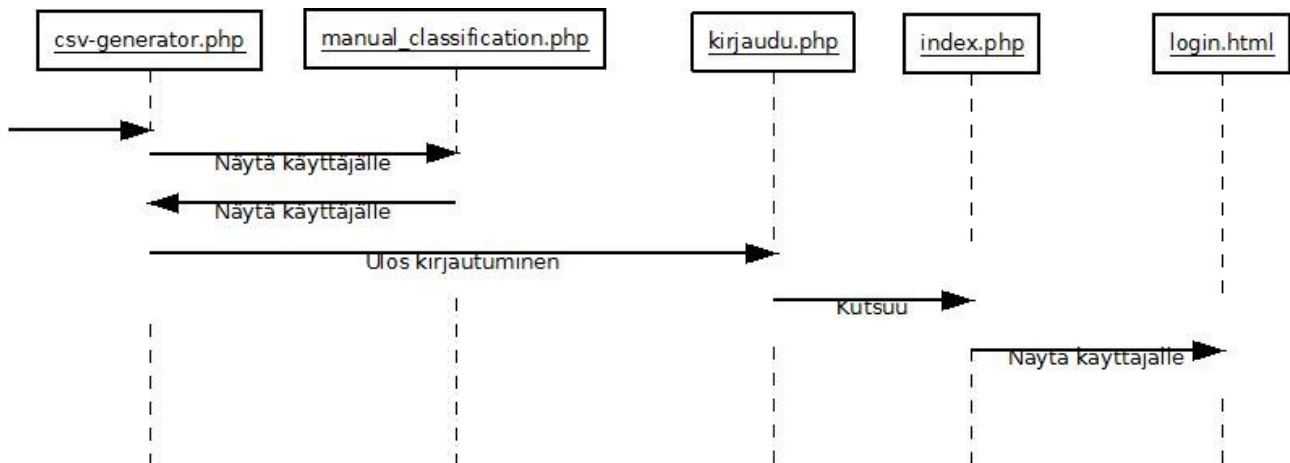
Navigointi järjestelmässä.



Komponenttien välinen viestintä käyttäjän kirjautuessa.



Komponenttien välinen viestintä käyttäjän navigoidessa palvelua.



Järjestelmän komponentit

Php-tiedostojen ohjelmakoodin logiikkaa on kommentoitu tarkemmin itse tiedostoon.

index.php

Vaatii apu.php:n ja sen avulla varmistaa onko käyttäjä kirjautunut ja ohjaa käyttäjän tämän mukaan joko login sivulle tai csv-generator sivulle.

apu.php

Kirjasto joka vaatii kyselyt.php:n. Sisältää funktiot sessioiden hallintaan, wrapperit osoitteiden ja kirjautumisen hallitsemiseen sekä filtterin joka muodostaa sql kyselyn käyttäjän inputin mukaan. Lisäksi sisältää kaksi taulua, toisessa on datasaitit ja toisessa on kide luokat ja niiden lyhenteet. Ohjaa epäonnistuneen kirjautumisen takaisin login.php sivulle.

kirjaudu.php

Hallinnoi sekä sisään- että uloskirjautumista. Vaatiin apu.php:n. Ohjaa index.php:hen

csv-generator.php

Vaatiin apu.php:n ja header.html:n. Varmistaa kirjautumisen. Näyttää käyttäjälle datasetin rajausvalinnat sekä navigointipalkin. Mahdollistaa csv tiedostojen latauksen. Lomake lähettää tietonsa csvgen.php:lle. Sisältää select dataset -kenttäkokoelman, time resolution -valitsimen, particle properties -kenttäkokoelman ja huonolaatuisen datan suodatusvalintaruudun.

csvgen.php

Vaatiin apu.php:n ja yhteys.php:n. Hoitaa csv-generatorin logiikkaa ja formuloi käyttäjän valinnoista SQL kyselyn. Muuttaa tuloksen csv tiedostoksi. Ohjaa käyttäjän lataamaan csv tiedoston.

manual_classification.php

Vaatiin apu.php:n, kidevalitsin.php:n ja header.html:n. Näyttää käyttäjälle Manual Classification sivun, kiteen kuvan sekä rajausvalinnat. Sisältää myös navigointipalkin. Lomake kutsuu manual_class.php:tä. Sisältää select dataset-, particle properties-, IC-PCA classification ja käyttäjän luokitus -kenttäkokoelmat, sekä huonolaatuisen datan merkintävalintaruudun.

manual_class.php

Vaatiin apu.php:n ja yhteys.php:n. Hoitaa manual_classification.php:n logiikkaa. Hoitaa tietokantahaun sekä lisäyksen käyttäjän syötteen mukaan. Ohjaa takaisin manual_classification.php:hen seuraavan kiteen id:llä.

statistics.php

Vaatii apu.php:n, yhteys.php:n ja header.html:n. Varmistaa kirjautumisen. Näyttää käyttäjälle статистиikkaa tämän käsin luokittelemista kiteistä.

yhteys.php

Yhdistämisfunktio joka luo yhteyden tietokantaan. Käyttää Jussin tietokantatunnuksia.

mail.php

Lähettaa Jussille sähköpostilla request account lomakkeen tiedot.

kidevalitsin.php

Vaatiin yhteys.php:n. Valitsee seuraavan sellaisen kiteen jota kukaan ei ole vielä luokitellut.

kyselyt.php

Vaatiin yhteys.php:n. Hoitaa käyttäjän tunnistukseen tarvittavan SQL-kyselyn.

login.html

Staattinen sivu joka sisältää kirjautumis lomakkeen ja request account lomakkeen. Kirjautumislomake kutsuu kirjaudu.php:tä ja request account lomake kutsuu mail.php:tä.

header.html

Sisältään navigointipalkin joka linkkaa csv-generator.php:hen, manual_classification.php:hen ja kirjaudu.php:hen avainsanalla ulos.

Asennustiedot

Tietokanta sijaitsee laitoksen Users-palvelimella. Kaikki järjestelmän komponentit ovat omissa alikansioissaan palvelimen htdocs-kansiossa.

Tietokanta on tarkoitettu yhdeltä yleiseltä palvelimelta, mutta jos haluat asentaa sen itse, näin se käy:

Komponentit sisältävä paketti puretaan public kansioon. Tämän jälkeen luodaan PosGreSQL tietokanta täältä löytyvin ohjein: <http://www.postgresql.org/docs/9.2/static/tutorial-createdb.html>

Lisää juuri luomasi tietokannan nimi, käyttäjätunnus ja salasana yhteys.php tiedostoon riville 5:

```
$yhteys = new PDO("pgsql:host=localhost;dbname=TIETOKANNANNIMI",  
"KÄTTÄJÄTUNNUS", "SALASANA");
```

Tämän jälkeen ajetaan kide_dbinit.sql joka alustaa tietokannan.

Nyt tietokanta on valmis vastaanottamaan dataa. Käyttäjiä voit lisätä SQL komentotulkista komennolla INSERT INTO users VALUES('käyttäjätunnus','salasana','email')

Käynnistys ja käyttöohje

Navigoi selaimella hakemistoon jossa järjestelmä sijaitsee. Meidän tietokantamme tapauksessa osoite on: jussitii.users.cs.helsinki.fi

Kun olet saanut käyttäjätunnukset (voit käyttää testitunnuksia user:anonymous, psw:testi), kirjaudu sisään ja olet csv-generator sivulla. Valikoiden avulla voit päättää minkälaisien kiteiden tiedot haluat ladata ja painamalla submit query saat ladattua tämän data csv tiedostona.

Data on seuraavassa muodossa: Aikaikkunan alkamishetki, kiteiden kokonaismäärä, platejen kokonaismäärä, bullettien kokonaismäärä, columnien kokonaismäärä, irregularien kokonaismäärä, rosettejen kokonaismäärä, rosette agregaatien kokonaismäärä, plate agregaatien kokonaismäärä, column agregaatien kokonaismäärä, maksimidimension keskivarvo, pinta-alalla painotettua area ration keskiarvo.

Kaikki luokitukset tässä ovat tehty c5nn metodilla joka on metodeista luotettavin.

Navigoimalla manual_classification sivulle voit luokitella itse kiteitä. Voit valita ensin datasetin ja ominaisuudet joiden mukaan rajata luokiteltavia kiteitä. Palvelu näyttää sinulle kiteen ja sinun tulee päättää mihin luokkaan se kuuluu. Jos olet uusi luokittelussa, tutustu julkaisuun Ice-cloud particle habit classification using principal components (H. Lindqvist et al. 2012)

User statistics -sivulta pääset tarkastelemaan käsin luokittelemiisi kiteisiin liittyvää статистиikkaa.

Logout painikkeella pääset kirjautumaan ulos.

Liitteet

Suunnitteludokumentti.odt, kide_db.zip,