

# 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.

## Toimintaympäristö

Järjestelmä vaatii toimintaympäristökseen Linux-käyttöjärjestelmän, Apache-palvelinohjelmiston, PostgreSQL-tietokannan ja PHP:n, eli ns. LAPP-kokoelman.

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

## Rajaukset

Järjestelmän toteutuksessa keskityimme 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 datan esikatselu html-tilana, ja käyttäjän käytettävissä olevan statistiikan laajentaminen.

## Toteutusympäristö

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

## Testaus

Ohjelma on testattu käsin. Seuraavaa testauksessa löytynyttä bugia 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

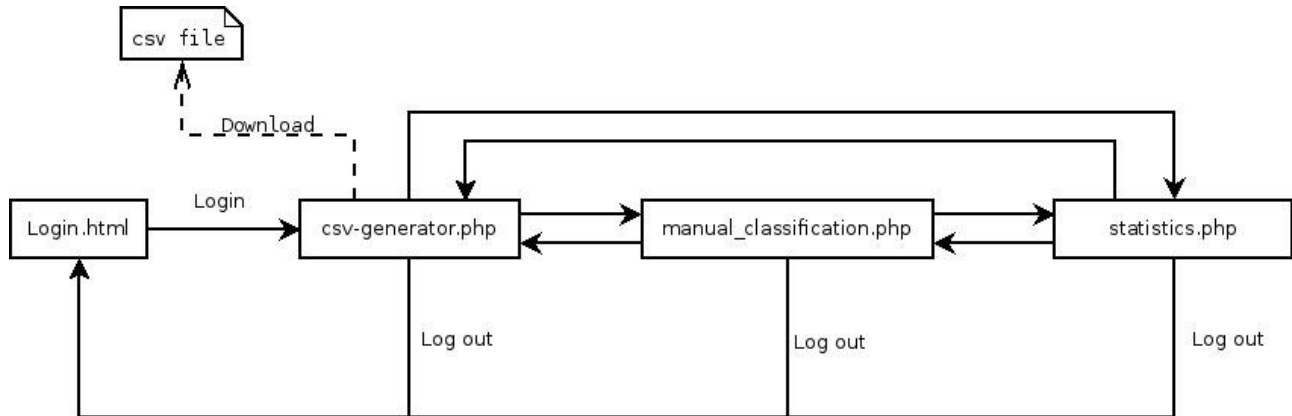
Näyttää kiteiden kuvia käyttäjälle. Sisältää dataseting rajausvalinnat sekä valikon josta käyttäjä voi luokitella annetun kiteen. Lisäksi sisältää navigointipalkin josta pääsee csv-generator sivulle sekä kirjautumaan ulos.

## statistics.php

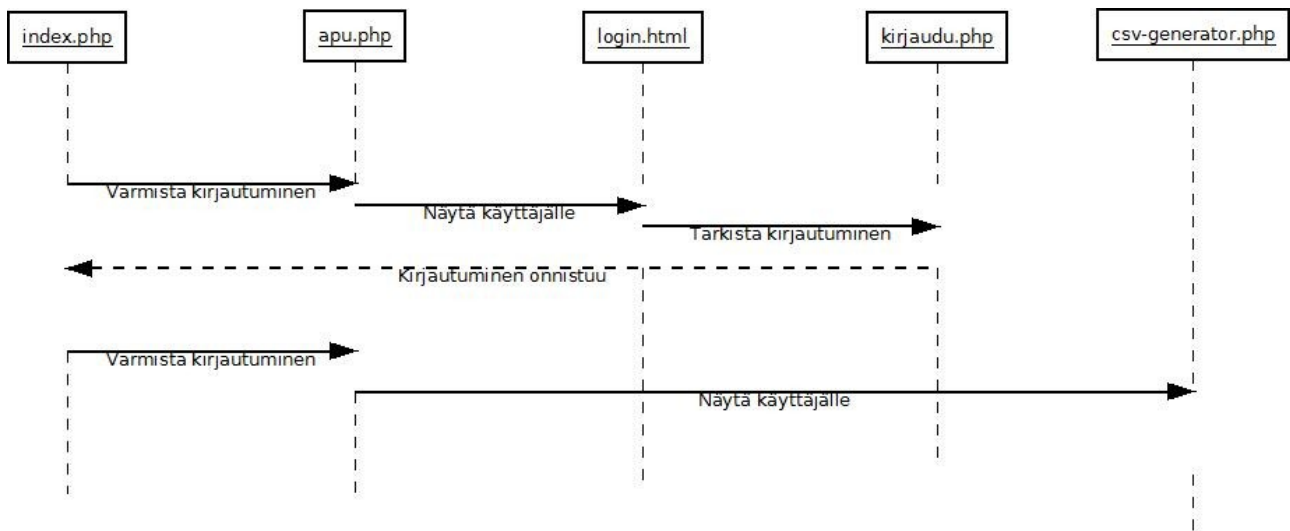
Yksinkertaista statistiikkaa käyttäjän käsin luokittelemista kiteistä.

## Järjestelmän toiminta

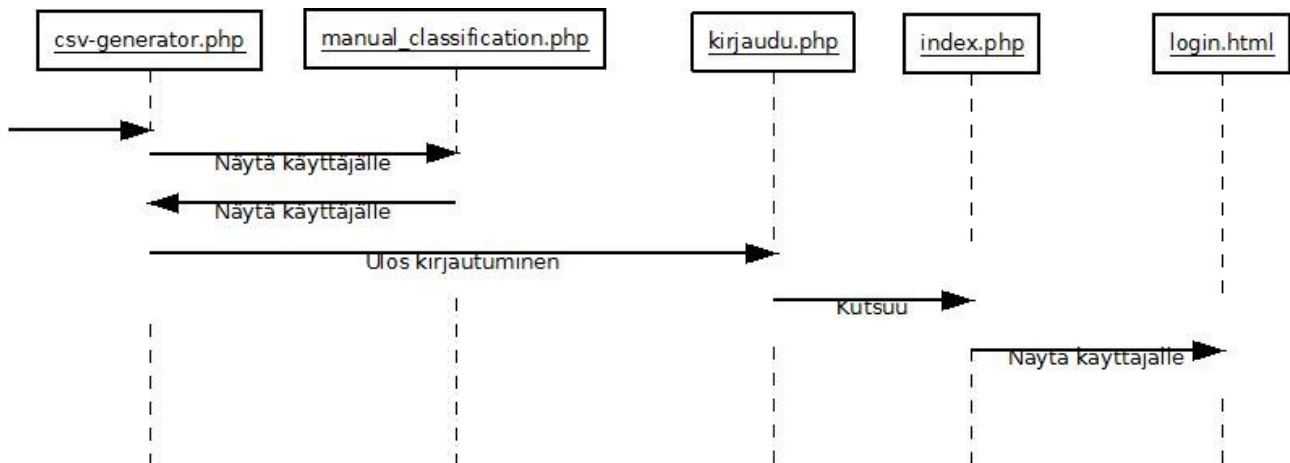
Navigointi järjestelmässä.



Komponenttien välinen viestintä käyttäjän kirjautuessa.



Komponenttien välinen viestintä käyttäjän navigoidessa palvelua.



## Järjestelmän komponentit

### PHP

Php-tiedostojen ohjelmakoodin logiikkaa on kommentoitu tarkemmin itse tiedostoon. Osa tiedostoista sisältää html-koodin sekaan sulautettua php-ohjelmakoodia.

#### 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 kidealuokat ja niiden lyhenteet. Ohjaa epäonnistuneen kirjautumisen takaisin login.php sivulle.

### **kirjaudu.php**

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

### **csv-generator.php**

Vaatii 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**

Vaatii 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**

Vaatii apu.php:n, kidevalitsin.php:n ja header.html:n. Näyttää käyttäjälle Manual Classification sivun, luokiteltavan 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.

\*kuvatiedostot oltava img/kide -hakemistossa, ks. ”Datan lisääminen” viimeisellä sivulla.

### **manual\_class.php**

Vaatii 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**

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

### **kyselyt.php**

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

## **HTML**

### **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.

## **JavaScript**

Kaikki .js -tiedostot sijaitsevat js-hakemistossa.

### **hidden\_fieldset.js**

Sisältää fieldset-elementtien avaamiseen ja piilottamiseen tarvittavan funktion.

### **jquery.js**

JavaScript -kirjasto. ks. <http://jquery.com/>

### **slimbox2.js**

jQueryä hyödyntävä JavaScript -kirjasto kuvien esittämiseksi smartbox-laatikoina.

## **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 hakemistoon, johon ohjelma halutaan asentaa. Vaihtoehtoisesti tiedostot voi kloonata suoraan Githubista käyttäen Git-versionhallintaohjelmaa. Ohjelma on saatavilla osoitteesta: [https://github.com/juhi24/kide\\_db](https://github.com/juhi24/kide_db)

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 sql-hakemistosta löytyvä kide\_dbinit.sql joka alustaa tietokannan.

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

## **Käynnistys ja käyttöohje**

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

Kun olet saanut käyttäjätunnukset (voit käyttää testitunnusta anonymous, salasanalla: testi), kirjaudu sisään ja olet csv-generator sivulla. Valikoiden avulla voit päättää minkälaisen 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 agregaattien kokonaismäärä, plate agregaattien kokonaismäärä, column agregaattien kokonaismäärä, maksimidimension keskivarvo, pinta-alalla painotettua area ration keskiarvo.

Kaikki luokitukset tässä ovat tehty 5NN-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 luokaan 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.

## ***Datan lisääminen***

Jääkidedata on tarkoitettu lisättäväksi tietokantaan erikseen toimitettavalla Matlab-ohjelmalla. Lisäksi manuaaliluokitusta varten on kuvat yksittäisistä kiteistä siirrettävä img/kide -hakemistoon. Kuvien tiedostomuodon on oltava muotoa <id>.jpg, missä <id> on tietokantaan tallennetun kiteen id.

Esimerkkiasennuksessa osoitteessa <http://jussitii.users.cs.helsinki.fi> on valmiiksi ladattuna runsaasti esimerkkidataa kuvatiedostoineen.

## **Liitteet**

Suunnitteludokumentti.pdf