

Kidetietokanta – Suunnitteludokumentti

Vesa Ylhäinen & Jussi Tiira

Johdanto

Järjestelmän tarkoitus

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 yläilmakehän jääkidetutkimuksessa.

Toimintaympäristö

Tieto on tallennettu posgresSQL tietokantaan ja siihen pääsee käsiksi html sivujen kautta toimivalla php ohjelmalla. Järjestelmä on tällä hetkellä asennettu laitoksen Users palvelimelle.

Rajaukset

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

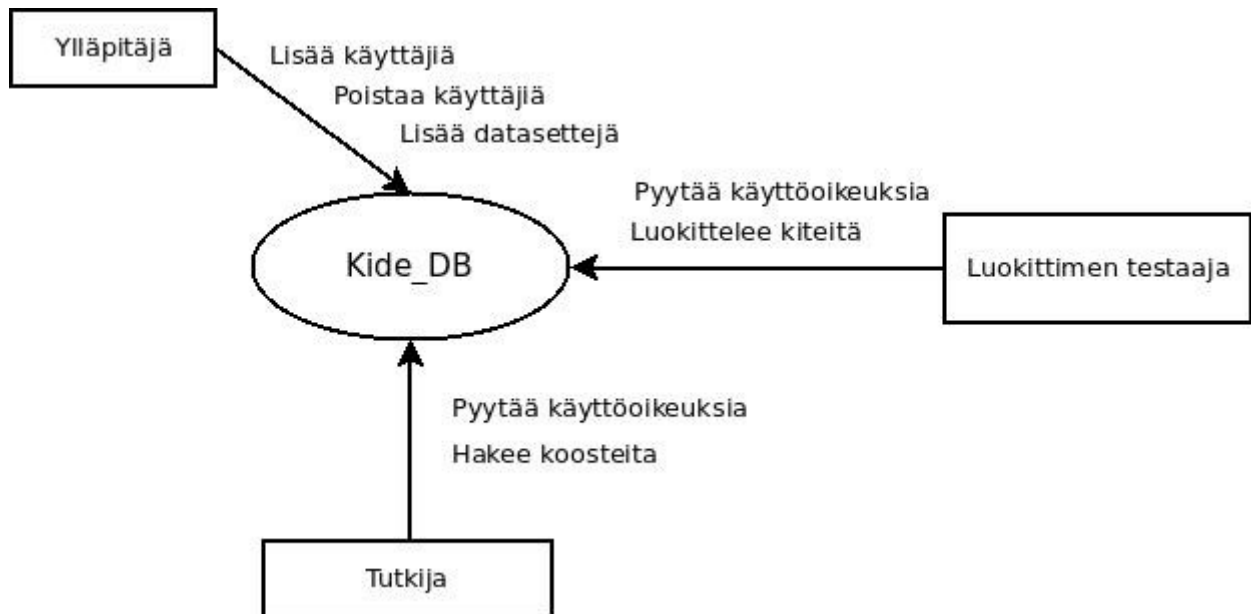
Kehitysmahdollisuuksia ovat muunmuassa automaattinen käyttäjien hallinta, csv-latauksen ohella tekstimuotoinen datan esikatselu ja koko datasettiä koskevan statistiikan näyttäminen.

Toteutusympäristö

Työ toteutetaan käyttäen Netbeans IDEä ja testaamalla sitä työn edetessä suoraan Users-palvelimella.

Yleiskuva Järjestelmästä

Järjestelmän sidosryhmät:



käyttäjärühmien esittely

Tutkijat haluavat saada pääsyn järjestelmään ja koota datasta yksilöityjä koosteita jotka ovat relevantteja heidän tutkimuksensa kannalta.

Luokittimen testaaja haluaa varmistua siitä, että automaattinen luokitin on luokitellut kiteet järjestelmässä oikein. Hän haluaa manuaalisesti luokitella jonkin määrän kiteitä ja sen jälkeen verrata näitä järjestelmässä oleviin luokituksiin.

Ylläpitäjä haluaa hallita järjestelmän käyttäjiä ja lisätä sekä poistaa dataa järjestelmästä.

Käyttötapaukset

Pyydä käyttöoikeuksia

Uusi käyttäjä pyytää käyttöoikeuksia järjestelmään. Toteutamme käyttöoikeuksien pyynnön sähköpostilomakkeella kirjautumissivulta.

Hae kooste

Käyttäjä haluaa koosteita kidetietokannasta valitsemillaan parametreillä. Tämä on järjestelmämme pääasiallinen käyttötarkoitus. Koosteet ladataan csv-muodossa.

Kiteiden luokittelu käsin

Käyttäjä haluaa tietoa luokittimen tarkkuudesta esimerkiksi eri mittauspaikoilta (site) saatujen kiteiden automaattiluokituksesta. Tätä varten käyttäjä määrittelee valitsimilla luokiteltavat kiteet ja luokittelee niitä haluamansa määrän käsin. Luokituksen tuloksia voi tarkastella tulossivulta.

Lisää datasettejä

Järjestelmän ylläpitäjä saattaa haluta lisätä datasettejä jos uusia tutkimustuloksia tulee saataville. Tämä on kuitenkin niin harvinainen ja tilanteesta riippuva käyttötapaus että ylläpitäjän tulee tehdä se siihen tarkoitetuilla erikseen saatavilla Matlab-työkaluilla.

Lisää käyttäjä

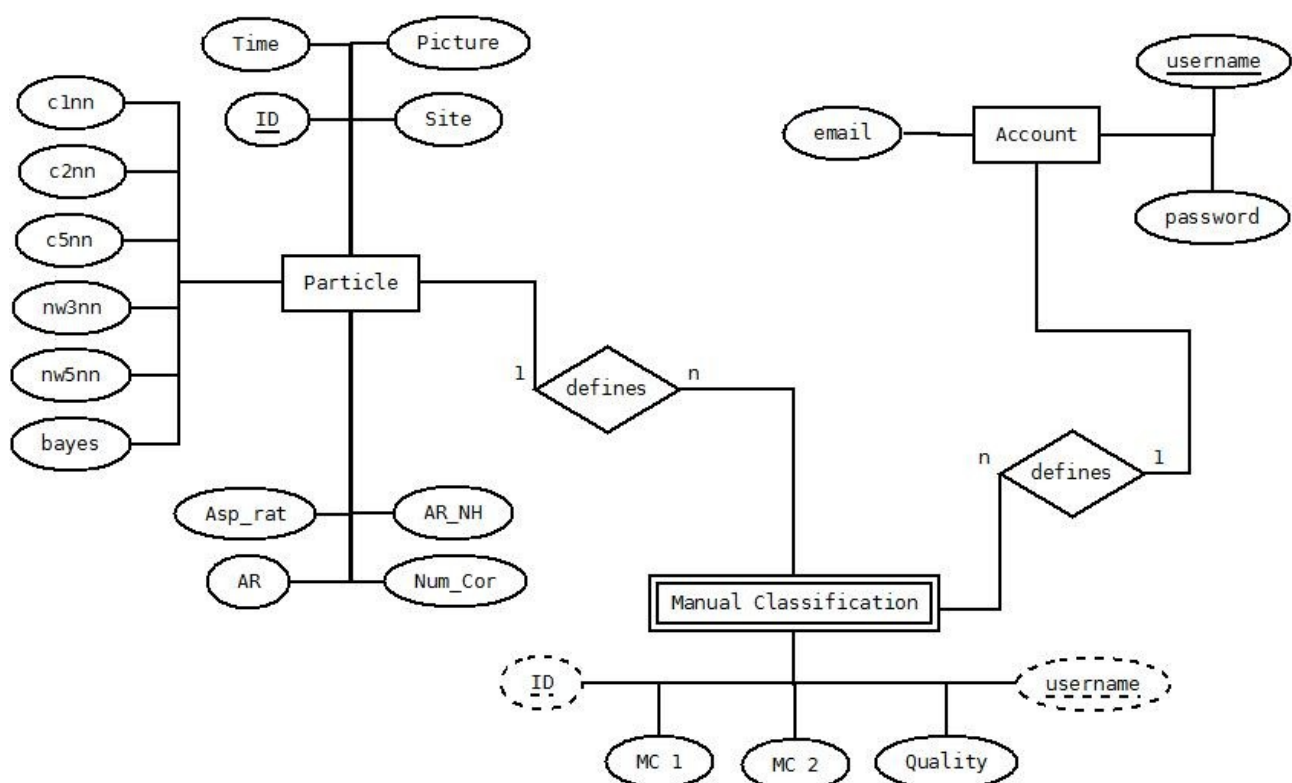
Ylläpitäjä lisää tietokantaan käyttäjän tämän lähettämän sähköpostin perusteella. Tämäkin käyttötapaus hoidetaan järjestelmässä manuaalisesti koska järjestelmän ylläpitäjä haluaa olla tietoinen järjestelmän käyttäjistä.

Poista käyttäjä

Ylläpitäjä haluaa poistaa käyttäjätunnukset joille ei enää ole käyttöä. Hoidetaan manuaalisesti.

Järjestelmän tietosisältö

ER kaavio



Tietokohteiden kuvaukset

Kiteen attribuutit

| <i>Attribuutti</i> | <i>Arvojoukko</i> | <i>Kuvaus</i> |
|--------------------|-------------------|--|
| id | varchar(50) | Primary key: Kiteen tunnistukseen käytettävä, kuvan tiedostonimestä generoitu tunniste joka toimii myös taulun avaimena. |
| Time | timestamp | Kertoo mittauksen tekoajan. |
| c1nn | nn | PCA luokitus yhden lähimmän naapurin mukaan |
| c3nn | nn | PCA luokitus kolmen lähimmän naapurin mukaan. |
| c5nn | nn | PCA luokitus viiden lähimmän naapurin mukaan. |
| nw3nn | nwn | Painottamaton PCA luokitus kolmen lähimmän naapurin mukaan. |
| nw5nn | nwnn | Painottamaton PCA luokitus viiden lähimmän naapurin mukaan. |
| bayes | bayesian | Bayesiläisen algoritmin luokitus. |
| dmax | real | Kiteen maksimi halkaisija. |
| ar | real | Kiteen pinta-ala verrattuna ympyrään jolla on sama maksimi halkaisija. |
| ar_noholes | real (0 < x < 1) | Sama kuin yllä, mutta mahdolliset reiät on täytetty. |
| asprat | real (0 < x) | Maksimi pituus jaettuna maksimi leveydellä. |
| n_corners | int | Kulmien määrä kiteessä |
| site | site | Kertoo miltä mittauspisteeltä kide on peräisin. |

Arvojoukot ovat määritelty seuraavasti:

nn (Enum(8)):
”B” = Bullet

”P” = Plate
 ”C”= Column
 ”I”= Irregular
 ”PA”= Plate Aggregate
 ”CA”= Column Aggregate
 ”R”= Rosette
 ”RA”= Rosette Aggregate

nwnn (Enum(9)):
 sama kuin yllä ja lisäksi
 ”ucl”= unclassified

bayesian (Enum (9)):
 Sama kuin nn ja lisäksi
 ”NaN” = Not a Number

site (Enum(6)):
 ”AAF” = ARM Aerial Facility
 ”AMF” = ARM Mobile Facility
 ”NSA” = North Slope of Alaska
 ”SGP” = Southern Great Plains
 ”TWP” = Tropical Western Pacific
 ”other” = Some other site.

Users taulun attribuutit

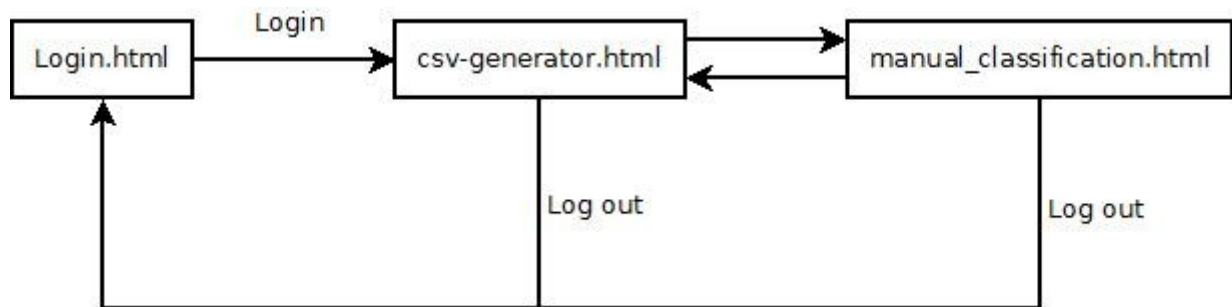
| <i>Attribuutti</i> | <i>Arvojoukko</i> | <i>Kuvaus</i> |
|--------------------|-----------------------|---------------------------|
| username | varchar(30) | Kyttäjänimi, Primary key. |
| password | varchar(30) | Salasana. |
| email | Varchar(50), nullable | Sähköpostiosoite |

Manuaalisen luokituksen attribuutit

| <i>Attribuutti</i> | <i>Arvojoukko</i> | <i>Kuvaus</i> |
|--------------------|-------------------|---|
| kide_id | Varchar(50) | Viittaa attribuuttiin id kide taulussa. Avain. |
| classified_by | varchar(30) | Viittaa users taulun attribuuttiin username. Avain. |
| class1 | nn | Käyttäjän ensisijainen luokitus kiteelle. |
| class2 | nn | Käyttäjän toissijainen luokitus kiteelle. |
| quality | boolean | Käyttäjän arvio onko kide hyvälaatuinen. |

Käyttöliittymän hahmotelma

Login ruudusta on mahdollista lähettää sähköpostilomakkeella pyyntö käyttöoikeuksien saamiseksi. Epäonnistunut kirjautumisyritys johtaa virheilmoitukseen. Csv-generator sivulta käyttäjä voi ladata koosteita ja manual_classification sivulla voi luokitella kiteitä.



Relaatiotietokantakaavio

Tässä tietokannan taulut ja niiden relaatiot.

