

# 1. JOHDANTO

## Järjestelmän tarkoitus

Kerätä tviittejä reaaliaikaisesti, mutta [Twitterin API:n](#) rajoitukset huomioiden käytännössä "ryöpyissä".

## Tiivis kuvaus siitä mistä on kyse

Verkkoselaimesta käytettävä Twitter-ryömiä, jolle voi antaa tiettyjä hashtageja tai avainsanoja kerättäväksi tietyllä maantieteellisellä tai muulla rajauksella. Sovellus ei ainakaan beta-versiona tule olemaan avoimessa verkossa, vaan se on tarkoitettu suljettuun käyttöön pienille tutkimustiimeille.

## Millaisen toiminnan tukemiseen järjestelmä on tarkoitettu.

Datajournalistien ja tutkijoiden käyttöön.

## Mitkä ovat järjestelmän tavoitteet

Kerätä tviittejä tietokantaan.

## Toteutus-/toimintaympäristö

Itse ryömiä kirjoitetaan hyödyntäen [muiden esimerkkejä](#) ja Pythonin [Beautiful Soup -kirjastoa](#) ja [Cronjobsia](#) ”ryöppyjen” ajastamiseksi. Verkkosivu toteutetaan käyttäen HTML:ää ja CSS:ää [Bootstrapilla](#) sekä mahdollisesti [Jekyllillä](#) (HTML) tai [Twigillä](#) (PHP). Mallit, näkymät ja kontrolleri kirjoitetaan PHP:llä kurssin ohjeiden mukaisesti.

## Missä ympäristössä työ toteutetaan (yleensä laitoksen users-palvelimella Tomcat- tai Apache- palvelimen alla)

Laitoksen users-palvelimella.

## Täytyykö web-sovelluksen alustajärjestelmän tukea jotain tiettyä ohjelmointikieltä. (esim. Java, Ruby, PHP..?)

Mahdollisesti Pythonia ja PHP:ta, en ole vielä varma missä määrin sovellus tulee vaatimaan selaintukea näille kielille vaikka käytän niitä.

## Täytyykö käyttäjän selaimen tukea jotain tiettyä ohjelmointikieltä (esim. javascript?)

En ole vielä varma missä itse ryömiä toimii, käyttäjän koneella vai sovelluksen käytössä olevilla palvelimilla. Jos käyttäjän koneella niin koneen on tuettava Pythonia. Muuten mahdollisesti javascriptia ja jQuerya, jos päädyn tuunaamaan graafista käyttöliittymää näillä.

## Edellyttääkö ohjelmisto jonkun tietyn tietokannan käyttöä vai voiko sitä vaihtaa helposti. Useimmat työt toimivat vain yhdellä kannalla

Tarkoitus olisi, että ohjelmisto voidaan siirtä toimimaan myös toisilla tietokannoilla. Esimerkiksi datajournalismiin erikoistunut [Knightlabs ohjeistaa](#) säilömään Twitter-ryömiöiden keräämää dataa [MongoDB](#)-tietokannassa.

## **2. KÄYTTÖTAPAUKSET**

### **A. Käyttäjät**

#### **Tilaa työ**

Käyttäjä voi tehdä sovellukselle tilauksen työstä, eli pyytää ryömijää noutamaan dataa tietyin rajoituksin.

#### **Rajaa hakua**

Rajauksiin sisältyvät ainakin hashtagit, maantieteellinen sijainti ja tviittien aika sekä mahdollisesti myös tviittien kirjoittajat eli tviittaaajat.

#### **Lataa tietokanta, tilaa "näkymä"**

Käyttäjät voivat tilata valmiista tietokannasta "näkymiä" tiettyjen rajausten perusteella tai ladata tietokannan kokonaisuudessaan CSV- ja XLS-muodoissa.

### **B. Ylläpito**

#### **Hyväksy työ**

Jotta järjestelmä ei ylikuormitu, ylläpidon on hyväksyttävä kaikki tilaukset.