

Traapija: Tietosisältö ja relaatiot

ATTRIBUUTIT, ARVOJOUKOT JA KUVAILUT

Tietokannan tietosisältö perustuu Twitteristä haettuun dataan. Twitterin kehittäjillä suunnatuilla sivuilla on perusteelliset kuvailut kaikkien tähänkin tietokantaan haettavista objekteista, sisältäen myös niiden attribuutit ja arvojoukot. Tämän dokumentin loppuosasta löytyvästä relaatiotietokantaakaaviosta löytyy myös siihen sisältyvien objektien attribuuttien arvojoukot sekä arvio siitä, mikä vastaava arvojoukko olisi PostgreSQL-tietokannassa.

Twitterin käyttämiin objekteihin voi tutustua seuraavista linkeistä:

- <https://dev.twitter.com/overview/api/users>
- <https://dev.twitter.com/overview/api/tweets>
- <https://dev.twitter.com/overview/api/entities>
- <https://dev.twitter.com/overview/api/entities-in-twitter-objects>
- <https://dev.twitter.com/overview/api/places>

TWITTER-OBJEKTIT TÄSSÄ TIETOKANNASSA

User

User, eli käyttäjä on Twitterin toimija. Käyttäjät twiittaavat, seuraavat toisiaan ja tekevät listoja, Heillä on oma aikajana, toiset käyttäjät voivat seurata heitä ja he voivat seurata toisia käyttäjiä.

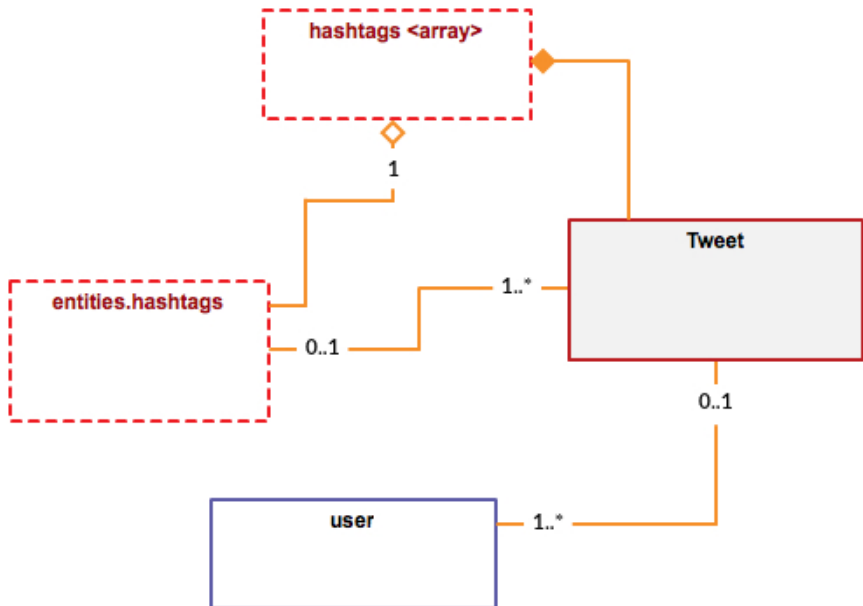
Tässä tietokannassa käyttäjiä käsitellään heidän tekemien twiittien perusteella. Kun Traapia noutaa twiitin se noutaa myös tietoja käyttäjistä. Kaikki tietokantaan sisältyvät käyttäjät koostetaan omaksi listakseen.

Tweet

Twitterin ytimessä ovat twiitit. Vaikka Traapija seuloo twiittejä hashtagien avulla, twiitit ovat tärkein objekti tietokannassa. Twiitit säilötään tietokannassa hashtagien perusteella.

Hashtag

Hashtagit ovat Traapijan ainut ja tärkein periaate seulomiselle. Twiittejä noudetaan Traapijan käyttäjän asettamin hakukriteerein, jotka perustuvat aina yhdelle tai useammalle hashtagille. Tietokannassa ylläpidetään yhtäältä lista kaikista hashtageista ja toisaalta twiitit on järjestetty listaksi sen mukaan, mihin hashtagiin tai hashtageihin ne sisältyvät.



hashtags <array>

+<key>text (String / text)
+<fkey>tweets.id (int64 / bigint)

entities.hashtags

+indices (Array of int / integer [])
+<key, fkey>text (String / text)

Tweet

+coordinates (Collection of Float / row_to_json)
+created_at (String / text)
+<fkey>entities.hashtags.text (String / text)
+entities.url.display_url (String / text)
+entities.user_mentions
+favorite_count (int / integer)
+<key>id (int64 / bigint)
+in_reply_to_status_id (int64 / bigint)
+in_reply_to_user_id (int64 / bigint)
+lang (String / text)
+place.attribute (Object / row_to_json)
+place.country (String / text)
+place.full_name (String / text)
+place.id (String / text)
+place.type (String / text)
+ place.url (String / text)
+quoted_status_id (int64 / bigint)
+quoted_status.text (String / text)
+retweet_count (int / int)
+retweeted_status (???)
+source (String / text)
+text (String / text)
+<fkey>user.id (int64 / bigint)

user

+created_at (String / text)
+description (String / text)
+followers_count (int / bigint)
+friends_count (int / bigint)
+geo_enabled (Boolean / Boolean)
+id (int64 / bigint)
+lang (String / text)
+listed_count (int / int)
+location (String / text)
+name (String / text)
+protected (Boolean / Boolean)
+screen_name (String / text)
+statuses_count (int / int)
+time_zone (String / text)
+url (String / text)