

Tekstianalytiikka Helsingin kaupungilla

Mikko Niemi

Helsinki

Mitä tekstianalytiikka on

Tekstianalytiikka

- Luonnollisen kielen käsittely edistynein menetelmin (Natural Language Processing)
 - Tekstien ryhmittely, luokittelu, tuottaminen, tunnesisältöjen havaitseminen
 - Datapohjaiset menetelmät ja säännölliset kieliovit
 - Datapohjaiset menetelmät voivat perustua neuroverkkoihin tai muihin tilastollisiin menetelmiin
-
- Yleensä tekstianalytiikan mallit eivät sisällä käsitteellistä ymmärrystä
 - Mallit ovat yleisiä sääntöjä, eivät pureudu yksittäistapauksiin

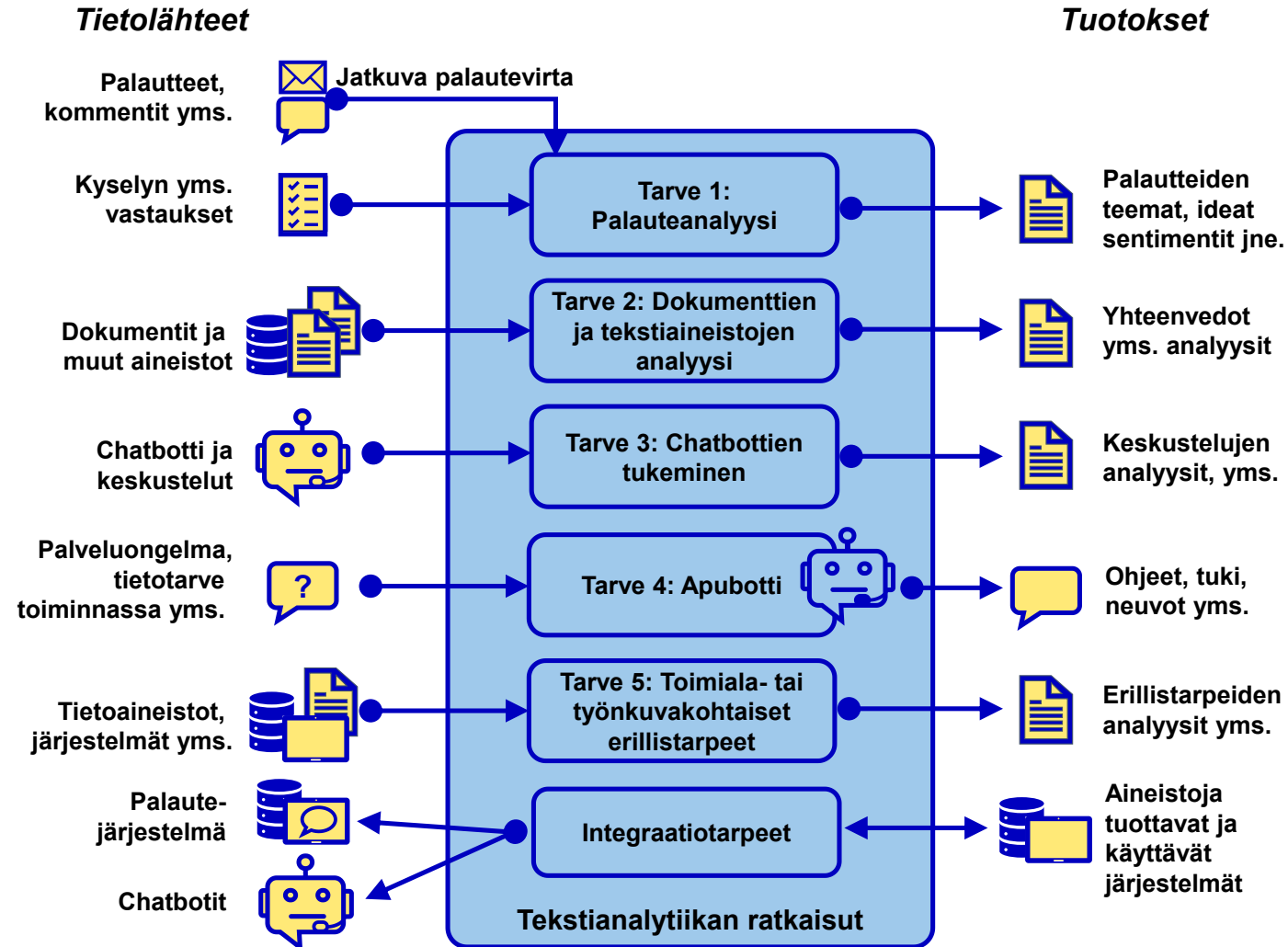
Tekstianalytiikka Helsingin kaupungilla

Tekstianalytiikka kaupungilla

- Palautteiden ja kyselyiden jatkuva analyysi valmiin työkalun avulla
- Useita kokeiluja kokeilukiihdyttämössä (kaupungin sisäinen toiminto nopeille digitalisaatiokokeiluille)
- Erilaista pienkehitystä
- Kuva: Kokeilukiihdyttämö - Tekoälyn hyödyntäminen Kerrokantasi vastausten luokittelussa ja analysoinnissa



Tarpeiden yleiskuva



Keskeisiä aineistoja

- Kaupunkilaispalautteet
- Palautekyselyt
- Toimintaan liittyvät asiakirjat kuten asemakaavat tai palotarkastuspöytäkirjat
- Operatiivisten järjestelmien tekstidata
- Päätöstekstit
- Talouden ja hankinnan asiakirjat

Omastadin budjettiehdotuksien analyysi

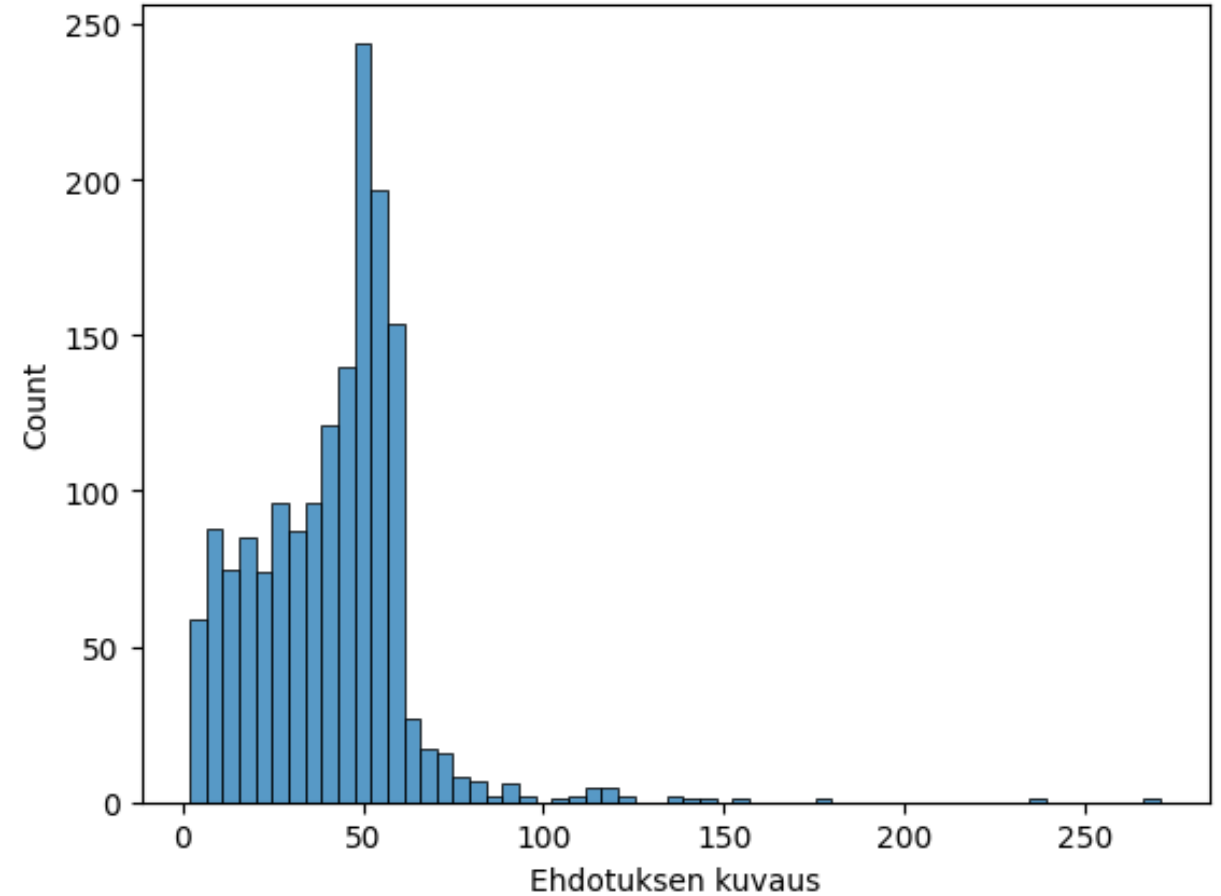
Tutkimuskysymys

- Budjettiehdotuksien ymmärtäminen kokonaisuutena - etäluenta
- Kiinnostavaa ainakin aihemallinnus ja klusterointi
- Ehdotusten automaattinen ryhmittely

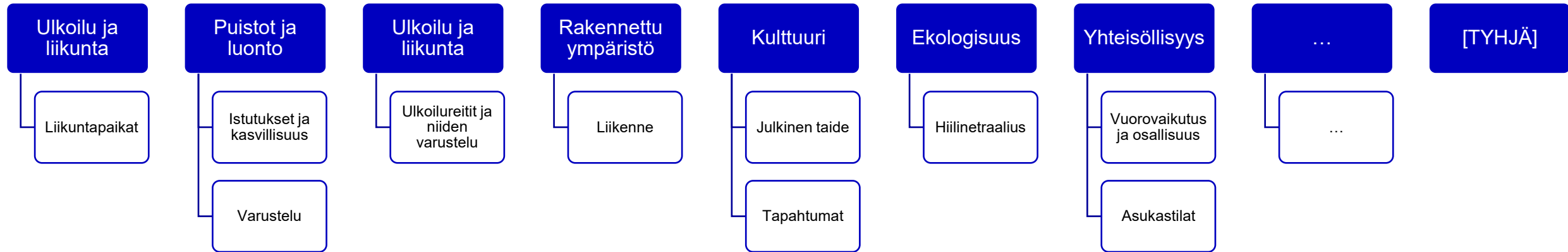
- Analyysi on keskeneräinen

Aineisto

- Omastadin kierroksen 2023 budjettiehdotukset
- Hyväksytyt ja hylätyt ehdotukset
- Otsikko ja sisältö
- Esikäsittelynä perusmuotoistaminen ja yleisimpien sanojen poisto



Käsin laaditut teemat



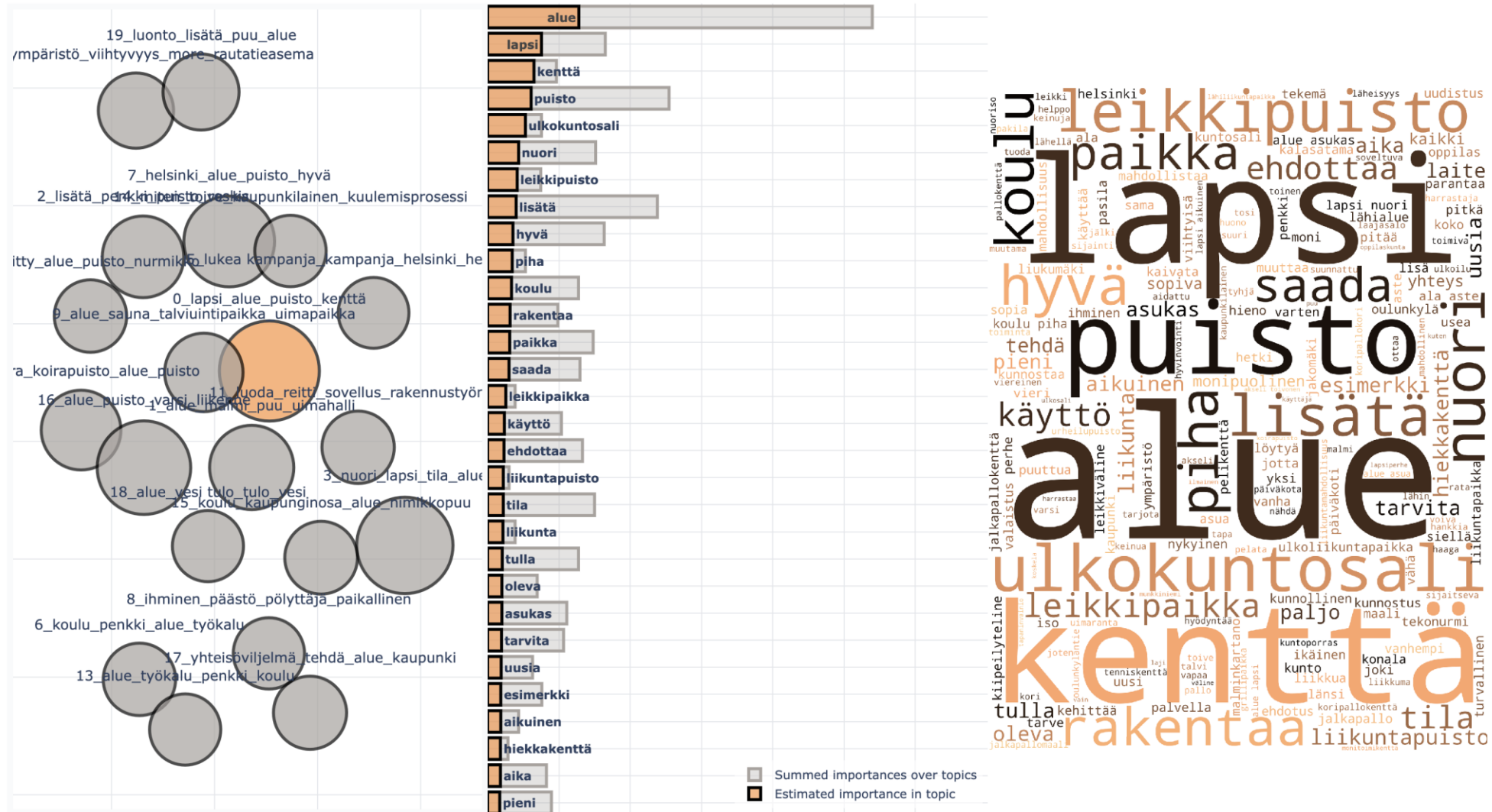
Aihemallinnus

- Lyhyille teksteille soveltuvia aihemallinnuksen ja klusteroinnin menetelmiä
- DMM, BTM
- Kirjastoina tweetopic ja topicwizard

- Sopiva aiheiden määrä arvioitava
- Seuraavassa alustavia tuloksia 20 aiheen DMM-mallista
- Bi-gram-malli

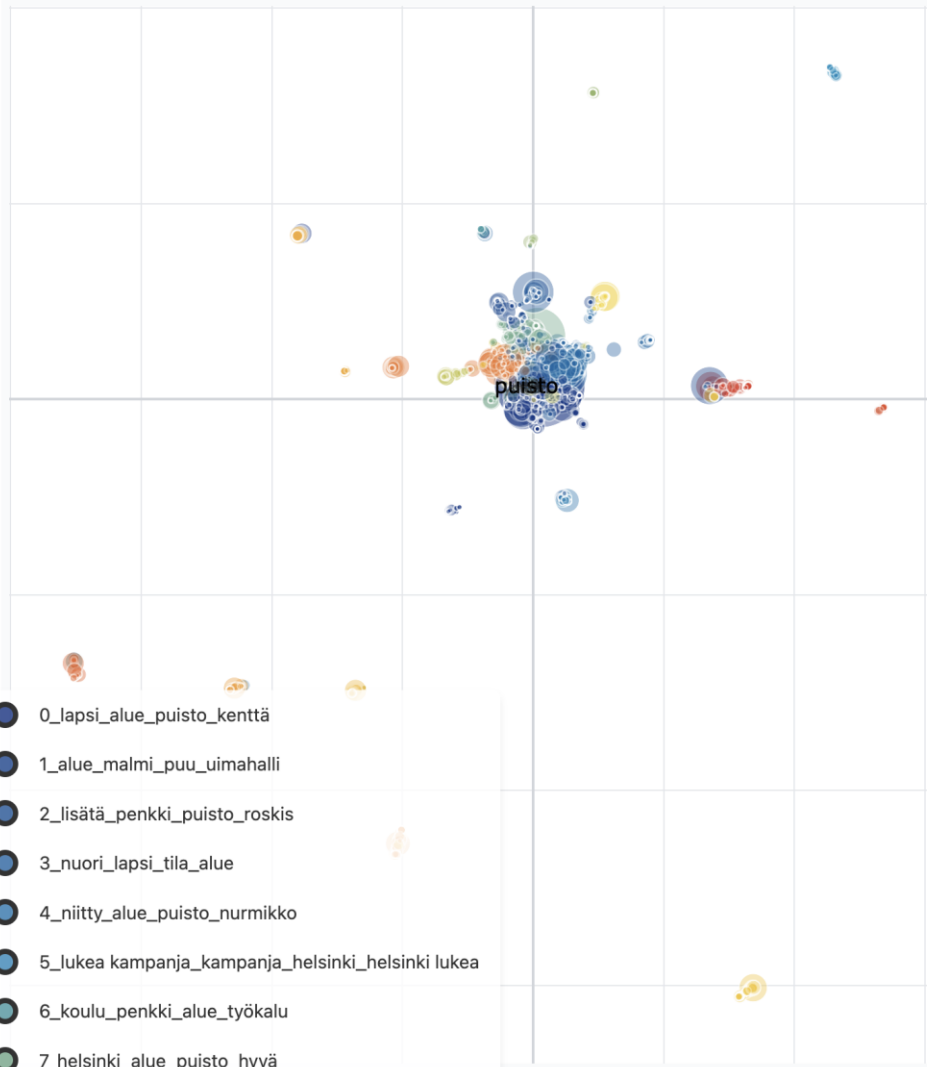
Alustavia tuloksia

- Puistot, ulkokuntosalit, leikkipuistot (349 ehdotusta)
 - alue puisto lapsi kenttä ulkokuntosali leikkipuisto hyvä nuori paikka rakentaa
- Saunat, uimapaikat (104 ehdotusta)
 - alue helsinki uimaranta sauna uimapaikka talviuintipaikka kaupunki koira hyvä meri
- Toiminta kouluissa ja päiväkodeissa (9 ehdotusta)
 - kampanja helsinki helsinki lukea lukea kampanja lukea päiväkotikoulu päiväkotikoulu työpaja kirjasto
- Puut (24 ehdotusta)
 - kaupunginosa alue nimikkopuu tulla kaupunki hyvä puu asukas kaupunginosa nimikkopuu puulaji



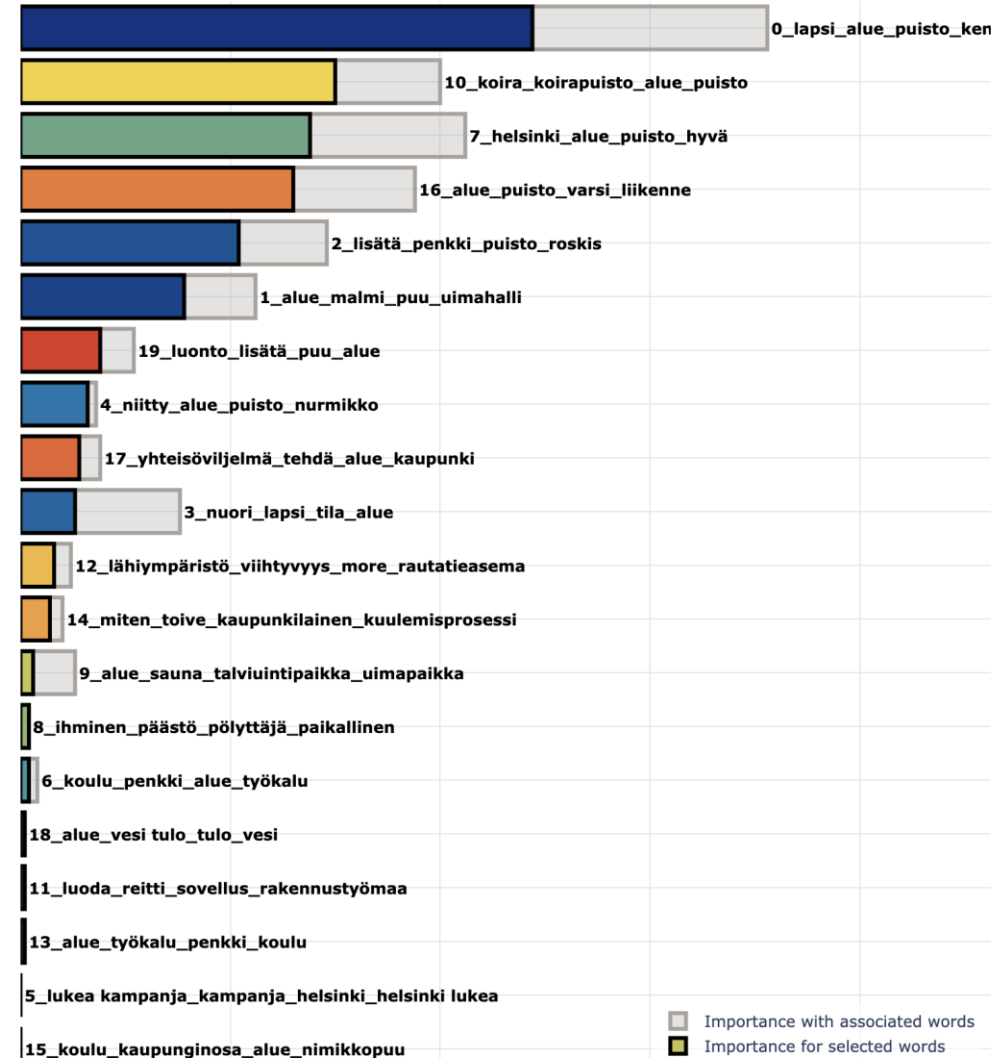


x puisto



- 0_lapsi_alue_puisto_kenttä
- 1_alue_malmi_puu_uimahalli
- 2_lisätä_penkki_puisto_roskis
- 3_nuori_lapsi_tila_alue
- 4_niitty_alue_puisto_nurmikko
- 5_lukea_kampanja_kampanja_helsinki_helsinki_lukea
- 6_koulu_penkki_alue_työkalu
- 7_helsinki_alue_puisto_hyvä
- 8_ihminen_päästö_pölyttävä_paikallinen
- 9_alue_sauna_talviuimaupaikka_uimaupaikka

ASSOCIATIONS:



Importance with associated words
Importance for selected words

Topics

Words

Documents



Helsinki

Etunimi Sukunimi

Yhteenveto

Yhteenveto

- Kaupunki on soveltanut tekstianalytiikkaa erityisesti palautteiden ja kyselyiden analysointiin
- Jatkuvia analyyseja ja yksittäisiä kokeiluja
- Kehitämme edelleen edellytyksiä tekstianalytiikan käyttöön