SimpleStats kokotekstien lataustilastointi ja sen JSON-rajapinta (DSpace-versio 5)

Simplestats on Kansalliskirjaston tekemä tilasto-ohjelma, joka pitää kirjaa DSpacen tietueiden latausluvuista. Latauslukuja voi tarkistella koko arkiston, tietyn yhteisön tai kokoelman tasoilla. Simplestats on asennettu jokaiseen Kansalliskirjaston Dspace-instanssiin, kuten Doriaan ja Theseukseen. **Näissä ohjeissa esimerkkinä käytetään Doriaa, mutta niitä voi soveltaa myös muille instansseille.**

Simplestatsin URL: http://www.doria.fi/simplestats/

(Tuolla näkyvät luvut päivitetään kerran vuorokaudessa, joten ne ovat 0-24 tuntia jäljessä reaaliaikaisesta tilanteesta. Varsinaisen Dorian puolella kunkin julkaisun sivulla näytettävä latausluku on jäljessä 0-48 tuntia reaaliaikaisesta tilanteesta.)

Simplestatsin käyttöliittymä

Käyttöliittymä antaa etusivulla ylimmän tason kokoelmahierarkian eli listan yhteisöjä, joista voi valita tarkasteluun halutun yhteisön tai koko julkaisuarkiston. Sivun alalaidasta voi valita tarkasteltavien latauslukujen aikavälin. Valintojen jälkeen "*Show statistics*" -nappia painamalla suoritetaan tilastojen haku.

Avautuvasta listauksesta näkyy kokoelmarakenne, josta voi klikata edelleen yhteisöä tai kokoelmaa nähdäkseen kyseisestä aineistosta latausluvut tietueen tarkkuudella. Listauksessa näytetään myös yhteisön ja kokoelman nimen perässä itemien määrä, tiedostojen (yleensä siis kokotekstien) määrä ja niiden yhteenlaskettu koko. Tässä listauksessa (koko arkiston tai valitun yhteisön tasolla) näytetään myös top 10 -lista ladatuimmista itemeistä.

Tietueen tarkkuudella avautuvassa taulukkomuotoisessa listauksessa tietueet ovat omilla riveillään ja sarakkeilla on kunkin kuukauden latausluvut. Viimeisessä sarakkeessa on itemin yhteenlaskettu latausluku tarkastelussa olevalta aikaväliltä.

Simplestatsin toimintaperiaate

Simplestats hakee öisin päivän aikana tulleet lataustapahtumat Dspacen lokista ja yhdistää lataustapahtumien tiedot Dspacen kokoelmarakenteeseen. Tiedot kuvaavat siten edellisen päivän tilannetta. Simplestatsin omassa tietokannassa ei ole Dspacen täydellisiä kuvailutietoja, joten sen avulla voi tehdä vain kokoelmarakenteeseen ja aikaväliin perustuvia tilastojaotteluja.

Simplestatsin JSON-rajapinta

Simplestatsin latauslukuja voi hyödyntää myös esim. omilla verkkosivullaan REST-tyyppisellä rajapinnalla, jossa kyselyt annetaan URL-muodossa ja vastaukset saadaan JSON-formaatissa. Yleisin tapa on käyttää tietoja JavaScriptin avulla.

URL:ään etusivun_url/simplestats/rest, esim. http://doria.fi/simplestats/rest voi antaa seuraavia parametreja:

handle: Kokoelman tai yhteisön handle, jonka tilastot haetaan. Pakollinen parametri, arvolla 0 saadaan koko Dorian tilastot.

Top-listauksia ohjaavat, vapaaehtoiset parametrit:
top: Kuinka monta top-itemiä listataan
top_start_time: Aloituskuukausi mistä top-lista lasketaan (jos tätä parametria ei anneta, niin aloituskuukausi aikaisin mahdollinen)
top_end_time: Lopetuskuukausi mistä top-lista lasketaan (jos tätä parametria ei anneta, niin lopetuskuukausi myöhäisin mahdollinen)

Käyttöesimerkkejä

Latauslukudata ilman top-listaa A. E. Nordenskiöldin irtokarttoja -kokoelmasta:

```
http://doria.fi/simplestats/rest?handle=10024/97216
```

Alkuvuoden top 20 samasta kokoelmasta:

```
http://doria.fi/simplestats/rest?handle=10024/97216&top=20&top_start_time=201301&top_end_time=201306
```

```
http://doria.fi/simplestats/rest?handle=0&top=100&top_start_time=201301
```

```
http://doria.fi/simplestats/rest?handle=0&top=10
```