

OpenSearch ja syötteet (RSS ja ATOM) (DSpace-versio 5)

OpenSearch-rajapinta

OpenSearch-rajapinnan avulla voidaan hakea URL:ään upotettujen [Lucene-hakusyntaksilla](#) tehtyjen hakujen mukaisia kyselyjä ja vastaus saadaan XML:nä.

OpenSearch -rajapinnan kautta julkaisuarkistossa tehtyjen hakujen tuloksia saa vastauksena **RSS-** tai **Atom-syötteinä**, joiden avulla julkaisuista on mahdollista tehdä automaattisesti päivittyviä listauksia esim. omille kotisivuille, muodostaa hakuja RSS-readereiden käyttöön yms. Kansalliskirjaston DSpace-arkistoihin on lisätty syötemuotoisten vastausten lisäksi **kk-** ja **kkf** formaatit, joilla OpenSearch haut saa XML-muodossa, jossa on tulostietueiden kaikki saatavilla oleva metadata näkyvillä. Jälkimmäisellä kkf-formaatilla saa lisäksi tiedostlinkit varsinaisiin tiedostoihin, kuten pdf:iin.

Oheassa OpenSearchin syntaksin esittelyä tarkemmin Doria-julkaisuarkistoon kohdistuvien hakuesimerkkien avulla. Ohjeet soveltuvat myös muihin Kansalliskirjaston ylläpitämiin julkaisuarkistoihin:

[Jukuri](#), [Julkari](#), [Kaisu](#), [Lauda](#), [LutPub](#), [Osuva](#), [Taju](#), [Trepo](#), [Theseus](#), [UTUPub](#), [Valto](#), [Fenno-Ugrica](#), [Fragmenta membranea](#). Korvaa vain <https://www.doria.fi> -alkuosa jonkin toisen julkaisuarkiston etusivun osoitteella:

```
Julkaisuarkisto/open-search/?query=...
```

OpenSearchin käyttöesimerkki: Tilataan MPKK:n kokoelmiin kohdistetun haun "sukellusvene" -hakutulokset RSS-syötteenä:

```
https://www.doria.fi/open-search/?query=sukellusvene&sort_by=3&order=desc&start=0&rpp=50&format=atom&scope=10024/72633
```

- **query**-parametrin perään voidaan tehdä haku samaan tapaan kuin Dspaceissa muutenkin, eli [Lucene-hakumootorin syntaksilla](#). Voidaan siis esim. kohdistaa haku tiettyyn kenttään hakuliaksien avulla, eli esim. title:sitäjatätä. Julkaisuarkistojemme kenttähakuun tarvittavat aliakset löytyvät [täältä](#). Sulkujen käyttöä ei kannata säästellä monimutkaisemmissa booleanhauissa.
- **sort_by**-parametrilla voi järjestää syötteen seuraavasti: 0 = Relevanssin, 1 = Nimekkeen, 2 = Julkaisuajan, 3 = Syöttöajan mukaan
- **order**-parametrilla voi muuttaa järjestyksen nousevaksi (asc) tai laskevaksi (desc)
- **start**-parametri kertoo, mistä hakutuloksesta aloitetaan näyttäminen.
- **rpp**-parametrilla kerrotaan, kuinka monta tulosta maksimissaan listataan per sivu.
- **format**-parametri voi olla joko atom, rss tai kk ja kkf. (kk ja kkf ovat Kansalliskirjastolla tehtyjä formaatteja, jotka palauttavat kaikki julkisesti saatavilla olevat metadatat tulostietueista. Formaateista kkf palauttaa myös tiedostolinkit, esimerkiksi kokotekstien ja kuvien linkit.)
- **scope**-parametrilla voi rajoittaa haun tiettyyn kokoelmaan tai yhteisöön, arvoksi annetaan kokoelman/yhteisön handle. Jos scopea ei määritellä, haku kohdistuu koko Dspace-arkistoon.

HUOM! Jotta saat kaikki tulokset, täytyy hakuja tehdä useita. Hakutuloksia iteroidaan kasvattamalla **start-parametria** **rpp-parametrin** verran. Kun tulosjoukko on tyhjä tai pienempi kuin **rpp:n** arvo, voidaan päätellä, että kaikki tulokset on iteroitu läpi.

Muita hakuesimerkkejä Doriasta.

Haussa voi käyttää fraaseja:

```
https://www.doria.fi/open-search/?query="vietnamin sota"
&sort_by=3&order=desc&start=0&rpp=50&format=atom&scope=10024/72633
```

MPKK:n julkaisemia koptereita tai panssarivaunuja käsitteleviä artikkeleita RSS-feedinä:

```
https://www.doria.fi/open-search/?query=publisher:maanpuolustuskorkeakoulu AND (abstract:helikopteri* OR abstract:panssariv*)&start=1&rpp=50&format=rss
```

MPKK:n julkaisemia julkaisuja, joissa mainitaan Israel tai Palestiina:

```
https://www.doria.fi/open-search/?query=publisher:maanpuolustuskorkeakoulu AND (abstract:israel OR abstract:palestiina)&start=1&rpp=100&format=kk
```

Yksi tapa käyttää Open Search -kyselyitä on siis integroida tuloksena saatava RSS-, Atom- tai XML-muotoinen tieto osaksi omia kotisivuja.



Koska OpenSearchin tuottama RSS-, Atom tai XML-muotoinen tieto ei sisällä HTML-muotoilutietoja, pitää syötteiden halutusta ulkoasusta huolehtia web-sivuston puolella. Tämän toteutus riippuu kotisivujen ylläpitoon käytetystä ohjelmistosta.

Syötteiden (feed) tilaus (RSS ja ATOM)

Syötteitä voi tilata kokolma-, yhteisö- tai julkaisuarkistotasolla.

Koko arkiston syötteiden haku rss:nä ja atomina:

```
https://www.doria.fi/feed/rss_2.0/site  
https://www.doria.fi/feed/atom_1.0/site
```

Haun kohdistaminen tiettyyn kokoelmaan tai yhteisöön (käytä yhteisön/kokoelman handlea):

```
https://www.doria.fi/feed/rss_2.0/10024/4194 http://www.doria.fi/feed/atom_1.0/10024/4194
```

Haku palauttaa niin monta työtä kuin DSpace-arkistossa on määritelty, määrä vaihtelee neljän ja seitsemän välillä. Tulokset ovat uutuusjärjestyksessä kentän dc.date.accessioned mukaan. Syötteet ovat kuitenkin huomattavasti rajoittuneempia verrattuna OpenSearch-rajapinnan käyttöön. Syöte-hakuja voi rajata vain kokoelmahierarkian perusteella tai kohdistaa ne koko arkistoon. Open Search -rajapinta on tässä suhteessa paljon monipuolisempi, ja sillä on mahdollista tuottaa esim. verkkosivuille päivittyviä uutuuslistoja aineistosta myös metadatasta tehtävien hakujen pohjalta, eli se ei ole sidoksissa kokoelmarakenteeseen.