



# Elasticsearch

uusi hakukone

Anneli Österman  
Pääkäyttäjä, Koha-Suomi Oy

# Taustaa

- Kaikki tiedonhauk kohdistuvat indeksiin ja hakukoneena käytetään Elasticsearchiä.
  - Vaski toimi Elasticsearchin käytössä pilottina
- Nykyinen Koha-versio ei tue täysin Elasticsearchia, joten puutteita on vielä. Kaikkia Elasticsearchin toiminnallisuuksia ei saada käyttöön kuin vasta versionpäivityksen yhteydessä.
  - Esimerkiksi Elasticsearchissa ei ole yhtä kattava tuki CCL eli Common Command Language Searching –hakuprotokollalle kuin Zebrassa
  - Myös fasettien määrittelyssä puutteita

# Taustaa

- Tiedonhaun hitaus johtuu pääsääntöisesti saatavuustietojen laskemisen hitaudesta. Hakukoneen vaihto Zebrasta Elasticsearchiin ei tuo siihen merkittävää muutosta.
  - Koha-Suomella on selvityksen alla eri keinoja saatavuuslaskennan nopeuttamiseen tai sen rajoittamiseen, esim. Finnan käyttämä Lazy load – toiminto.

# Mitä ne indeksit ovat?

## Indeksi

Tietokanta

Tiedonhakija tekee haun indeksiin ja indeksi palauttaa listan teoksista, jotka vastaavat hakua.

### Sanahaku (kw)

- sisältää kaiken indeksoidun erottelematta, mikä termi on
- teokset

### Tekijät (au)

- Päätaalo, Kalle
- teokset
- Lindgren, Astrid
- teokset

### Nimekkeet (ti)

- Seitsemän veljestä
- teokset
- Sinuhe egyptiläinen
- teokset

### Julkaisuvuodet (pb)

- 1976
- teokset
- 2016
- teokset

### Asiasanat (su)

- mustikka
- teokset
- koiraeläimet
- teokset

Teokset indeksoidaan tietokannasta indeksiin. Tiedonhakuja ei tehdä suoraan tietokantaan.

Tiedonhakija

# Miten haku toimii?

- Tiedonhakija valitsee, mistä indeksin osasta haluaa tietoa hakea eli onko kyse esim. tekijä tai nimeke.
- Tiedonhaun käyttöliittymä kohdistaa haut aina tiettyyn indeksiin
- Sanahalla saa haettua kerralla kaikista indeksistä, mutta silloin ei voi tietää, mihin MARC-kenttään/indeksiin haku kohdistuu.
- Haussa voi käyttää alavetovalikoiden lisäksi myös indeksien tunnisteita kirjoittamalla ne sanahakuun, esim. ti:seitsemän veljestä AND mtype:DVD
  - tässä haetaan nimekkeistä sanoja "seitsemän veljestä" sekä aineistotyyppiä DVD
  - aineistotyyppihaussa pitää käyttää aineistotyyppin indeksin nimeä mtype

# Mitä indeksejä on olemassa?

- [Täydellinen lista olemassa olevista indekseistä](#) löytyy Redminestä Teknisestä dokumentaatiosta.
- Indeksityöryhmä, johon kuului edustajia useammasta Koha-Suomen kimpasta sekä Varastokirjastosta ja Kansalliskirjastosta, loi kevään aikana Koha-yhteisön ja tieteellisten kirjastojen ”mäppäysten” pohjalta kattavamman mäppäyksen. Ryhmä mielti myös tarkan haun valikoihin tarvittavia muutoksia.
  - Kaikki valikkomuutokset eivät ole näkyvillä Vaskissa.
  - Mäppäyksiä testataan Vaskissa ja tieteellisissä kirjastoissa kesän ajan ja ryhmä kokoontuu syyskuun alussa ja käy läpi saadut kommentit ja kokemukset.

# Tarkka haku

- Tarkassa haussa pystyy erilaisista valikoista valitsemaan tarvittut indeksit
- Ylimmäisenä on alavetovalikoita, joissa on tarjolla yleisimmin käytetyt indeksit.

## Tarkka haku

[Siirry nidehakuun](#)

Haku [Vähemmän vaihtoehtoja](#) [Tyhjennä kentät](#)

**Hakusanat**

Nimeke		Hae indekseistä: <input type="checkbox"/>
ja	Tekijä	
ja	Asiasana	[+]

Huom! Alempia kenttiä ei voi käyttää tällä hetkellä yksinään vaan tiedot on täytettävä ylhäältä lähtien. Eli ylempiä ei saa jättää tyhjäksi.

# Tarkka haku

- Seuraavassa osiossa voi rajata aineistotyyppin, hyllypaikan, kokoelman, (hyllytarkenteen) sekä ikärajan mukaan.
  - Välilehtien välillä tehdään AND-haku
- Voit myös jo tässä vaiheessa rajata haun **emo- tai osakohteisiin**.

Aineistotyyppi Hyllypaikka Kokoelma Hyllytarkenne Ikärajat

**Rajoita seuraaviin:**

<input type="checkbox"/> Äänikasetti	<input type="checkbox"/> Äänilevy	<input type="checkbox"/> Äänilevy, puhe	<input type="checkbox"/> Äänitallenne	<input type="checkbox"/> Äänitallenne, puhe
<input type="checkbox"/> Aikakauslehti	<input type="checkbox"/> Artikkel	<input type="checkbox"/> Atlas	<input type="checkbox"/> Blu-ray	<input type="checkbox"/> Braille
<input type="checkbox"/> CD	<input type="checkbox"/> CD-ROM	<input type="checkbox"/> CD, puhe	<input type="checkbox"/> DAT-kasetti	<input type="checkbox"/> Dia
<input type="checkbox"/> DVD	<input type="checkbox"/> eKirja	<input type="checkbox"/> Elektroninen	<input type="checkbox"/> Elokuva	<input type="checkbox"/> Esine
<input type="checkbox"/> eVideo	<input type="checkbox"/> Kaavio	<input type="checkbox"/> Kalvo	<input type="checkbox"/> Kartta	<input type="checkbox"/> Kasetti, puhe
<input type="checkbox"/> Käsikirjoitus	<input type="checkbox"/> Kaukokartoituskuva	<input type="checkbox"/> Kausijulkaisu	<input type="checkbox"/> Kirja	<input type="checkbox"/> Kirjan osa
<input type="checkbox"/> Kokoelma	<input type="checkbox"/> Kollaasi	<input type="checkbox"/> Konsolipeli	<input type="checkbox"/> Levyke	<input type="checkbox"/> Maalaus
<input type="checkbox"/> Magneettinauha	<input type="checkbox"/> Mikrofilmi	<input type="checkbox"/> Moniviestin	<input type="checkbox"/> Musiikkitalenne	<input type="checkbox"/> Muu aineisto
<input type="checkbox"/> Muu painate	<input type="checkbox"/> Nauhakasetti	<input type="checkbox"/> Nuotti	<input type="checkbox"/> Optinen levykasetti	<input type="checkbox"/> Päivittyvä julkaisu
<input type="checkbox"/> Piirikotelo	<input type="checkbox"/> Piirros	<input type="checkbox"/> Raina	<input type="checkbox"/> Sana- tai kuvakortti	<input type="checkbox"/> Sanomalehti
<input type="checkbox"/> Sarjan osa	<input type="checkbox"/> Työpiirustus	<input type="checkbox"/> Valokuva	<input type="checkbox"/> Valokuvanegatiivi	<input type="checkbox"/> Video
<input type="checkbox"/> Videokasetti	<input type="checkbox"/> Videokela	<input type="checkbox"/> Videolevy	<input type="checkbox"/> Videosilmukkasetti	

**Rajaa emo/osakohteisiin:**

Vain emokohteet

Vain osakohteet



# Tarkka haku

- Seuraavassa osiossa voi rajata julkaisuvuoteen tai vuosiväliin, vastaanottovuoteen sekä kielen mukaan.
  - Kieli-valikossa ei ole kaikkia mahdollisia kieliä tarjolla. Puuttuvia pitää hakea indeksin tunnisteella In. Esim. In:ara

## Rajoitukset

Vuosi:  (muodossa: vvvv-vvvv)

Vastaanottovuosi:  (muoto: yyyy)

Kieli:

# Tarkka haku

- Seuraavassa osiossa on Lisärajaukset.
  - Keskimmäiseen Sisältö-valikkoon on lisätty vaihtoehdot **Päivittyvä julkaisu** (bib-level:i) ja **Kausijulkaisu** (bib-level:s). Muuten valikoiden sisältö on sama kuin ennenkin.
  - Rajaukset perustuvat luettelointitietojen kiinteämittäisiin kenttiin (000-008), jotka voivat olla erityisesti vanhemmalla aineistolla puutteelliset tai virheelliset.

**Lisärajaukset**

Mikä tahansa kohdeyleisö ▼ Mikä tahansa sisältö ▼ Mikä tahansa materiaali ▼

Muut sisältötyypit

**Sijainti ja saatavuus**

Vain saatavilla olevat niteet

Niteen 'Ei lainata' -tila: Mikä

- Mikä tahansa sisältö
- Kaunokirjallisuus
- Tietokirjallisuus
- Elämäkerta
- Musiikkiäänite
- Puheäänite
- Päivittyvä julkaisu
- Kausijulkaisu

# Tarkka haku

- Seuraavassa osiossa voi rajata **sijaintiin ja saatavuuteen**.
  - ”Vain saatavilla olevat niteet” –rajaus toimii niin, että mukaan tulee vain ne niteet, jotka eivät ole lainassa. Mukaan tulee myös kuljetettavana olevat tai ei lainattavat.
  - Saatavuutta ja kirjastotietoa ei pysty yhdistämään toisiinsa. Jos teoksen yksikin, minkä tahansa kirjasto nide on saatavana, tulee teos mukaan hakutulokseen.

**Sijainti ja saatavuus**

Vain saatavilla olevat niteet

Niteen 'Ei lainata' -tila: Mikä tahansa ▾

Rajoita kirjastoihin: Kaikki kirjastot ▾

TAI

Kirjastoryhmät: -- ei mikään -- ▾

# Tarkka haku

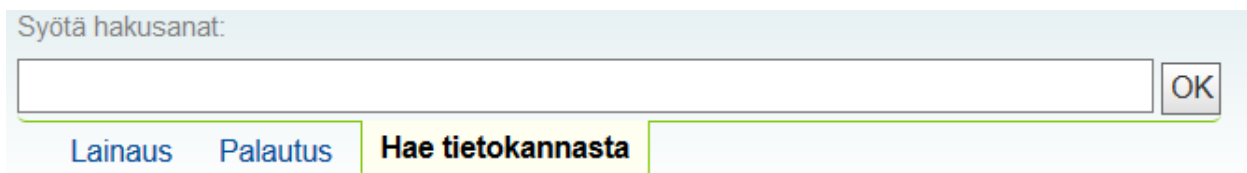
- Viimeisessä osiossa voi määrittää **hakutulosten järjestyksen**.
  - Oletuksena on julkaisuvuosi uusimmasta vanhimpaan
  - *Tällä hetkellä ei toimi järjestäminen suosion mukaan*
  - Järjestystä voi muuttaa myös hakutulossivulla
- Huomioita
  - **Julkaisuvuoden mukaan järjestäminen** tehdään 008-kentän mukaan, jos siellä on väärä/puutteellinen julkaisuvuosi, järjestyys teos väärin.
  - **Luokan mukaan järjestäminen** tehdään niteen signumissa olevan luokan mukaisesti. Se voi poiketa 084a-kentässä olevista luokista tai niiden järjestyksestä.
  - **Nimekkeen mukaan järjestettäessä** tiedot järjestetään ensin MARC-kenttien 240/130/245n mukaan, vasta sitten 245a-kentän mukaan. Erikoismerkit ensimmäisenä. Tapa ei ole suomalaisessa toimintakulttuurissa paras mahdollinen, mutta toiminee esim. englanninkielisissä Koha-kirjastoissa.

Järjestys

Järjestä

# Tunnisteiden käyttö

- Tarkan haun lisäksi on olemassa ns. Perushaku, jossa voi hakea suoraan indeksien nimillä tai niiden tunnisteilla.



Syötä hakusanat:

[Lainaus](#) [Palautus](#) [Hae tietokannasta](#)

- Tunnisteita voidaan käyttää yksinkertaisiin hakuihin, mutta se on myös tehokas tapa tehdä monimutkaisia hakuja, sillä ne mahdollistavat monipuolisten hakujen rakentamisen.
- Aloita haku kirjoittamalla ensin haluamasi indeksin tunniste ja sen jälkeen yhtäsuurikuin-merkki tai kaksoispiste:
  - ti=(seitsemän veljestä)
  - au:kivi aleksis
  - su=(suomenkielinen kirjallisuus)
  - kw=sienet

# Toimivat tunnisteet

- Nykyisessä versiossa Elasticsearchin tuki tunnisteille on puutteellinen. Vain osa Zebrassa toimivista tunnisteista toimii Elasticsearchissa.
- Vieressä listattuna tunnisteita, jotka on testattu toimivaksi.
- Jos indeksillä ei ole tunnistetta, pitää kirjoittaa koko indeksin nimi. Esim. other-classification:84.2
- Tarkassa haussa tunnisteilla voi hakea ylimmässä hakulaatikossa valitsemalla ensin Sanahaku alavetovalikosta.

Tunniste	Indeksi
ti	Nimeke
au	Tekijä
su	Aihe, asiasana
cpn	Yhteisötekijä
pn	Henkilötekijä
se	Sarjan nimeke
bc	Viivakoodi
nb	ISBN-numero
ns	ISSN-numero
lcn ja callnum	Niteen signum
ln	Kieli
pb	Julkaisija
yr	Julkaisuvuosi
loc	Hyllypaikka

# Boolean operaattorit

- Haussa voi käyttää Boolean operaattoreilla 'AND', 'OR', tai 'NOT'.
  - ti=seitsemän veljestä AND au=kivi aleksis
  - su=suomenkielinen kirjallisuus NOT pb=kekkonen
  - kw=metsä AND su=sienet
- **Boolean operaattorit pitää kirjoittaa isolla, jotta ne toimivat.**
- Boolean operaattorilla alkavat haut
  - hakulausetta ei voi aloittaa boolean operaattorilla, vaikka sen kirjoittaisi pienellä
    - ti:and she laughed
  - operaattori pitää "suojata" hipsuilla tai sulkeilla
    - ti:"and she laughed"
    - ti:(and she laughed)

# Haun rakentaminen

- ti:(seitsemän veljestä) AND au:kivi, aleksis AND mtype:kirja
  - Aleksis Kiven kirjoittamat Seitsemän veljestä teokset, joiden aineistotyyppi on kirja
- ti:(seitsemän veljestä) AND au:kivi, aleksis AND (mtype:kirja OR mtype:DVD)
  - Toimii myös ilman mtype:n toistamista
  - Aleksis Kiven kirjoittamat Seitsemän veljestä teokset, joiden aineistotyyppi on kirja tai DVD
- su:retkeily NOT mtype:alehti
  - Teokset, joissa on asiasanana retkeily ja aineistotyyppi ei ole aikakauslehti
- ti:"oma suomi" AND ti:2
  - Oma Suomi –nimisen teoksen osa 2



# Haun rakentaminen

- Ryhmittelyllä on merkitystä hakutulokseen
  - Boolean operaattoreita käytettäessä pystyy hakutermejä ryhmittelemään sulkeita käyttämällä

su:matematiikka AND su:oppikirjat NOT kw:pitkä AND  
su:lääkelaskenta

antaa eri tuloksen kuin

(su:matematiikka AND su:oppikirjat) NOT (kw:pitkä AND  
su:lääkelaskenta)

- sulkeiden vaikutusta ja logiikkaa kannattaa testaila

# Haun rakentaminen

- Elasticsearch hakee hakufraasissa olevia kaikkia sanoja riippumatta siitä, missä kohtaa fraasia sanat ilmenee.
  - ti:toiset kengät tuo tulokseen mukaan myös nimekkeen "Tiussun taussun kengät rikki, kyllä suutari toiset tekkee"
- Jos haluaa hakea termejä tismalleen siinä järjestyksessä kuin ne on, pitää ne laittaa hipsuihin.
  - ti:"toiset kengät"

# Hakutulokset

- Jos ihmettelet, miksi jokin teos tulee hakutuloksiin, tutki teoksen luettelointitietoja. Todennäköisesti hakutermin löytyy jostain kentästä.
  - Voit tarkistaa [Elasticsearchin indeksit](#) –wikistä, mihin indeksiin kyseinen kenttä on indeksoitu
- Hakutuloksia näytetään oletuksena 50/sivu. Eri kimpoissa voi olla eri oletusarvo.
- Tuloksia voi rajata faseteilla, mutta se on hidasta, koska se tekee käytännössä uuden haun.
- Teoksista näytetään muutamia keskeisiä tietoja hakutuloslistalla. Täydelliset tiedot saa näkyville avaamalla teoksen tiedot.
- Kansikuva näytetään, mikäli sellaiseen on linkki 856u-kentässä.

# Niteiden suodattaminen

Turku, Pääkirjasto kokoelmat (52)

Muut kokoelmat (1213)

▼ Suodata | ✓ Valitse kaikki | ✕ Tyhjennä kaikki

▲ Nidetyyppi ▼ Nykyinen paikka ▼ Kotikirjasto ▼

- Suodata-linkkiä voi käyttää niteiden suodattamiseen teoksen tarkat tiedot -näytöllä
  - Esim. Suomen kuvalehti 2021 sisältää paljon niteitä
    - voit suodattaa näkyviin niteet, joiden numero on esim. 2021 : 9
    - tai tietyn kirjaston niteet. Kaikkien rivien perusteella voi tehdä suodatusta. Tila-sarakkeeseen voi kirjoittaa "saatavana", jolloin saatavana-olevat suodattuvat näkyviin.

Turku, Pääkirjasto kokoelmat (52)

Muut kokoelmat (1213)

Hankintatiedot

Kuvat

▼ Poista suodattimia | ✓ Valitse kaikki | ✕ Tyhjennä kaikki

▲	Nidetyyppi ▼	Nykyinen paikka ▼	Kotikirjasto ▼	Luokka ▼	Tila ▼	Viimeisin havainto ▼	Viivakoodi ▼	Kausijulkaisun numerointi ▼
	Nidetyyppi	Nykyinen paikka	Kotikirjasto	Luokka	Tila	Viimeisin havainto	Viivakoodi	2021 : 9
<input type="checkbox"/>	Ei lainattavissa	Turku, Pääkirjasto	Turku, Pääkirjasto <i>Tieto</i>	TIE 05.1 SUO	Ei lainata (Käytettävissä lukusalissa)	24.05.2021	853N24447977	2021 : 9

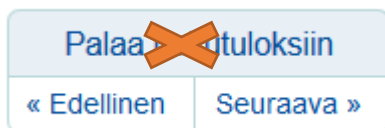
# Poimintoja hyödyllisistä indekseistä

Indeksi	Sisältö
ln	Kieli
ln-subtitle	Tekstityksen kieli
language-original	Alkuperäiskieli
ln-audio	Puheen kieli
agelevel	Ikäraja
title	Nimeke
author	Tekijä
subject	Asiasana
note	Huomautus
other-classification	YKL-luokitus
barcode	Niteen viivakoodi
mtype tai itype	Aineistotyyppi
itemtype	Aineistotyyppi ja nidetyyppi
publisher	Julkaisija

Tarkista koko lista [Elasticsearchin indeksit](#) -wikistä

# Vinkkejä

- Käytä selaimen takaisin ja eteenpäin –nappuloita ennemmin kuin Palaa hakutuloksiin –linkkiä, kun haluat palata takaisin hakutuloslistaan.



- Haussa voi hakea diakriittisia merkkejä niiden perusmuodossa eli esim. á:n voi hakea a:lla.
  - Tämä koskee myös merkkejä äöö, jolloin esim. haulla "Liian lyhyt häme" saa hakutulokseksi "Liian lyhyt hame".
- Koita rakentaa haku niin, että mukaan ei tule aikakauslehtiä, jos et nimenomaan etsi niitä. Lehdissä on paljon niteitä, joiden hakeminen hidastaa hakuja merkittävästi.
- Hakusana katkaistaan \*-merkillä

# Vinkkejä

- Jos nimeke tai tekijän nimi päättyy ?- tai !-merkkiin, täytyy merkki jättää pois, koska haku ei toimi merkin kanssa. (Sama ongelma jo Zebrassa)
- Jos hakusana sisältää erikoismerkkejä, kannattaa ne suojata hipsuilla. Muuten haku ei toimi. Korjattu uudemmassa versiossa.
  - ti:fahrenheit "1 1/9"
- Jos mahdollista, kannattaa jo hakusivulla tehdä rajaukset niin pitkälle kuin mahdollista, koska hakutuloksen rajaaminen faseteilla on hidasta.
- Kannattaa hyödyntää Finnan hakua ja Kohaan vieviä kirjanmerkkejä. Ohje kirjamerkkien lisäämiseen omaan selaimeen löytyy Kohan etusivulta uutisista.
  - Finnalla on oma hakukone ja indeksi, jotka toimivat eri tavalla kuin Kohan haku. Finna-haku on tällä hetkellä nopeampi kuin Kohan tiedonhaku.

# Tiedossa olevat ongelmat

- Jos nimeke tai tekijän nimi päättyy ?- tai !-merkkiin, täytyy merkki jättää pois, koska haku ei toimi merkin kanssa. (Sama ongelma jo Zebrassa)
- Jos hakusana sisältää erikoismerkkejä, kannattaa ne suojata hipsuilla. Muuten haku ei toimi. Korjattu uudemmassa versiossa.
  - ti:fahrenheit "11/9"
- Perustiedot-näytöllä linkkejä klikatessa kaikki haut ei toimi
  - Sarjan nimeä klikatessa, esim. Sarja: Syrjästäkatsojan tarinoita ; IX.
  - Asiasanoissa fiktiivisen hahmon nimeä klikatessa, esim. Aaltonen, Viena (fiktiivinen hahmo)
- Julkaisuvuoden ja vastaanottovuoden mukaan rajaus
  - Julkaisuvuosi hakee copydate-indeksistä, joka tällä hetkellä puutteellinen uudemman aineiston osalta. Yhteisössä käytetään date-of-publication-indeksiä, joka on parempi. Ongelma korjataan täydentämällä copydate-indeksi vastaamaan date-of-publication-indeksiä.
  - Vastaanottovuosi-rajaus suunnitellaan piilotettavaksi tarkasta hausta.
- Fasetteja muodostuu vain 10/otsikko.
- 651-indeksoitu kokonaan -> fasetteihin voi tulla finto-osoitteita
  - Korjataan mäppäykset



# Vinkkejä

Kohan hakupalkin pikanäppäimet:

- tiedonhaku Alt+Q
- lainaus Alt+U
- palautus Alt+R

Lainauksessa

- lainakuitti Alt+P



# Kiitos!

Anneli Österman

Pääkäyttäjä

[anneli.osterman@koha-suomi.fi](mailto:anneli.osterman@koha-suomi.fi)