



# Hakutulosten relevanssi EBSCO Discovery -haussa

Veera Mujunen, Regional Sales Manager, EBSCO Health, Nordic  
Countries

# Yhden hakuikkunan järjestelmät

Mitä niistä pitää tietää?



# Esimerkkejä erilaisista hakujärjestelmistä

## Federoitu haku

- Lähettää hakukyselyn eri aineistoihin ohjelmointirajapintojen kautta
- tuo tulokset järjestelmään
- mahdollisesti sulauttaa vastaukset yhteen hakutulostaan

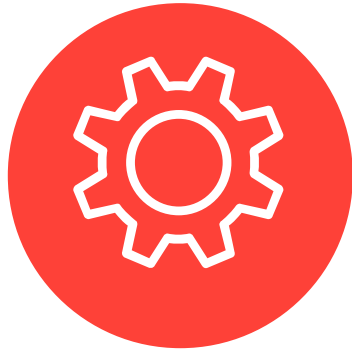
## Haku yhdessä indeksissä

- Paikallinen haku yhdessä indeksissä
- Metadatan runsaus vaihtelee (esim. ”ohut” metadata kuten sisällysluettelotiedot)
- Erilaisia painotuksia hakualgoritmeissa

## EBSCO Discovery –haku

- Paikallinen haku yhdessä indeksissä
- Lisensoitu runsas metadata asiasanoineen (viitetietokannat!)
- Asiasanoja hyödyntävä hakualgoritmi

# Hakujen ongelmia tieteellisten artikkelien haussa



## Epävarma toimintaympäristö

Riippuvuus rajapinnoista ja eroavaisuudet eri järjestelmien välillä



## Epätäydellinen tieto

Vrt. tietokannan metadata asiasanoineen vs. samojen artikkelien "ohuemmat" tiedot esim. pelkkä sisällysluettelo.



## Tuloksena tulosten vaihteleva laatu

Tulosten kattavuus ja järjestäminen luotettavasti on vaikeaa, jos sisältö on puutteellista (metadata) ja tekninen ympäristö sirpaleinen (rajapinnat, erilaiset tietueet).

# EBSCO Discovery –haku



Runsas metadata on lisensoitu yhteistyökumppaneilta

Tarvittavat tiedot



Metadata on "säilötty yhteen siiloon"

Yhtenevä alusta haulle



Indeksi koostuu osista

Muokattavuus



Runsasta metadataa hyödyntävä hakualgoritmi

Luotettava tulosten järjestäminen



Relevantit hakutulokset

Käyttäjä löytää etsimänsä

# Miten hakutulokset järjestetään EBSCO Discovery -haussa?

EBSCO Discovery –haun relevanssi



# Tulosten relevanssia määrittävät tekijät EBSCO Discovery -haussa:

Järjestelmä noudattaa seuraavia prioriteettia **eikä suosi mitään sisällöntuottajaa:**

1. **Täsmäävät asiasanat** kontrolloiduista asiasanastoista
2. Täsmäävät artikkelien otsikot
3. Täsmäävät kirjoittajien antamat hakusanat
4. Täsmäävät hakusanat asbtraktissa
5. Täsmäävät hakusanat kokotekstissä

# Lisäksi hienosäätöä varten käytetään toissijaisia kriteerejä

(käyttäjätestaukseen ja palautteeseen  
perustuvia)





# Toissijaiset kriteerit

- Ajantasaisuus: Äskettäin julkaistu sisältö saa hieman enemmän painoarvoa kuin vanhempi sisältö
- Dokumenttityyppi: kokooma-artikkeli on halutumpi kuin kirja-arvostelu
- Dokumentin pituus: 4 sivun teksti on halutumpi kuin ¼ sivun teksti

# Metadatan vertailu

Edellä mainittu on mahdollista vain, jos asiasanat ja muu runsas metadata on saatavilla.



# Verrataanko indeksin laajuutta vai syvyyttä?

**Esim. Mitkä tiedot tietokannasta X ovat mukana haussa?  
/ Mikä metadata tietokannasta X on mukana haussa?**

VS.

Onko tietokanta X mukana haussa?

# EBSCO Health

## Kysymyksiä?

Veera Mujunen



[vmujunen@ebSCO.com](mailto:vmujunen@ebSCO.com)



+358 40 754 7503