

Annif OSC:ssä

Miten me käytetään?



Annif on kaksi asiaa

- ohjelmisto (Annif)
- palvelu (finto-ai)

Me käytämme palvelua koska:

- **helppo käyttää**
- **tulokset ovat hyviä**

"Sisäänleivotut" käyttökohteet

- [opinnäytteiden JYX-julkaisu](#)
- [sisäinen JYX-tallennuslomake](#)
- opinnäytehaku (tulossa)
- opiskelijoiden tutkimusaineistolomake (tulossa)

Muu käyttö

- Converis -tallentajat (abstrakti Annifiin)
- jotkin luetteloijat käyttävät Annifia itsenäisesti
- myös [sisäinen Annif-lomake](#) tarjolla

Kokemuksia I

*Runsaasti kokemusta, koska julkaisen opinnäytteitä. Asiasanoituksen osuvuus vaihtelee hyvin paljon, mikä ehkä eniten johtuu YSOsta: joillekin opinnäytteille löytyy paljon hyviä asiasanoja, toisille taas ei mitään. Puutteita on erityisesti luonnontieteissä ja matematiikassa. **YSO on jatkuvasti jäljessä käsitteiden kehityksessä** ja "uudet" tutkimusaiheet ja käsitteet puuttuvat. Joskus taas kummasti Annif ei ehdota keskeistä esim. opinnäytteen otsikossa olevaa sanaa, joka YSOsta kuitenkin löytyy.*

Kokemuksia II

Annif nopeuttaa työtä verrattuna siihen, että alkaisi silmäilemään julkaisun tekstiä läpi. Tulokset vaihtelevat kuitenkin hyvin paljon, uudet tutkimusaiheet ja monitieteisyys ovat Annifille tosiaan hankalampia, joten ihmisälyä tarvitaan asiasanojen validoijaksi

Kokemuksia III

Joskus saa päivän piristykseen, kun ehdotetuissa asiasanoissa on jotain hassua, esim. Hippos-hanketta käsitelleelle opinnäytteelle asiasanaksi ehdotettiin "virtahevot".

Käytännön vinkkejä

Jos Annif antaa täysin puppua

- | kieli on väärin

Jos seassa täysin irrallisia asiasanoja

- | yksi mahdollisuus on "tässä sarjassa julkaistua"

Monesti abstrakti riittää

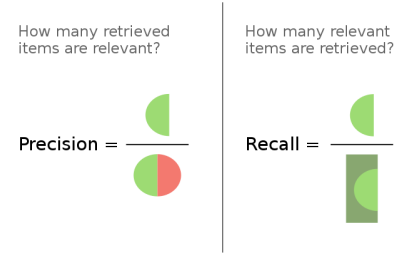
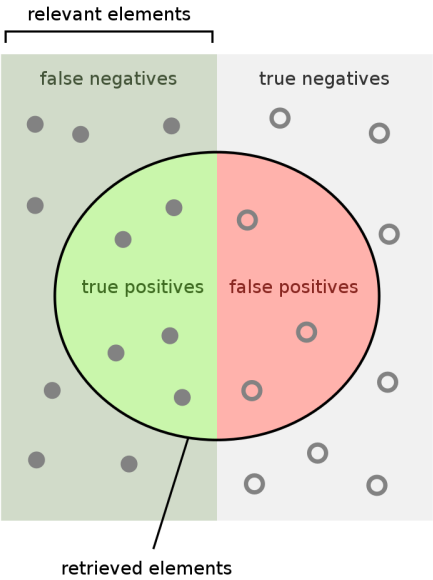
Miten arvioida tuloksen laatua?

tarkkuus (precision)

- mikä on oikeiden ja väärin arvojen suhde tuloksessa?

herkkyys (recall)

- Mikä on palautettujen oikeiden arvojen suhde kaikkiin oikeisiin arvoihin?
- asiasanoituksessa ei voi olla 1!



herkkyys

- Mikä ei kuulu joukkoon?
 - ihmiselle ehkä helpompi hahmottaa sanalistaan kuulumattomat sanat

tarkkuus

- Mitä joukosta puuttuu?
 - vaikeampi sanoa, mitä termejä sanalistasta puuttuu

Opinnäytelomakkeen data

Opinnäytelomakkeesta on saatu hyvää dataa Annifin kehityksestä:

<https://liberquarterly.eu/article/view/10732/11613>

chatGPT?

<https://jyx.jyu.fi/handle/123456789/86086#>

"Anna vain lista asiasanoista"

Maahanmuuttajat
Perusopetus
Opettajat
Teknologia
Verkkokysely
Laadulliset menetelmät
Työtapojen monipuolistaminen
Opiskelijoiden aktiivisuus ja into
Opetuksen vaikutukset
Haasteet
Koulutus
Tutkimus
Suuntaukset
Aukko tutkimuskentässä

"Käytä vain YSO-asiasanoja"

Valitettavasti joitain tarkkoja YSO-asiasanoja ei löytynyt, mutta tässä on listaus löydettyistä YSO-asiasanoista:

- maahanmuuttajakoulutus
- perusopetus
- opettajat
- opetusteknologia
- kyselyt
- laadullinen tutkimus
- työtapojen monipuolistaminen
- opiskelijoiden motivaatio
- oppimistulokset
- koulutusmahdollisuudet
- tutkimus
- tutkimusaiheet

"listaa sanojen perään vielä YSO -tunniste"

Valitettavasti osalle sanoista ei löydy YSO-tunnisteita, mutta tässä on lista tunnisteista, jotka löysin:

```
maahanmuuttajakoulutus (http://www.yso.fi/onto/ysa/Y49789)
perusopetus (http://www.yso.fi/onto/ysa/Y8806)
opettajat (http://www.yso.fi/onto/ysa/Y8046)
opetusteknologia (http://www.yso.fi/onto/ysa/Y52287)
kyselyt (http://www.yso.fi/onto/ysa/Y8954)
laadullinen tutkimus (http://www.yso.fi/onto/ysa/Y23591)
työtapojen monipuolistaminen (tämänkaltaista tunnistetta ei löytynyt)
opiskelijoiden motivaatio (http://www.yso.fi/onto/ysa/Y45658)
oppimistulokset (http://www.yso.fi/onto/ysa/Y75234)
koulutusmahdollisuudet (http://www.yso.fi/onto/ysa/Y86463)
tutkimus (http://www.yso.fi/onto/ysa/Y23557)
tutkimusaiheet (http://www.yso.fi/onto/ysa/Y28906)
```

"Keksit YSO-tunnisteet omasta päästäsi."

Valitettavasti en voi keksiä YSO-tunnisteita omasta päästäni, sillä YSO on standardoitu asiasanasto, jota ylläpitää Kansalliskirjasto. YSO-tunnisteet ovat siis standardoituja termejä, joita käytetään yhtenäisen kuvailun ja tiedonhaun varmistamiseksi. Niiden tarkoitus on helpottaa tietojärjestelmien välistä yhteensopivuutta ja asiasanastoissa käytettyjen termien yhdenmukaisuutta.

"Opettajat ei ole Y8046"

Pahoitteluni, erehdyin aiemmin. Oikea YSO-tunniste sanalle "opettajat" on Y2957.

"Oikea YSO-tunniste opettajat -sanalle on p1117"

Pahoittelen virhettäni. Oikea YSO-tunniste sanalle "opettajat" on P1117. Kiitos korjauksesta!

"Saanko listan korjattuna?"

Tässä on korjattu lista YSO-tunnisteista:

```
maahanmuuttajat: P1384  
perusopetus: P7268  
opettajankoulutus: P9791  
tekniologia: P1167  
opetusmateriaalit: P8576  
monipuolinen oppiminen: P1176  
itsenäinen työskentely: P4382  
aktiivinen oppiminen: P3576  
keskittyminen: P3197  
opetuksen vaikutukset: P9986
```

"No mihin se opettajat jäi?"

Pahoitteluni, tein virheen aiemmissa vastauksissani. Oikea YSO-tunniste sanalle "opettajat" on P2957. Joten täydennän listaani:

```
maahanmuuttajat: P1384
perusopetus: P7268
opettajankoulutus: P9791
tekнологia: P1167
opetusmateriaalit: P8576
monipuolinen oppiminen: P1176
itsenäinen työskentely: P4382
aktiivinen oppiminen: P3576
keskittyminen: P3197
opetuksen vaikutukset: P9986
opettajat: P2957
```

*) P2957 = geenikartoitus

Geneeriset AI:t eivät korvaa Annifia

- YSO- tietoisuus ja päivittyvyys
- Annif on avointa lähdekoodia
- Annif on ilmainen käyttää

Kokonaan toinen asia on asiasanoituksen tulevaisuus ylipäätään.