

# YSON linkitys Wikidataan

## Taustaa

[In English](#)

YSON käsitteitä on linkitetty Wikidataan vuodesta 2019 lähtien. Linkin ilmaisemiseen Wikidatassa käytetään Wikidata-ominaisuutta "YSO-tunniste" (Property:P2347). Linkityksestä ovat vastanneet YSON ylläpitäjät. Vuosien 2019-2020 aikana käytiin läpi ensimmäinen otos eli reilu 7300 kpl YSON käsitettä ja luotiin linkit sopivissa tapauksissa. Ensimmäisen otoksen linkitys tehtiin käymällä läpi lista automaattisesti luotuja linkkiehdotuksia, jotka joko hyväksyttiin tai hylättiin. Linkkiehdotukset pohjautuivat seuraaviin kolmeen lähteeseen:

1. LCSH-käsitteistä löytyvät Wikidata-linkit niissä tapauksissa, joissa YSON käsite on linkitetty LCSH:n käsitteeseen.
2. Ylen sisällönkuvailussa käytetyt käsitteet, joilla on vastine sekä YSOssa että Wikidatassa<sup>1</sup>.
3. Lista Finnassa esiintyvistä sisällönkuvailussa eniten käytetyistä YSO-käsitteistä sekä näiden vastineista Wikidatassa (vastineiden perustana termien vastaavuus).

## Tilannekatsaus 2022

Vuoden 2022 tammikuun lopussa YSO-tunniste ([P2347](#)) on lisätty **14785** Wikidata-käsitteeseen. YSO-paikkojen tunnisteita on lisäksi **6591** Wikidata-käsitteessä. Vuoden 2021 lopulla laadittiin listaus YSO-käsitteistä, joiden tunnistetta ei ole Wikidatassa. Tällaisia oli noin 20.000. Näiden YSO-käsitteiden käytettävät termit ladattiin monikko- ja yksikkömuodossa Openrefine 3.5 -ohjelmaan, jossa niitä mapattiin Reconcile-toiminnolla Wikidatan samankielisten termien kanssa. Osumia saatiin noin 3500, joiden tunnisteet ja labelit ladattiin Wikidataan.

Helmikuussa 2022 jäljellä on vielä noin **12500 YSO-käsitettä**, joiden käytettävät termit eivät mappautuneet suoraan Wikidataan, joten niiden vastaavuus pitää yksitellen tarkistaa. Kaikille ei tule löytymään vastinetta. Tällaisissa tapauksissa voi kuka tahansa Wikidatan toimittaja luoda vastaavan käsitteen Wikidataan ja lisätä tarvittavat ominaisuudet sekä YSO-ID tunnisteeseen.

Vuonna 2022 tulemme poimimaan Wikidatasta niiden käsitteiden URI:t, joissa on YSO-tunniste, ja lisäämme ne YSON käsitetietueisiin. Helmikuusta 2022 alkaen Wikidata-tunnisteita lisätään YSON uusiin käsitteisiin vastaavalla tavalla kuin LCSH-sanaston tunnisteet. Linkkiä ei välttämättä lisätä Wikidataan. Myöhemmin valmistettavien Wikidatan botti-ohjelmien avulla Wikidata-tunnisteita voidaan poimia Finton YSO-käsitteistä ja täydentää YSO-tunniste automaattisesti vastaavaan Wikidata-käsitteeseen.

## Linkityisperiaatteita

YSON Wikidata-linkityksessä sovelletaan seuraavia periaatteita:

- Linkityksessä käytetään skos:closeMatch-suhdetta.
- Linkitys tehdään YSO-käsitettä tarkimmin vastaavaan Wikidata-kohteeseen.
- Linkitys tehdään vain, jos sopiva vastine Wikidatasta löytyy.
- Yksi YSO-käsite voidaan linkittää useaan Wikidata-kohteeseen silloin, kun YSON käsitteen merkitys kattaa nämä kohteet.
- Usea YSO-käsite voidaan linkittää samaan Wikidata-kohteeseen silloin, kun Wikidata-kohde kattaa YSO-käsitteiden merkityksen.
- YSO-käsitteistä ei tehdä linkityksiä Wikipedian disambiguationisivua vastaaviin Wikidata-kohteisiin.
- Linkityksessä hyödynnetään myös [YSON LCSH-linkityksen periaatteita](#) soveltuvien osin.
- Wikidata-kohteiden ja Wikipedia-artikkelien sisältöä voidaan muokata kohtuullisessa määrin silloin, kun se on hyödyllistä linkityksen kannalta. Tässä noudatetaan yleisiä Wikidatan periaatteita. Muokkaaminen koskee ilmeisten virheiden ja epätarkkuuksien korjaamista ja sen tarkoituksena on parantaa kohteiden laatua ja semanttista yhdenmukaisuutta. Muokkauksia tehdään vain suomen-, ruotsin- tai englanninkielisen sisällön osalta. Muokkaustoimenpiteitä ovat:
  - Wikidata-kohteen termin, määritelmän tai ohjaustermin muokkaus, lisääminen tai poisto.
  - Wikipedia-artikkelilinkin lisääminen, poistaminen tai vaihtaminen Wikidata-kohteessa.
  - Kahden Wikidata-kohteen yhdistäminen käyttämällä Wikidatan merge-toimintoa.
  - Wikidata-kohteeseen linkitetyn Wikipedia-artikkelin otsikon muuttaminen.

<sup>1</sup> = Yleltä saatu koostetiedosto, joka sisältää kaikki Ylen maaliskuuhun 2018 mennessä sisällönkuvailussa esiintyneet käsitteet. Nämä sisältävät linkejä eri resurssisiin kuten KOKO-ontologiapilveen, Wikidataan ja Freebaseen. YSON Wikidata-linkityksessä on hyödynnetty näistä sellaisia käsitteitä, joilla on linkit sekä KOKOon että Wikidataan silloin kun KOKOn käsitteellä oli vastine myös YSOssa. Tällaisia oli yhteensä noin 2500 kpl.