

Miten kuvia ja dataa käytetään Wikimediassa?

HRI Loves Developers: Kartalla historiasta
6.10.2026

Kimmo Virtanen, Zache



Tausta

Wikimedia Suomella oli vuosina 2018–2023 jälleentalokuvausprojekti.

- Helsingin kaupunginmuseon kuvista valittiin käsin noin 10 000 kuvan otos.
- Suurimmassa osassa kuvista oli valmiiksi osoitetason sijaintitieto.
- Kuvat tuotiin Ajapaikiin (virolainen jälleentalokuvaussivusto).
- Ajapaikissa ihmiset korjasivat kuvien metatietoja ja kävivät jälleentalokuvaamassa kuvien kohteita.



Tuntematon, valokuvaaja

Marian sairaala, pihamaa kesällä.

Maria's hospital, yard in summer.

1920



[Helsingin kaupunginmuseo hkm.HKMS000005:km0000m9d0](#)



[Helsinki \(Finland\)](#)

[Finland](#)

[Add to album](#) +



[Mauri Tuovinen](#) , 28.05.2022



[Admin](#)

- Add a geotag
- Date
- Add a rephoto
- Annotate
- Share
- Categorize
- Transcribe
- Edit picture
-



Add a
geotag



Date



Add a
rephoto



Annotate



Share



Categorize



Transcribe



Edit
picture



Tuntematon, valokuvaaja

Marian sairaala, pihamaa kesällä.

Maria's hospital, yard in summer.

1920



[Helsingin kaupunginmuseo hkm.HKMS000005:km0000m9d0](#)



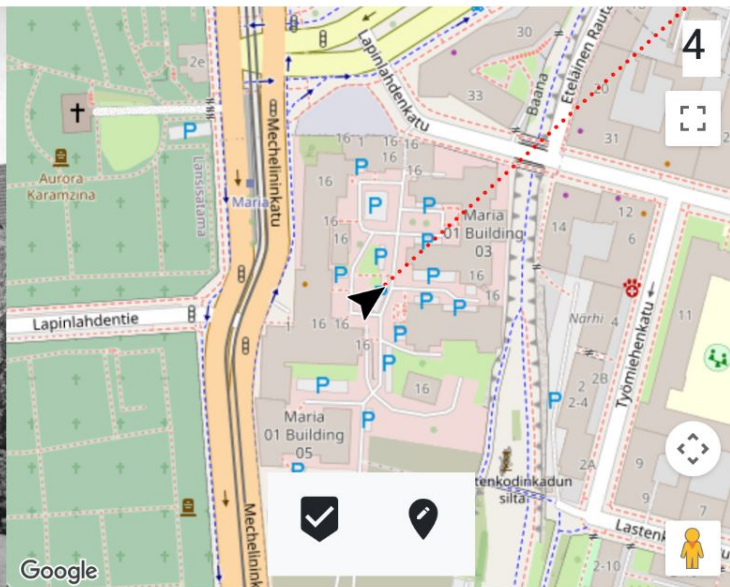
[Helsinki \(Finland\)](#)

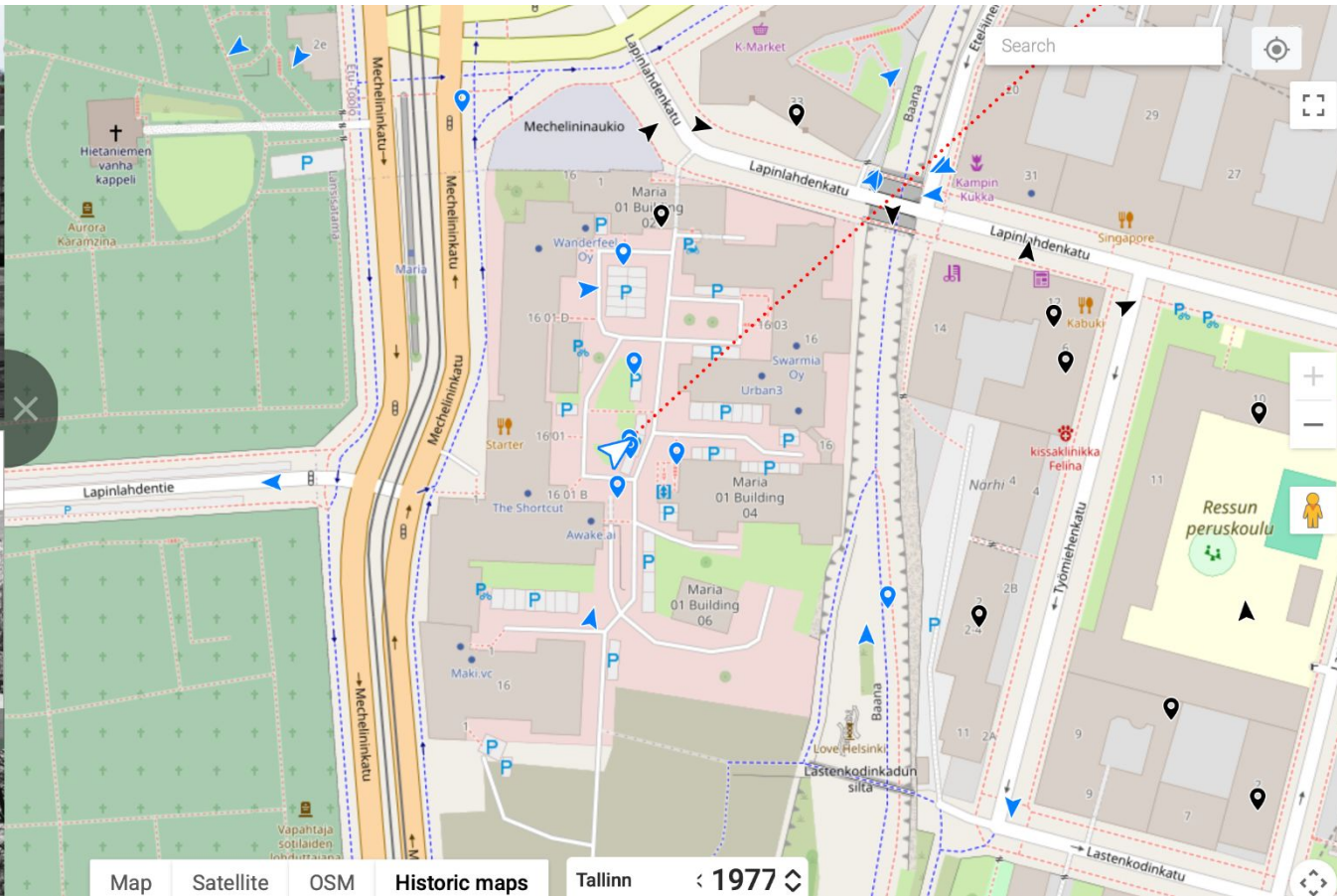
[Finland](#)

[Add to album](#) +



[Tag a person](#) +







Ajapaikin joukkoistusfilosofia

- Keskitytään pelkästään koneellisesti luettavaan tietoon.
- Kuvat luokitellaan ensin automaattisesti, ja ihminen korjaa luokitusvirheitä niitä huomatessaan.
 - Luokkia ovat esimerkiksi ilmakehä, taivasta, maanrajasta, sisäkuva ja ulkokuva. Lisäksi luokitellaan kuvassa näkyviä ihmisiä: lapsia, aikuisia, vanhuksia, naisia ja miehiä....
 - Luokat on tarkoitettu filttäilyyn, ja ne ovat tarkoituksella karkeita, jotta virheiden määrä pysyy pienenä.
- Korjattua datasettiä voidaan käyttää uuden luokittelijan kouluttamiseen.
- Ihmisten tekemissä arvioissa tulos on vastausten ”keskiarvo” painotettuna sen mukaan, miten hyvin vastaaja on vastannut aiemmin.

Mitä joukkoistus lisää lähtodataan nähden

Sijainti

- Alkuperäinen sijaintitieto ja mahdollinen tarkennettu kuvauspaikka

Kuvaussuunta

- Mihin suuntaan kamera on osoittanut

Ajoitus

- Alkuperäinen ajoitus ja mahdollinen tarkennus

Mitä paikalla on nykyään

- Mahdollinen jälleenalokuva toimii muutoksen hahmottamisen lisäksi tietojen varmistuksessa

Tunnisteet ja linkit

- Miten sama kuva löytyy museon aineistosta, Ajapaikista ja Commonsista. (Perceptual hashes, dhash+phash)

Tilastoja Ajapaikista

Ajapaikin Helsingin kaupunginmuseo -albumin tilanne 6.10.2026:

- Albumissa oli noin 13 000 kuvaa, joista suurimmalla osalla oli vähintään karkea sijainti.
- 283 käyttäjää oli tarkentanut 5 000 kuvan sijaintia tallennuksen jälkeen.
- 3 600 kuvalle oli asetettu tarkka sijainti ja kuvan ottosuunta.
- 7 000 kuvalle oli asetettu tarkempi päiväys tai vuosi.
- Tallennettuna oli 2 420 jälleentalokuvaa 131 kuvaajalta.

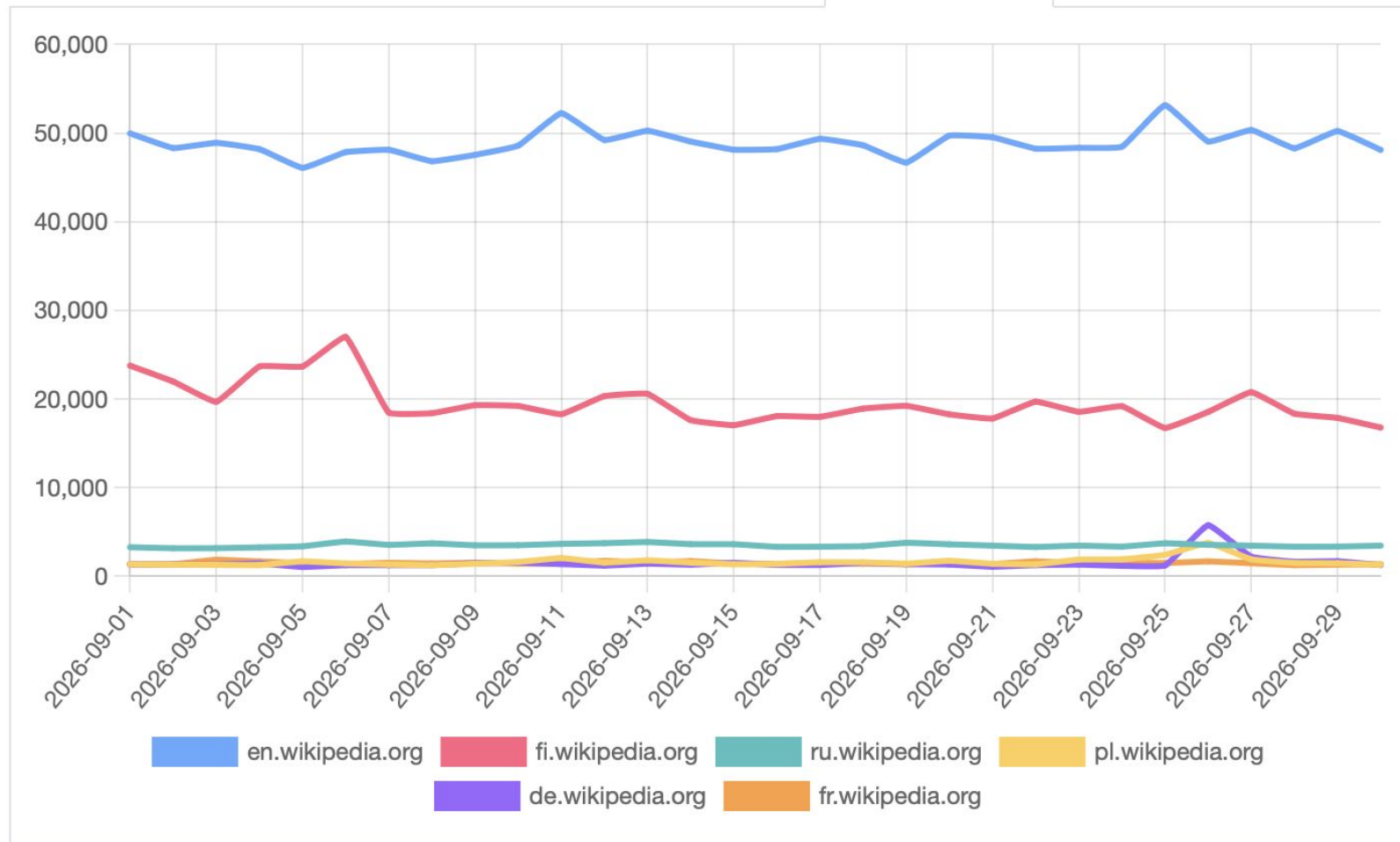
Kuvat Wikimedia Commonsissa

Vuonna 2023 kuvat tallennettiin systemaattisesti myös Wikimedia Commonsiin.

Luokka [Files from the Helsinki City Museum](#) tilanne 6.10.2026:

- Luokassa oli 21 400 kuvaa. Eri Wikimedia-projekteissa oli näistä käytetty 2 188 kuvaa (10,2 %).
- Kuvia oli katsottu syyskuussa yhteensä 2 460 312 kertaa.
- Ajapaikiin tallennettujen kuvien lisäksi luokassa on alkuperäisestä kuvasta rajattuja kuvia ja käyttäjien käsin tallentamia kuvia.

Lähde: [Glamorous-työkalu](#)

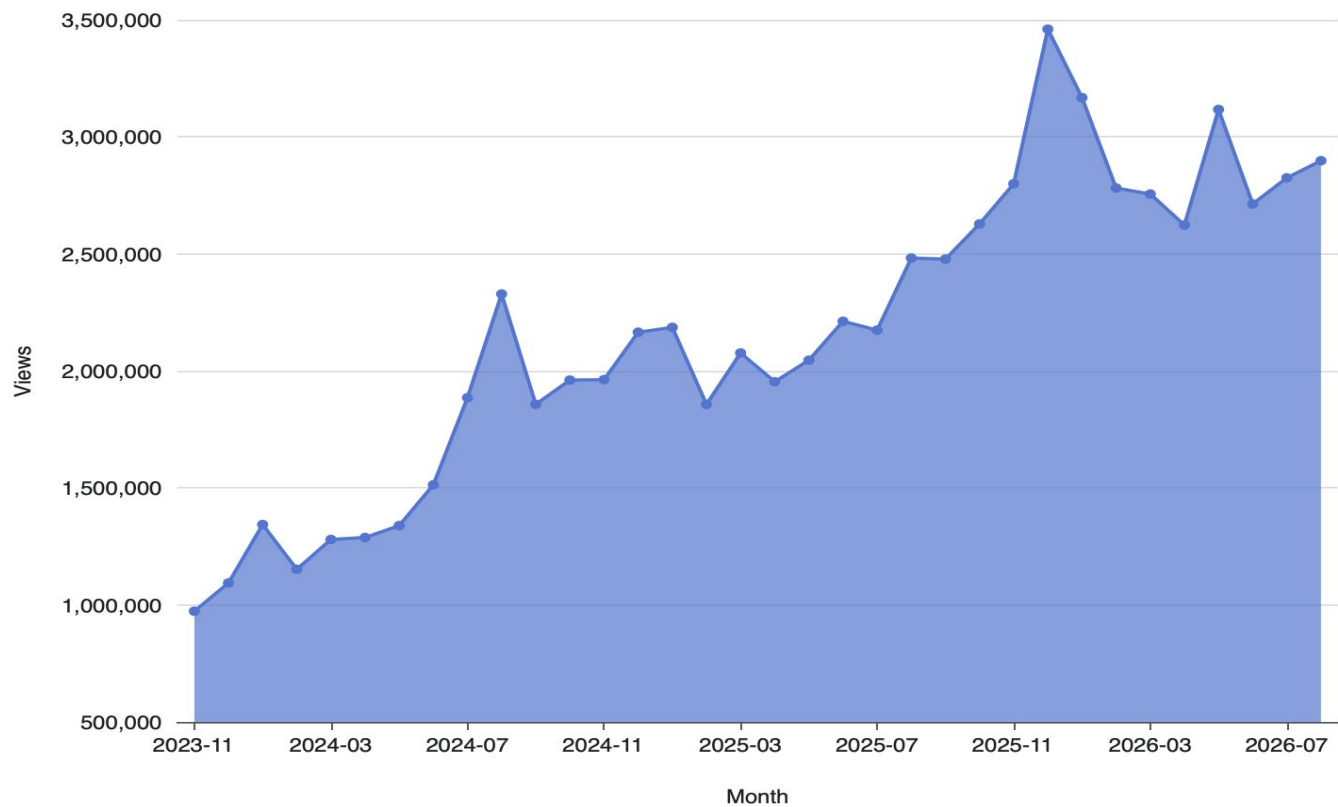
[Process](#)[Global file usage](#)[File usage details](#)[Daily views](#)

[Process](#)[Global file usage](#)[File usage details](#)[Daily views](#)

Wiki ↓	Distinct files ↓	Pages using files ↓	Total file usages ↓	Page views 2026-09 ↓
fi.wikipedia.org	1,410	1,081	1,723	585,692
www.wikidata.org	456	480	531	9,740
en.wikipedia.org	401	353	521	1,465,450
sv.wikipedia.org	301	247	384	22,972
fr.wikipedia.org	207	165	233	44,837
cs.wikipedia.org	117	42	126	1,671
de.wikipedia.org	97	88	113	45,522
ru.wikipedia.org	70	73	78	104,790
fa.wikipedia.org	55	32	55	8,069
es.wikipedia.org	50	45	59	13,987
it.wikipedia.org	38	30	42	7,859
zh.wikipedia.org	35	28	37	9,057
uk.wikipedia.org	33	32	35	2,657

● Pageviews

Files uploaded by FinnaUploadBot pageviews 2023-2026





Kolme romaninaista Helsingissä
1930-luvulla. Kuvaaja tuntematon. PD.
Kuvaa on käytetty Suomen romanit
-artikkelin eri kieliversioissa

https://commons.wikimedia.org/wiki/File%3AKolme_romaninaista.jpg



Kevyen liikenteen väylä Baana. Kuvaaja Sakari Kiuru. CC BY 4.0.
Kuvaa on käytetty englanninkielisen wikipedian artikkelissa
"Basketball".

https://commons.wikimedia.org/wiki/File%3AKevyen_liikenteen_v%C3%A4yl%C3%A4_Baana_-_G8537_-_hkm.HKMS000005-km0000n5j4.jpg



Diskotanssiesitys Kontulan 30-vuotispäivillä vuonna 1994. Kuvaaja tuntematon. CC BY 4.0. Kuvaa on käytetty artikkelin "Disco" eri kieliversioissa.

[https://commons.wikimedia.org/wiki/File%3ADiskotanssiesitys_Kontulan_30-vuotisp%C3%A4ivill%C3%A4_1994_-_N262306_\(hkm.HKMS000005-km0036in\).jpg](https://commons.wikimedia.org/wiki/File%3ADiskotanssiesitys_Kontulan_30-vuotisp%C3%A4ivill%C3%A4_1994_-_N262306_(hkm.HKMS000005-km0036in).jpg)



Paavo Nurmi sytyttää olympiatulen avajaisissa vuonna 1952. Kuvaaja tuntematon, Olympia-kuva Oy. CC BY 4.0. Kuvaa on käytetty artikkeleissa "Paavo Nurmi" ja "Kesäolympialaiset 1952".

https://commons.wikimedia.org/wiki/File:Olympiatuli_1952.jpg



Suomen ETYK-delegaatio vuonna 1975. Kuvaaja Tapio Korpisaari. CC BY 4.0. Kuvaa on käytetty artikkelissa "ETYK". [Derek Longmuir](https://commons.wikimedia.org/wiki/File%3ABay_City_Rollers_-_yhtyeen_rumpali_Derek_Longmuir_on_saapunut_hotelli_Hesperian_edustalle_-_G35573_(hkm.HKMS000005-km000nhk4).jpg)

[https://commons.wikimedia.org/wiki/File%3ABay_City_Rollers_-_yhtyeen_rumpali Derek Longmuir on saapunut hotelli Hesperian edustalle - G35573 \(hkm.HKMS000005-km000nhk4\).jpg](https://commons.wikimedia.org/wiki/File%3ABay_City_Rollers_-_yhtyeen_rumpali_Derek_Longmuir_on_saapunut_hotelli_Hesperian_edustalle_-_G35573_(hkm.HKMS000005-km000nhk4).jpg)

Käytetty artikkelissa ETYK

<https://commons.wikimedia.org/wiki/File%3AETYK-Finland-delegation-1975.jpg>

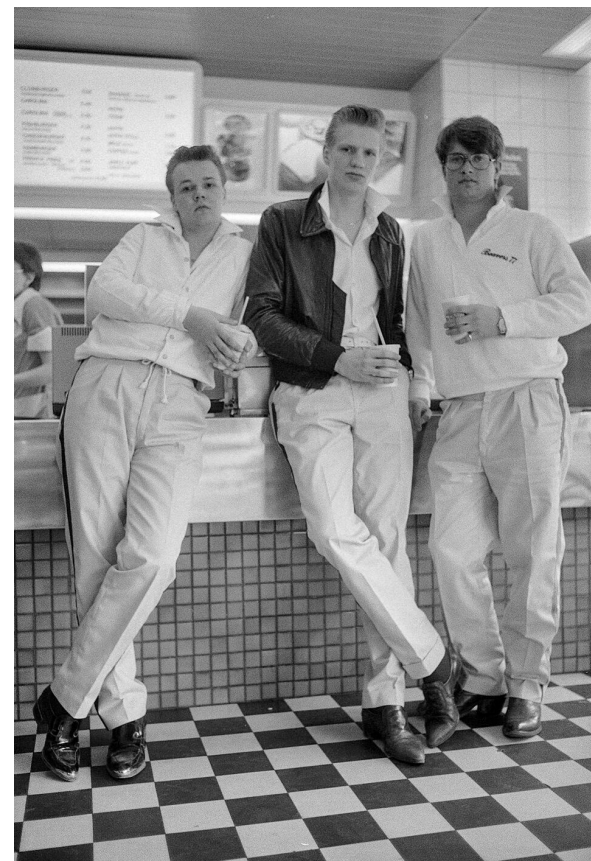




Hedvig Gebhard, oik. istumassa Miina Sillanpää. Kuvaaja tuntematon, PD
<https://commons.wikimedia.org/wiki/File:%3AGebhard-Sillanpaa-1910.jpg>



Suomi–Ruotsi-maaottelu Stadionilla. Lasse Virén johtaa. CC-BY-4.0, Reijo Pasanen
<https://commons.wikimedia.org/wiki/File:%3ALasse-Viren-1980.jpg>



Rockabilly-yhtye Teddy & The Tigers Citykäytävän Carrolsissa. CC-BY-4.0, Harri Ahola
<https://commons.wikimedia.org/wiki/File:Teddy-et-the-Tigers-1978.jpg>

Mitä HRI:n aineistoja ollaan käytetty

- Kaupunginmuseon ja muiden Helsingin museoiden kuvia ja historiallisia karttoja.
- Luettelot suojellusta rakennetusta ympäristöstä ja rakennustunnisteet.
- Vektorirajat suojelluista ja arvokkaista ympäristöistä.
- Digitransit API rajapinnat koordinaatti-osoite muutoksiin.
- Finna ja Finto, koska lisenssi on CC0

Mikä on tavoite

- Mahdollisimman pitkälle automatisoida metadatan rikastus ja jättää vastaantulevien virheiden korjaus ihmiselle
- Esimerkiksi siten, että käyttäjä voi kuvan käsittelyn yhteydessä suoraan kartalta klikata mitä kuva käsittelee.
- Tehdä aineistosta löydettävämpää. Pystytään hakutyökaluilla yhdistämään aiheet joista puuttuu kuvat ja kuvat joita on olemassa
- Wikimedia on monikielinen joten aineistoja pitäisi pystyä hakemaan usealla kielellä.
- Ratkaisu tähän on muuntaa metadata linkitetyksi dataksi

