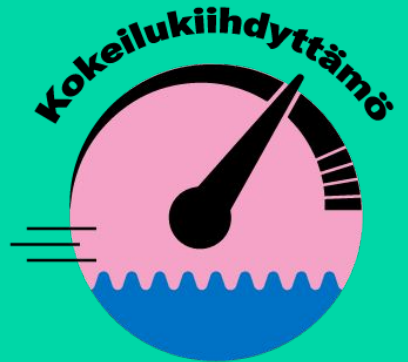


Tekoälykokeilun loppuraportti

Suojateiden tunnistaminen ilmakuvista
koneoppimisen avulla

Helsinki



Tukea digitalisaatiokokeiluihin
kaupungin työntekijöille

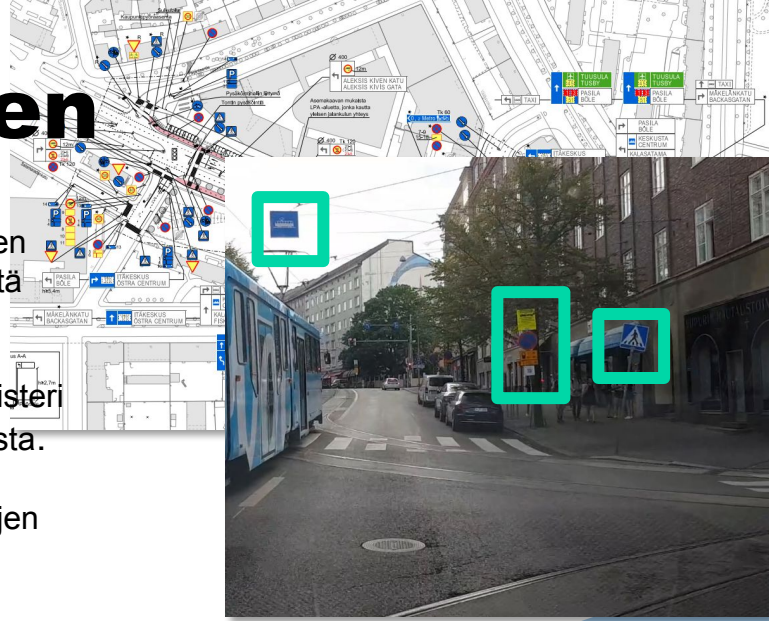
Suojateiden tunnistaminen ilmakuvista koneoppimisen avulla

Kaupunkiympäristön toimiala: Fanny Taxell, Saska Lohi,
Virpi Vertainen, Minna Leinonen, Esko Laiho

Integrify: Katja Lairikko, Alex Jung

1. Kokeilun onnistuminen

- Kaupungilla ei tällä hetkellä ole kattavaa rekisteriä liikenteenohjauslaitteiden suunnitelma- ja toteumatiedoista. Yksi tämän kokeilun tavoitteena on kerätä dataa kehitettyä liikenteenohjauslaitteiden tietokantaa varten. Kaupunkiympäristön toimialalla on käynnistynyt kehityshanke avoimesta liikenteenohjauslaitteiden tietokantaratkaisusta, joka on toteutuessaan rekisteri liikenteenohjauslaitteiden (ml. tiemerkinnyt) suunnitelma- ja toteumatiedoista.
- Kaupunkiympäristön toimialalla on tehty vuosien mittaan tiemerkinntöjen inventointeja, mutta aineistot ovat vanhentuneita ja suunnitelmaohjelmistojen tuottamat aineistot ovat olleet vaikeasti hyödynnettävissä ja hallittavissa.



Tämän kokeilun tavoitteena on tunnistaa koneoppimisen avulla tiemerkinntöjä kaupungin vuosittain kuvatuista ortoilmakuvista.


- Selvitetään mahdollisuutta luoda työkalu, joka tunnistaa, käsittelee ja luokittelee tiemerkinntöjä ilmakuvista. Kokeilun tiemerkinntädataksi valikoitui suojateiden tunnistaminen.
- Kuinka helppoa on kaupungin omasta avoimesta datasta louhia ja luokitella haluttuja aineistoa koneoppimisen avulla?

1. Kokeilun onnistuminen

Kokeilussa onnistuneesti todennettiin mahdollisuus automaattiseen suojaiteiden tunnistamiseen ilmakuvista koneoppimisen avulla.

- Kokeilussa luotiin koneoppimisen malli, joka tunnistaa ilmakuvista suojaiteita.
- Kokeilun aikana opittiin lisää tekoälyn ja koneoppimisen hyödyntämisestä ja käytännössä kokeiltiin kaupungin oman datan tehokkaampaa hyödyntämistä.
- Saatu suojaiteaineisto viedään kaikkien kaupungin työntekijöiden ja kaupunkilaisten nähtäville kartta.hel.fi – sivustolle.
- Aineisto voidaan päivittää joka vuosi kun kaupunki kuvaa uuden ortoilmakuvan.
- Aineisto toimii testiaineistona digitalisaatio-ohjelman kärkihankkeen ”Liikenteenohjauslaitteiden alustan” kehittämistyössä.
- Mallia voidaan opettaa vielä paremmin tunnistamaan oikeat kohteet.

Pidemmillä tähtäimellä ajantasainen tieto liikenteenohjauslaitteista mahdollistaa älyliikenteen toteutumisen sekä suunnittelua tukevat analyysit ja parantaa omaisuudenhallintaa.



”Omaisuuden tila tulee tuntea, jotta voidaan ennakoida toimenpiteitä”

2. Kokeilun eteneminen

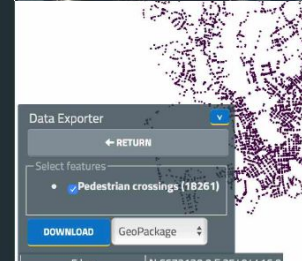
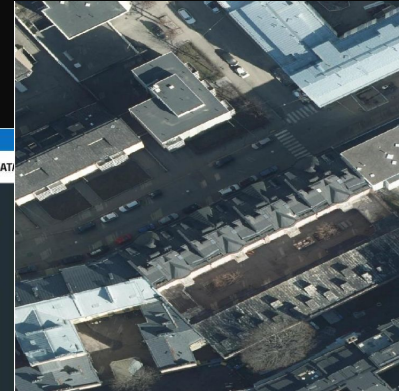
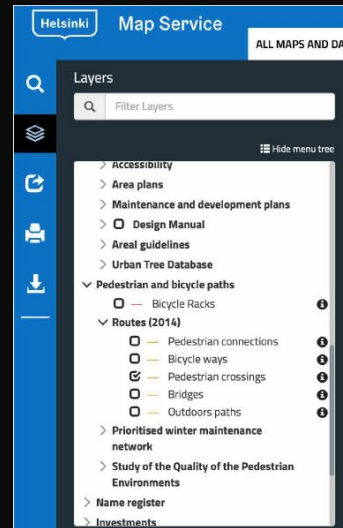
- Kumppanina **Integrify yritys**
- Periaatteina ja lähtökohtina avoimuus ja läpinäkyvyys sekä molemminpuolinen oppiminen - hyödynnetään kaupungin avointa dataa ja avointa lähdekoodia
- Hyödynnettyinä aineistoina toimivat:
 - Ortoilmakuvat
 - Vääräväri-ilmakuvat
 - Inventoidut suojatievektorit (2014)
 - Muut karttapalvelusta saatavat kartta-aineistot

Helsinki



We develop code(rs)

We teach international talent software development and connect them with work.



2. Kokeilun eteneminen

HANKINTA

1. Datan kerääminen

2. Koneoppimisen mallin luominen

3. Mallin arviointi ja kehitys

AINEISTO VISUALISOINTI

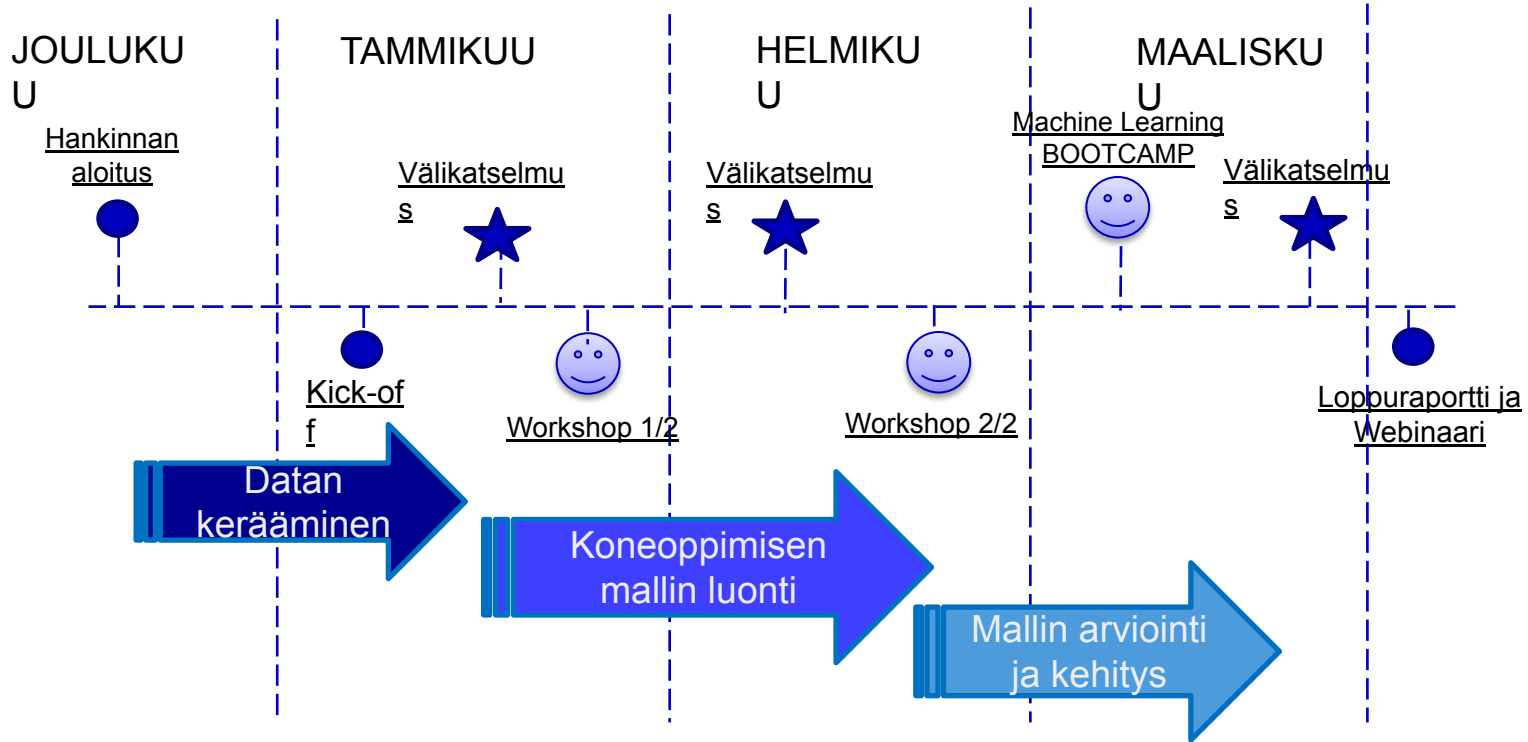
- 1. Datan kerääminen.** Kartta.hel.fi kaupungin karttapalvelusta kerättiin ohjelmallisesti n. 3000 ilmakuvaa Helsingin alueelta. Mallin tarkkuuden parantamiseksi olisi ollut parempi jos saatavilla olisi ollut enemmän dataa. Tilannetta auttoi se, että pystyttiin hyödyntämään myös Espoon ja Vantaan kaupungin ilmakuvia. Mitä enemmän ja laadukkaampaa dataa on saatavilla, sitä parempi malli saadaan aikaiseksi.

2. Koneoppimisen mallin luonti: Data jaettiin opetusdataan ja testidataan. Opetusdataan eli jokaiseen ilmakuvaan merkittiin manuaalisesti suojateiden paikka.

Malli luotiin hyödyntämällä opetusdataa suojateiden tunnistamiseen eri koneoppimisen arkkitehtuurien avulla joita hienosäädettiin eri parametrien avulla. Näitä olivat RetinaNet, faste R-CNN ja Mask R-CNN. Parhaan tarkkuuden tuotti RetinaNet.

3. Mallin arviointi ja kehitys: Testidatan avulla mallin toimivuus validoitiin. Malli pystyy tunnistamaan suojatien noin 70-80% tarkkuudella. Tätä tarkkuutta voidaan parantaa opetusdatan määrän lisäämisellä.

2. Kokeilun eteneminen

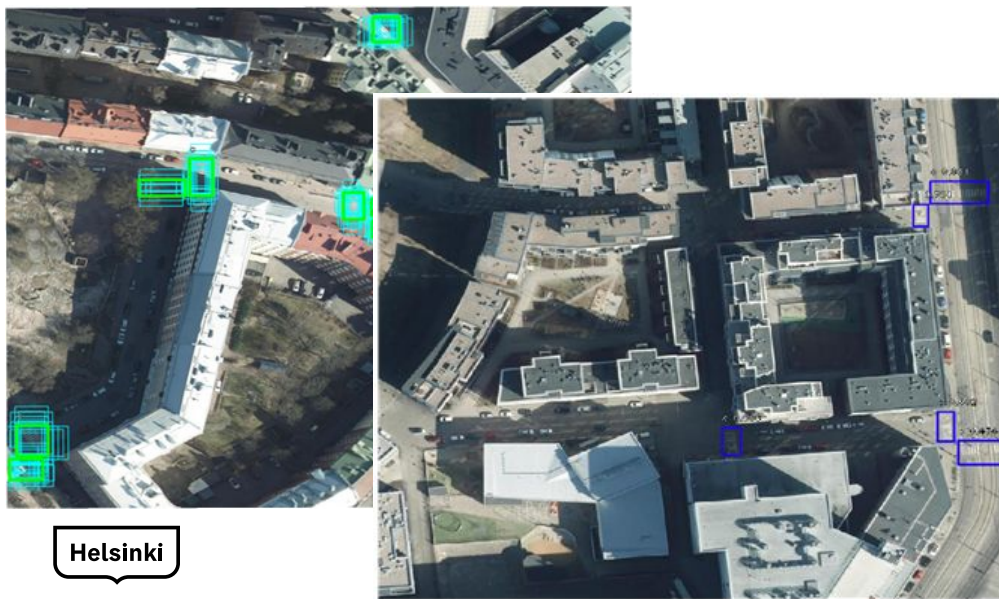


3. Kokeilun tuotokset

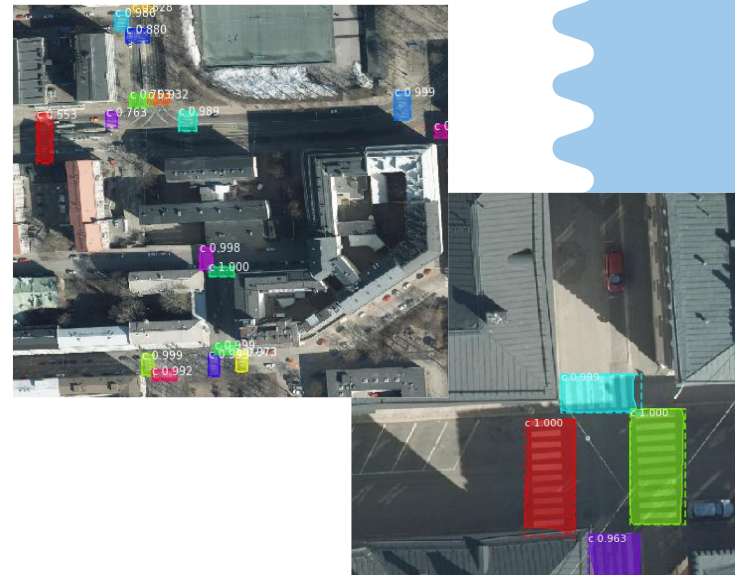
Kokeilun lopputuotoksena syntyi koneoppimisen malli jonka avulla pystytään tunnistamaan suojatiet ilmakuvista ja listaamaan niiden koordinaatit.

1. Tunnistetaan ilmakuvista todennäköiset suojatiet ja annetaan niille todennäköisyysarvot sekä määritellään tunnistuskehykset

RetinaNet



Mask R-CNN



3. Kokeilun tuotokset

2. Listataan koneoppimismallin tunnistetuista suojateistä tunnistuskehysten koordinaatit

	gmI_id	id	vaylatyyppi	luokka	featurecode	html_color	min_x	max_y	max_x
0	klinj_suojatie.16704	16704	Suojatie	Suojatiet	12	400040	2.549730e+07	6672929.268	2.549730e+07
1	klinj_suojatie.16713	16713	Suojatie	Suojatiet	12	400040	2.549717e+07	6673374.486	2.549718e+07
2	klinj_suojatie.16741	16741	Suojatie	Suojatiet	12	400040	2.549733e+07	6672931.001	2.549733e+07
3	klinj_suojatie.16908	16908	Suojatie	Suojatiet	12	400040	2.549736e+07	6672770.811	2.549735e+07
4	klinj_suojatie.16910	16910	Suojatie	Suojatiet	12	400040	2.549735e+07	6672758.490	2.549734e+07
...

- Eteen tulleita haasteita kokeilun edetessä oli:
 - Kvantunnistusmallien luotettavuus: esim. Vaaleat kattomateriaalit tunnistettiin virheellisesti suojateiksi –aineiston rajaamista esim. rakennusten osalta pohdittiin
 - Opetusdataa ei saatu tarpeeksi, jotta luokittelua olisi voitu tehdä luotettavasti.
 - Kokeilun alkuvaiheeseen kannattaa varata tarpeeksi aikaa. Erityisesti datan valmistelua ja nimiöintiä varten.

Helsinki

Aineistot saatavilla githubista: <https://github.com/City-of-Helsinki/road-marking-detector>.

Tekninen ympäristö:

MALLIN LUOMINEN: Python – ohjelmointikielen työkalut: **RetinaNet, faster-R-CNN ja mask R-CNN** syväoppimisen arkkitehtuureja.

DOKUMENTOINTI: **Jupyter Notebook –applikaatio**

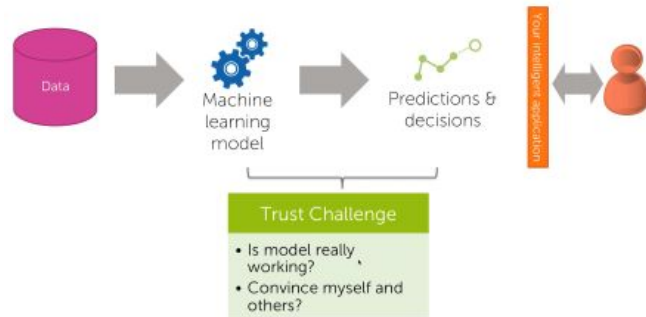
MALLIN HARJOITTAMINEN: **Amazon EC2 – pilvipalvelun avulla.**

4. Opit tekoälyn mahdollisuuksista

- Helsingin kaupungilla on paljon avointa dataa jota kannattaa hyödyntää palveluiden kehittämisessä. Jo pienillä panostuksilla voidaan saada tuloksia. Kokeilut ja haastekilpailut ovat helppo ja nopea tapa näiden mahdollisuuksien selvittämiseksi.
- Tärkeintä ehkä tässä kokeilemisessä on se, että se haastaa muuttamaan yhteisiä kaupungin sisäisiä toimintatapoja ja antanut uutta tietoa prosessin laadusta ja haasteista datan hyödyntämisestä läpi organisaatorajojen. Kaupunkiympäristöön liittyvän tiedon tallentamisessa, seurannassa, analysoinnissa ja arvioinnissa on haasteensa.

5. Opit tekoälyn kehittämisestä

- Koneoppimisen ratkaisun kehittäminen on **iteratiivinen prosessi** joka vaatii monen eri algoritmin testaamista hyvien ja luotettavien tulosten saavuttamiseksi.
- Koneoppimisen **mallin rakentaminen on paljon aikaa vievä prosessi**. Erityisesti datan manuaalinen nimeäminen vie aikaa. Tässä kokeilussa datan nimeäminen kesti n. 5 päivää ja mukana oli yhteensä 20 henkilöä. Mallin kouluttamiseen tarvitaan myös paljon **laskentatehoa**, kokeilun mallin kouluttamiseen käytettiin satoja tunteja aikaa.
- On tärkeää, että mallin kehittämisessä on mukana henkilö jolla on ratkaistavan haasteen alan osaamista.



5. Opit tekoälyn kehittämisestä

- Tekoäly toimii enemmänkin ”tukiälynä” – Ihminen on korvaamaton
- Ihmisen rooli korostuu etenkin datan muokkaamisessa ja validoinnissa. Ihminen pystyy opettamaan mallin toimintaa ja ohjaamaan oikeaan suuntaan.

”Paras malli on yhtä hyvä kuin käytetty opetusdata”

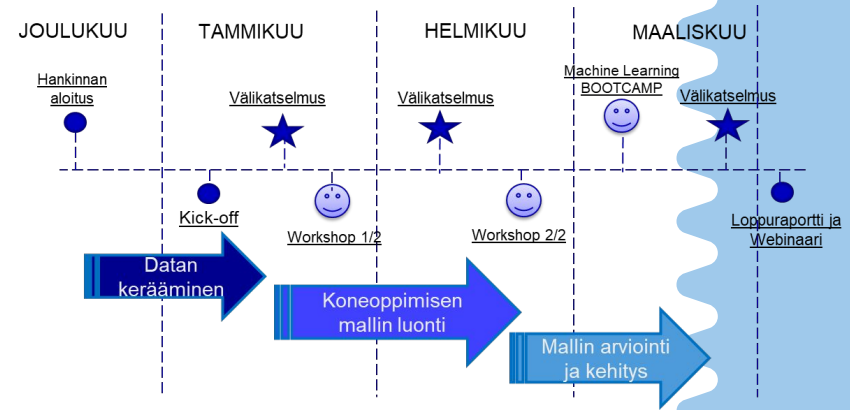
”60 % projektin ajasta kannattaa käyttää datan esitarkistukseen”
– Työ, tekoäly ja ihminen kurssi

6. Opit kokeilemisesta

- Kokeileminen antaa nopeasti ja suhteellisen pienillä resursseilla lisätietoa kehitettävän asian kannattavuudesta
- Hyvä kommunikointi kaikkien sidosryhmien välillä on tärkeää. Tässä kokeilussa oli käytössä slack – ryhmä jossa kaikki sidosryhmät olivat mukana. Tämä mahdollisti tehokkaan ja nopean kommunikaation.
- Toimivat verkostot – yhteistyö ja oikeiden henkilöiden kontaktointi tärkeää jo heti kokeilun alussa.
- Työpajat koneoppimisen mallin kehittäjien ja alan asiantuntijoiden eli Helsingin kaupungin kanssa olivat onnistunut tapa vaihtaa tietoa molemmin puolin ja viedä mallin kehitystä eteenpäin.
- Mallia voi kehittää pidemmälle käyttämällä enemmän kerroksia kartta.hel.fi palvelusta. Esimerkiksi vertaamalla ilmakuvia eri ajanjaksoilta voidaan suojateiden kuntoa ja niiden muuttumista seurata tarkemmin.

7. Opit resursoinnista

- Aikaa kokeiluun meni **n. 4 kk**
- Kokeilu suoritettiin aikataulussa ja budjetti riitti hyvin.
- Opetusdatan luomiseen meni yllättävän paljon aikaa, mitä ei ehkä alussa osattu ottaa huomioon. Kaikkia tavoitteita ei saavutettu, sillä opetusdataa ei saatu nimettyä tarpeeksi.
 - Datan keräämisessä mukana pääsääntöisesti **3 henkilöä**
 - Koneoppimisen mallin luonti
 - Datan manuaalinen nimeäminen kesti **n. 5 päivää. Mukana n. 20 henkilöä.**
 - Koneoppimisen mallin kehittämisessä mukana **3 henkilöä**
 - Mallin arviointi ja kehitys
 - Mallin kouluttamiseen **n. 200 h.**
- Kokeilun alkuvaiheeseen kannattaa varata tarpeeksi aikaa. Erityisesti datan valmistelua ja nimiöintiä varten.



8. Kokeilun tekninen ympäristö

- Mallin luomisessa käytettiin Python – ohjelmointikielen työkalut: RetinaNet, faster-R-CNN ja mask R-CNN syväoppimisen arkkitehtuureja.
- Dokumentoimiseen käytettiin Jupyter Notebook –applikaatiota.
- Amazon EC2 – pilvipalvelua käytettiin mallin harjoittamiseen.
- Tilaajan puolelta pystytetään palvelinympäristö, jossa aineistoa voidaan testata ja jatkokehittää.
- Aineistot saatavilla githubista: <https://github.com/City-of-Helsinki/road-marking-detector>. Githubista saatava dokumentaatio täydentää loppuraporttia.
- Vuosittain saatava inventointiaineisto on tarkoitus viedä karttapalveluun.

9. Jatkopäätökset ja -ideat

- Kokeilun tuloksia hyödynnetään testiaineistona toisessa kehittämissuunnitelmassa.
- Kokeilu oli toimiva ja siitä saatiin lisätietoa koneoppimisen hyödyntämiseen kaupunkiympäristön omaisuudenhallinnan kehittämisessä. Tarkoituksena on laajentaa kokeilua koskemaan mahdollisesti muitakin tiemerkinkeitä.
- Ensimmäisenä lähdetään oppilastyönä selvittämään tiemerkinkeitä kuntotietojen kartoitusta koneoppimisen keinoin.
- Mahdollisesti lähdetään edistämään tiemerkinkeitä tarkempaa luokittelua. Tämä vaatisi kuitenkin lisäresursseja, jotta tarvittava määrä opetusdataa saadaan luotua.