

Tien leveys AI

Pilottiprojektilla kohti kansallisten aineistojen
tehokäyttöä

Matti Pesu / Väylävirasto

Tekoäly

pilotti

Tiedon lajit

Yhteisen datan käyttö

Yhteistyö

Koneoppiminen (Machine Learning)

Valvottu oppiminen (Supervised Learning)

Regressio: Ennustetaan jatkuvaa arvoa. Esimerkkejä: Lineaarinen regressio, logistinen regressio

Luokittelu: Ennustetaan kategorista arvoa. Esimerkkejä: Tukivektorikoneet (SVM), päätöspuut, satunnaismetsät, k-lähimmät naapurit (k-NN), Naive Bayes

Valvomaton oppiminen (Unsupervised Learning)

Klusteroituminen: Ryhmitellään dataa klustereihin. Esimerkkejä: K-means, hierarkkinen klusterointi, DBSCAN

Dimensioiden vähentäminen: Vähennetään datan muuttujien määrää.

Esimerkkejä: PCA (Principal Component Analysis), t-SNE (t-Distributed Stochastic Neighbor Embedding)

Puoli-valvottu oppiminen (Semi-supervised Learning)

Yhdistää pieni määrä merkittävää dataa suureen määrään merkitsemätöntä dataa.

Vahvistusoppiminen (Reinforcement Learning)

Agentti oppii toimimaan ympäristössä palkkioiden ja rangaistusten perusteella.

Esimerkkejä: Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO)

Syväoppiminen (Deep Learning)

Käyttää monikerroksisia neuroverkkoja monimutkaisten mallien oppimiseen.

Syvä syöttöverkko

(Feedforward Neural Networks) Perustyyppi, jossa tiedot liikkuvat vain yhteen suuntaan kerroksesta toiseen.

Konvoluutioneuroverkot

Convolutional Neural Networks, CNN). Erikoistunut kuvien ja visuaalisten tietojen käsittelyyn.

Toistuvat neuroverkot

(Recurrent Neural Networks, RNN). Erikoistunut sekvenssidatan, kuten tekstin ja puheen, käsittelyyn. Esimerkkejä: LSTM (Long Short-Term Memory), GRU (Gated Recurrent Units)

Generatiiviset verkot

(Generative Networks) Tuottaa uutta dataa oppimansa jakauman perusteella.

Esimerkkejä: Generatiiviset vastakkaisverkot (GAN), Variational Autoencoders (VAE)

Tekoäly

Tilastolliset menetelmät

Käytetään tilastollisia teorioita ja malleja tietojen analysointiin ja ennustamiseen.

Bayesiläiset menetelmät

Naive Bayes, Bayesiläiset verkot

Markov-mallit

Piilotetut Markovin mallit (Hidden Markov Models, HMM)

Kombinaatiomenetelmät ja hybridimallit

Yhdistävät eri tekoälymenetelmiä parantaakseen suorituskykyä.

Esimerkkejä: Ensemble-menetelmät (bagging, boosting, stacking), hybridi-neuroverkot (käyttävät sekä CNN- että RNN-kerroksia)

Luonnollisen kielen käsittely (Natural Language Processing, NLP)

Menetelmät, jotka mahdollistavat koneiden ymmärtävän ja tuottavan inhimillistä kieltä.

Tekstin esikäsittely

Tokenisaatio, sanastojen luominen

Sanastojen ja lausekkeiden mallit

Bag of Words, TF-IDF

Syväoppimismallit

Word2Vec, BERT, GPT

Evoluutiomenetelmät

Käytetään luonnonvalinnan periaatteita optimointiongelmien ratkaisemiseksi.

Geneettiset algoritmit, evoluutiostrategiat



Koneoppiminen (Machine Learning)

Valvottu oppiminen (Supervised Learning)
 Regressio: Ennustetaan jatkuvaa arvoa. Esimerkkejä: Lineaarinen regressio, logistinen regressio
 Luokittelu: Ennustetaan kategorista arvoa. Esimerkkejä: Tukivektorikoneet (SVM), päätospuut, satunnaisetsat, k-lahimmat naapurit (k-NN), Naive Bayes

Valvomaton oppiminen (Unsupervised Learning)
 Klusterointilumen: Ryhmitellään dataa klustereihin. Esimerkkejä: K-means, hierarkkinen klusterointi, DBSCAN
 Dimensoiden vähentäminen: Vähennetään datan muuttujien määrää. Esimerkkejä: PCA (Principal Component Analysis), t-SNE (t-Distributed Stochastic Neighbor Embedding)

Puoli-valvottu oppiminen (Semi-supervised Learning)
 Yhdistää pieni määrä merkittävää dataa suureen määrään merkitsemätöntä dataa.

Vahvistusoppiminen (Reinforcement Learning)
 Agentti oppii toimimaan ympäristössä palkkioiden ja rangaistusten perusteella. Esimerkkejä: Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO)

Luonnollisen kielen käsittely (Natural Language Processing, NLP)

Menetelmät, jotka mahdollistavat kielten ymmärtävän ja tuottavan inhimillistä kieltä.

Tekstin esikäsittely
 Tokenisaatio, sanastojen luominen
 Sanastojen ja lausekkeiden mallit
 Bag of Words, TF-IDF

Syvaoppimismallit
 Word2Vec, BERT, GPT

Tilastolliset menetelmät

Käytetään tilastollisia teorioita ja malleja tietojen analysointiin ja ennustamiseen.

Bayesiläiset menetelmät
 Naive Bayes, Bayesiläiset verkot

Markov-mallit
 Piilotetut Markovin mallit (Hidden Markov Models, HMM)

Kombinaatiomenetelmät ja hybridimallit

Yhdistävät eri tekoälymenetelmiä parantaakseen suorituskykyä.
 Esimerkkejä: Ensemble-menetelmät (bagging, boosting, stacking), hybridi-neuroverkot (käyttävät sekä CNN- että RNN-kerroksia)

Evoluutiomenetelmät

Käytetään luonnonvalinnan periaatteita optimointiongelmien ratkaisemiseksi.
 Geneettiset algoritmit, evoluutiostrategiat

Tekoäly Syväoppiminen (Deep Learning)

Käyttää monikerroksisia neuroverkkoja monimutkaisten mallien oppimiseen.

Syvä syöttöverkko

(Feedforward Neural Networks) Perustyyppi, jossa tiedot liikkuvat vain yhteen suuntaan kerroksesta toiseen.

Konvoluutioneuroverkot

Convolutional Neural Networks, CNN).

Erikoistunut kuvien ja visuaalisten tietojen käsittelyyn.

Toistuvat neuroverkot

(Recurrent Neural Networks, RNN). Erikoistunut sekvenssidatan, kuten tekstin ja puheen, käsittelyyn.

Esimerkkejä: LSTM (Long Short-Term Memory), GRU (Gated Recurrent Units)

Generatiiviset verkot

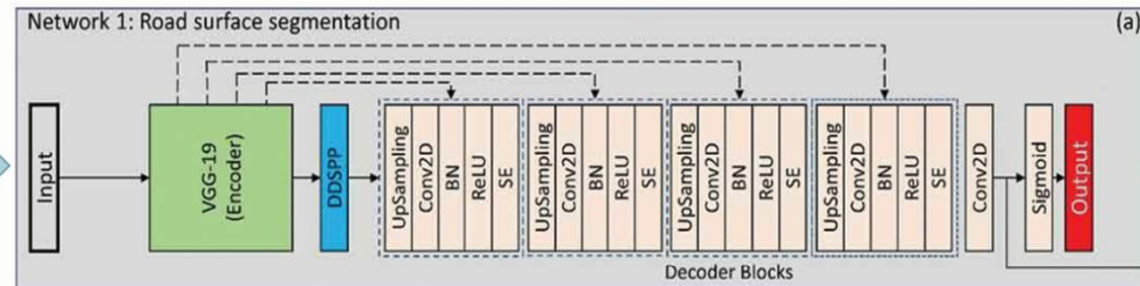
(Generative Networks) Tuottaa uutta dataa oppimansa jakauman perusteella. Esimerkkejä: Generatiiviset vastakkaisverkot (GAN), Variational Autoencoders (VAE)



pilotti : Malli

RoadVecNet

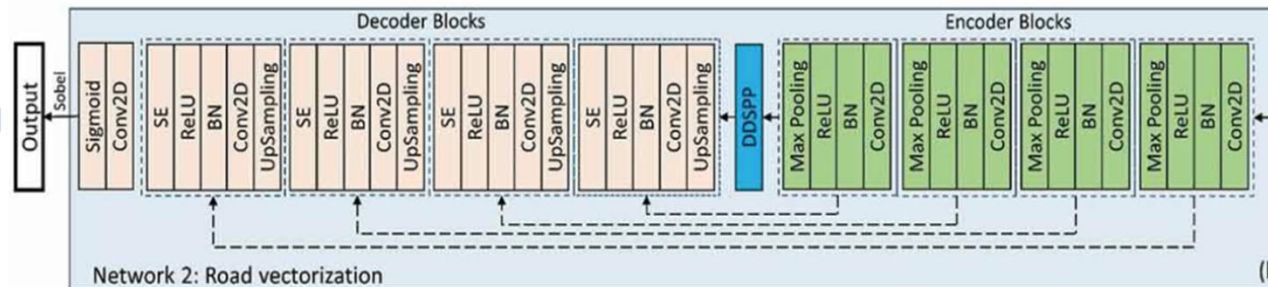
Syöte:
ortokuva



1.Tuotos:
Segmentoitu
tiestö

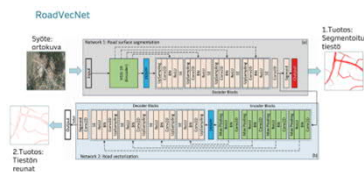


2.Tuotos:
Tiestön
reunat



pilotti : Opetusdata

Malli



Aiemmasta ATMU-projektista:

MML 2020 ortokuvat. 3456 kuvaa: 2592 koulutusaineisto, 518 testiaineisto, 346 vertailuaineisto.

Tiestömaski, Digiroad keskilinjat ja leveys.

Tässä projektissa

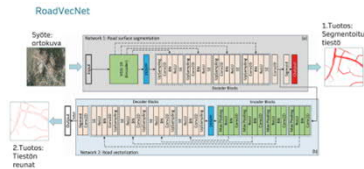
2023 ortokuvia.

Lidar-tuotteita: hillshade, aspect, slope, DTM, intensity, roughness



pilotti : Alustavia tuloksia

Malli



Slopella koulutettu malli sai parhaat tulokset, hillshade ja aspect heikoimmat.

	RoadVecNet						
	RGB	hillshade	aspect	slope	DTM	intensity	Roughness
F1	68,96	66,92	66,68	71,74	69,47	69,86	70,53
Edge F1	56,28	55,51	55,26	57,12	56,46	56,64	56,64
Val f1	67,91	67,54	66,69	70,54	68,55	70,01	69,12
edge val f1	55,69	55,07	55	56,58	55,96	56,05	56,02

Alemmat tieluokat ja polut hankalia, samoin leveät monikaistaiset tiet. Lidar auttaa peitteisissä ja varjoisissa paikoissa.

Opetusdata

Aiemmasta ATMU-projektista:

MML 2020 ortokuvat. 3456 kuvaa: 2592 koulutusaineisto,
518 testiaineisto, 346 vertailuaineisto.
Tiestömaski, Digiroad keskilinjat ja leveys.

Tässä projektissa

2023 ortokuvia.
Lidar-tuotteita: hillshade, aspect, slope, DTM, intensity,
roughness



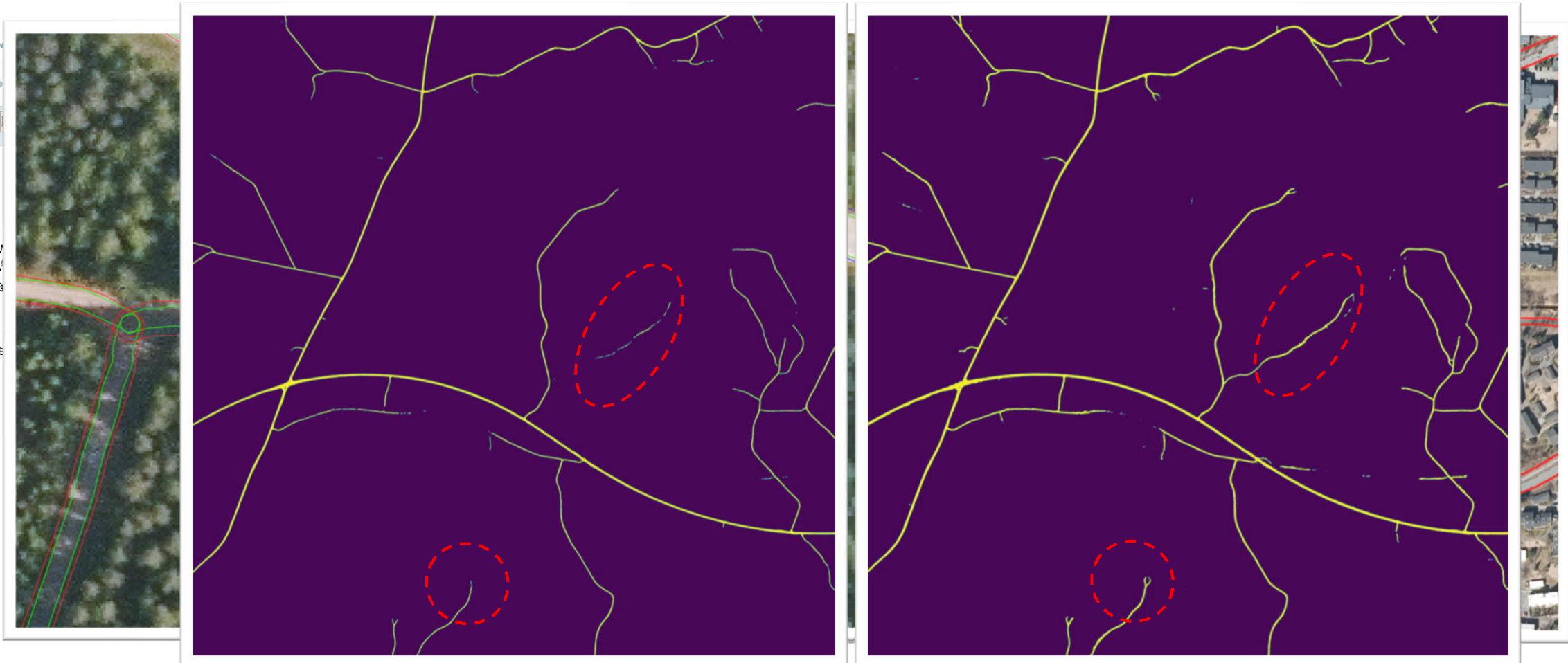
pilotti : Alustavia tuloksia

Malli

Opetus

Aiemmasta

Tässä proje



Tiedon lajit

Rekisteritieto

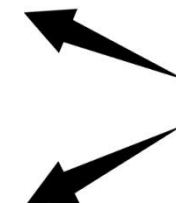


Ensisijaisia

Tien reuna
Leveys ja päällyste
Keskilinja

Toissijaisia

Asfaltin reuna
Ajoratamaalaukset
Kohtaamispaikka
Kapea kohta
Kävely- ja pyörätien luokka
Kääntöpaikka



Vihjetieto



Tiedon lajit

Tien reuna
Leveys ja päällyste
Keskilinja

Asfaltin reuna
Ajoinalaukset
Kohtaamispaikka
Kapea kohta
Kävely- ja pyörätien luokka
Kääntöpaikka

Rekisteritieto

Vihjetieto

Väylien suunnittelu / rakentaminen / kunnossapito

Omaisuuksienhallinta

Digitaalinen kaksonen

Ennusteet

Ajantasainen ja
automaattinen tietovirta



Yhteisen datan käyttö

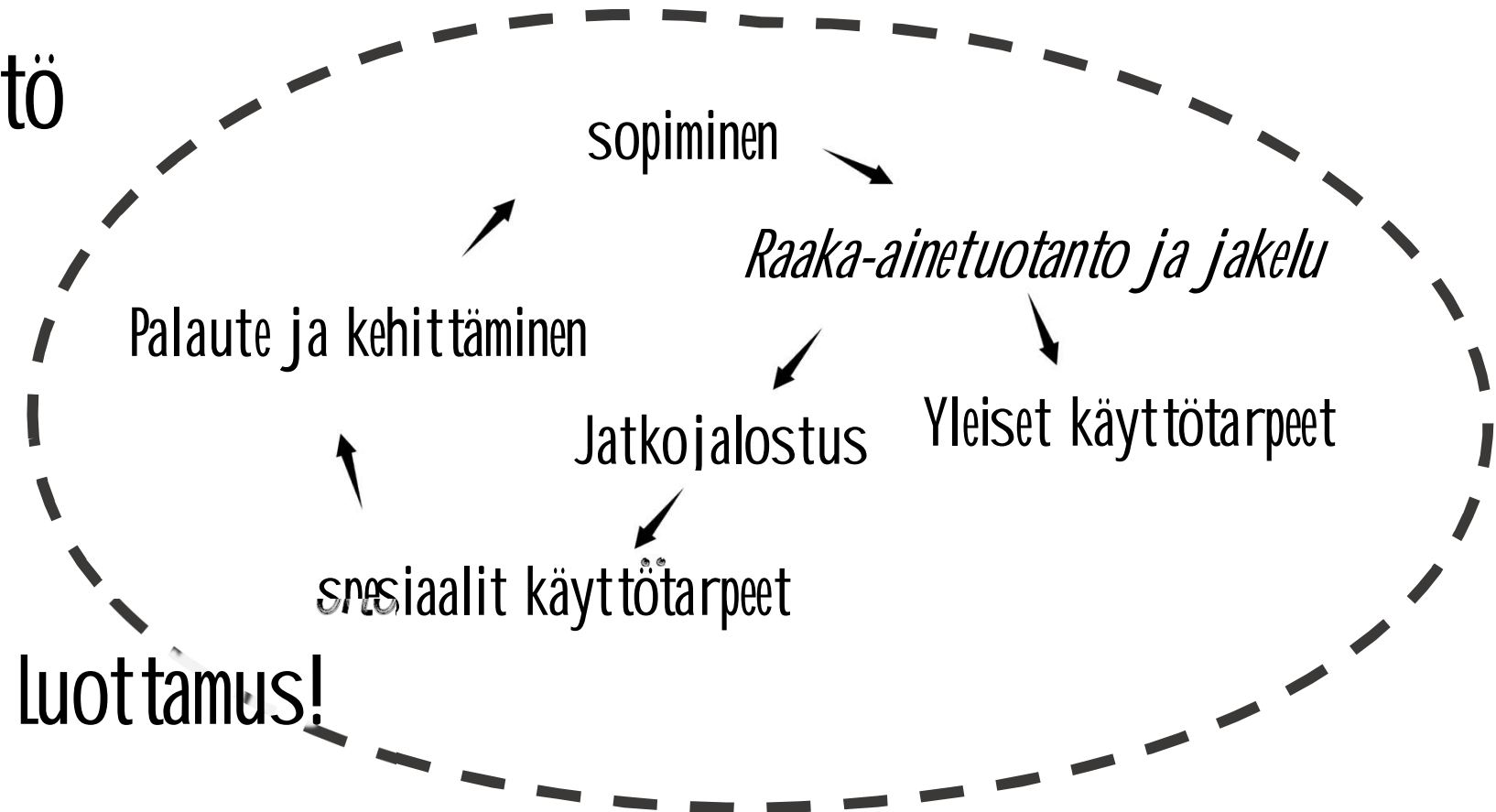
Kansallisia laserkeilaus- ja ilmakuvausohjelmia ohjaa ja rahoittaa KALLIO-niminen yhteistyöelin. Kansallisia ohjelmia rahoittavat Maanmittauslaitoksen lisäksi [Suomen metsäkeskus](#), [Metsähallitus](#), [Ruokavirasto](#) ja [Suomen ympäristökeskus](#).

Yhteistyöelimessä ovat lisäksi edustettuina maa- ja metsätalousministeriö, ympäristöministeriö ja puolustusministeriö.

Kansalliset laserkeilaus- ja ilmakuvausohjelmat on suunniteltu siten, että ne palvelevat kaikkien yhteistyöelimessä toimivien tarpeita. Kansallisista ohjelmista saatavia aineistoja hyödynnetään esimerkiksi 3D-rakennusten tuottamisessa, metsävaratiedon keruussa sekä EU-peltovalvonnassa ja tulvakartoituksessa.



Yhteisen datan käyttö



Yhteistyö = Luottamus + yhteinen tarve

Luottamus = $\frac{\text{Uskotavuus} + \text{luotettavuus} + \text{aitous}}{\text{itsekkyyys}}$

