



Vesihuoltolaitosten
kehittämisrahasto



Esiselvitys tekoälyn hyödyntämisestä vesihuollossa

Projektikuvaus

Hankkeen esittely

- Hankkeen tavoitteena oli analysoida tekoälyn (Artificial Intelligence) potentiaalia, uhkia ja haasteita vesihuollossa sekä tunnistaa vesihuoltoalan yleiset valmiudet hyödyntää tekoälyä laitosarjessa. Yksi hankkeen tärkeimmistä tavoitteista oli tunnistaa kohteet, joissa tekoälyä kannattaa hyödyntää ja mitä AI:n hyödyntäminen vaatii. Hankkeessa laadittiin julkaisu *Esiselvitys tekoälyn hyödyntämisestä vesihuollossa*.
- Hankkeen yhteydessä haastateltiin useita vesihuoltoalan toimijoita sekä mm. energiayhtiöitä.
- Hankkeen ohjausryhmä piti kolme työpajaa, joissa käsiteltiin mm. datan laadun merkitystä, tietoturvaa ja olemassa olevia tekoälyratkaisuja.
- Hankkeen ohjausryhmään kuuluivat HS-Vedeltä Jarno Laine, Lahti Aqualta Sauli Pihamaa ja Janne Mäki-Petäjä, Kymen Vedeltä Jani Väkevä ja Joonas Leinonen, Napapiirin Vedeltä Joni Hemmilä, Kuusamon energia- ja vesiosuuskunnasta Kimmo Viinikka sekä VVY:ltä Saijariina Toivikko. Swecolta hankkeen työryhmässä toimivat pääasiassa Jaana Pulkkinen, Joonas Lepistö ja Anna Rouhiainen.
- Oman panoksensa hankkeeseen antoivat Aalto-yliopiston työelämäprofessori, Fluidit Oy:n CTO Markus Sunela sekä Traficomien Kyberturvallisuuskeskuksesta Kimmo Saarni, Jukka Isosaari ja Aleksis Blomqvist.



Julkaisun sisällysluettelo

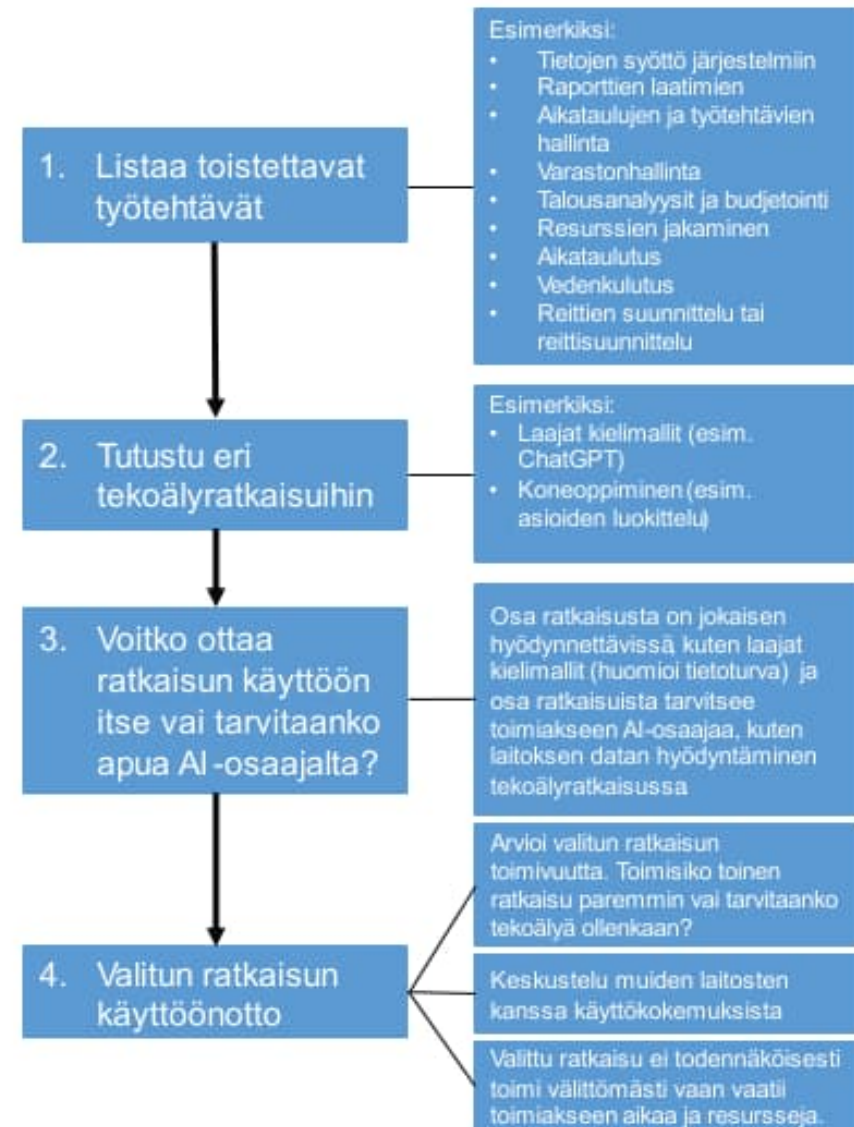
- Tekoäly on yleistermi laajalle joukolle erilaisia todennäköisyysfunktioihin perustuvia ratkaisuja, joista tällä hetkellä pinnalla ovat etenkin laajat kielimallit ja ChatGPT. Jotta tekoälyn laajuus olisi hahmotettavissa, koettiin tärkeäksi esitellä erilaisia tekoälyn ratkaisuja: koneoppimista, neuroverkkoja sekä luonnollisen kielen käsittelyä.
- Tekoälyä pohjautuu dataan ja sen käsittelyyn, joten datan merkitystä ei voi olla korostamatta. Julkaisuun pyrittiin saamaan tiivis, mutta kattava paketti datan osalta tärkeimmistä asioista.
- Julkaisussa käydään kevyesti läpi myös tekoälyratkaisujen energiankulutusta sekä tekoälyä käytettäessä huomioitavia eettisiä näkökulmia ja lainsäädäntöä. Työskennellessä AI-ratkaisuiden parissa on lisäksi aina huomioitava tieto- ja kyberturva.

Sisällysluettelo

1	Johdanto	7
1.1	Mitä on tekoäly?	8
1.2	Tekoälyn historia	9
1.3	Tekoälyn hyödyntäminen organisaatioissa	10
2	Data	12
2.1	Laatu ja määrä	12
	CASE – Laadukkaan datan merkitys ja sen hyödyntäminen liiketoiminnassa	13
2.2	Luokittelu	14
	CASE – Tekoälyn hyödyntäminen kasvihuonekaasupäästöjen vähentämisessä	15
2.3	Massadata (Big Data)	16
2.4	Algoritmit	17
	CASE – Poikkeamien havaitseminen ja ennustaminen datan avulla	18
	CASE – Ennusteita ihmisten hyödynnettäväksi	19
	CASE – Tekoälyn rooli vesihuollon tehostamisessa	20
2.5	Tärkeimmät nostot: Data	21
3	Koneoppiminen	22
	CASE – Tekoäly jätevedenpuhdistuksen tehostajana	24
	CASE – Algoritmit ja koneoppimismalli vesihuoltolaitoksen apuna	25
3.1	Neuroverkot ja syväoppiminen	26
3.2	Tärkeimmät nostot: Koneoppiminen	28
	CASE – Vesijohtoverkoston vuotojen havaitseminen Lontoossa	29
4	Luonnollisen kielen käsittely (NLP)	30
4.1	Suuret kielimallit (Large Language Models)	31
4.2	Arviointi ja validointi	34
	CASE – Tekoälyn hyödyntäminen operatiivisessa tehokkuudessa	35
	CASE – Asiakaspalvelun työtehtävien tehostaminen	36
4.3	Tärkeimmät nostot: Luonnollisen kielen käsittely (NLP)	37
5	Energiankulutus ja kestävyys	38
5.1	Tärkeimmät nostot: Energiankulutus ja kestävyys	39
6	Turvallisuus	40
6.1	Etiikka	41
6.2	Lainsäädäntö	41
6.3	Tietosuoja	42
6.4	Tietoturva	43
6.5	Kyberturva	44
6.6	Tärkeimmät nostot: Turvallisuus	45
7	Yhteenveto: Tekoäly vesihuollossa	46
7.1	Tulevaisuus	48
	Liite 1 – AI-sanasto	49

Avuksi tekoälyn hyödyntämismahdollisuuksien hahmottamiseen

- Vaikka kyseessä on esiselvitystasoinen raportti, pyrittiin vesihuoltolaitoksille laatimaan askelmerkkejä tekoälyn hyödyntämiseksi (kuva 1). Hankkeen tärkeimpiä viestejä on, että tekoälysovellus on vain yksi osa prosessia eikä ratkaise kaikkia haasteita kerralla. Lisäksi on tärkeää muistaa, että tekoälysovellusta on koulutettava ennen kuin se vastaa vesihuoltolaitoksen tarpeisiin.



Kuva 1. Miten aloittaa tekoälyn hyödyntäminen.

Nostoja julkaisusta

- Julkaisuun tuotiin konkretiaa esittelemällä eri toimijoilta saatuja case-esimerkkejä. Case-esimerkit ovat omissa laatikoissaan ja niihin on myös lisätty kuvakkeet selkeyttämään, mitä tekoälyratkaisua kussakin on hyödynnetty.
- Jotta julkaisu tavoittaisi myös kiireiset lukijat, laadittiin jokaisen teeman osalta tärkeimmät nostot, joiden avulla aiheeseen pääsee hieman käsiksi lukematta koko julkaisua.

CASE – ASIAKASPALVELUN TYÖTEHTÄVIEN TEHOSTAMINEN



Luonnollisen kielen käsittely Puhe

”Lahti Aqua on tehostanut asiakaspalvelua tekoälyn sekä ohjelmistorobotin avulla. Laajaan kielimalliin perustuvaa tekoälysovellusta hyödynnetään Lahti Aquassa sähköposteihin vastaamisessa sekä ohjelmistorobottia vesimittarilukematietojen kirjaamisessa asiakaspalvelujärjestelmään.

Lahti Aqua on selvittänyt vuosittain asiakkailtaan mieluisimpia yhteydenottotapoja, joista sähköposti on Minun Aquani -palvelun rinnalla suosituin. Lahti Aquan asiakaspalvelun yhteiskäyttö-sähköpostiin saapuukin vuosittain noin 12 000 postia.

Tekoälysovellus muodostaa saapuneisiin sähköposteihin vastausehdotuksen, jonka asiakaspalveluhenkilöstö tarkastaa ja tarvittaessa muokkaa ennen lähettämistä asiakkaalle.

Tekoälysovellus on suunniteltu vuorovaikutteiseen keskusteluun ja sen käyttämään äänensävyyn sekä vastausten pituuteen pystytään vaikuttamaan. Tekoälysovelluksen kielimalli perustuu OpenAI ChatGPT:n. Mallia tulee siis kouluttaa, jotta se vastaa vesilaitok-

3.2 TÄRKEIMMÄT NOSTOT: KONEOPPIMINEN



Koneoppimisen peruseriaatteet

Koneoppiminen mahdollistaa tietokoneiden oppimisen ja päätöksenteon datan perusteella itsenäisesti. Datan laatu on kriittinen tekijä, sillä huono data johtaa huonoihin tuloksiin.

- Datan laatu ja määrä
- Ongelman määrittely ja mallin tarkoitus
- Sopivat algoritmit
- Koulutus ja testaus



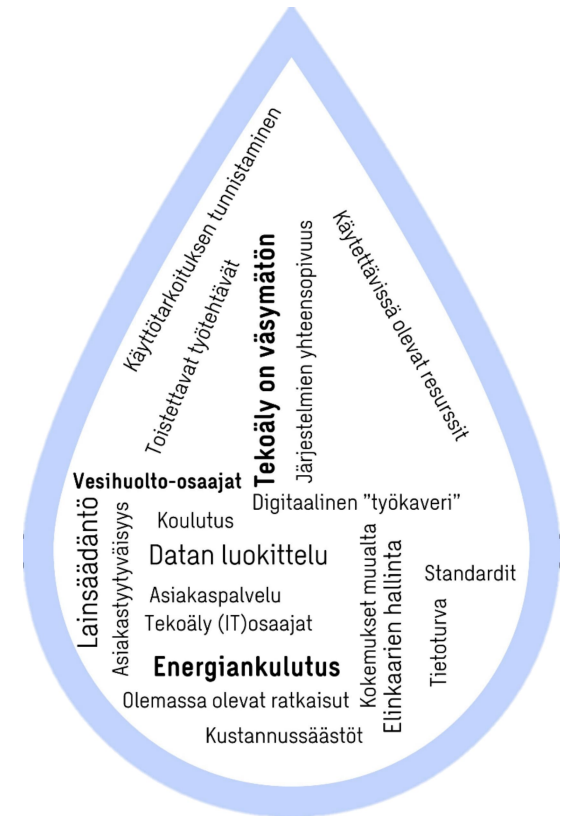
Koneoppimisprosessin vaiheet

Mallia kehitetään kouluttamalla algoritmeja datan avulla. Koulutusdata opettaa mallia, ja testidataa käytetään suorituskyvyn arviointiin.

- Ongelman määrittely
- Datan keruu, esikäsittely ja jakaminen
- Mallin koulutus
- Seuranta ja ylläpito

Tuloksia

- Tekoälyn hyödyntämiselle vesihuollossa on suuri potentiaali ja monia ratkaisuja on jo käytössä. Tekoälyä voikin verrata työkalupakkiin, josta löytyy erilaisia työkaluja, eli tekoälyratkaisuja.
- Tekoälyratkaisun valitsijan on tunnistettava erilaiset ratkaisut sekä tunnettava haaste, joka on tarpeen ratkaista, jotta osaa valita oikean työkalun. Käytettävän tekoälyratkaisun valintaan vaikuttaa, pohjautuuko haaste esimerkiksi automaattiodataan vai asiakaspalveluun.
- Tekoälyratkaisu on osa prosessia, ei yleisvastaus kaikkiin ongelmiin. Ongelma voi myös olla kannattavampaa ratkaista muuten kuin tekoälyn avulla.
- Tekoälyn käyttöönottoon sekä käyttöön on varattava resursseja, sillä ratkaisu on koulutettava jokaisen vesilaitoksen tarpeisiin. Kouluttaminen vaatii vesihuollon sekä tekoälyn substanssiosaamista. Usein tekoälyratkaisun toteuttamiseen tarvitaan osaamista vesilaitoksen ulkopuolelta.
- Tekoälyratkaisusta saatavien tulosten luotettavuus on riippuvaista lähtödatan laadusta. Käytettävän datan laatuun ja määrään on siis kiinnitettävä huomiota.
- Tekoälyä käytettäessä on aina kiinnitettävä huomiota tieto- ja kyberturvaan. Lisäksi saatuihin vastauksiin tulee suhtautua kriittisesti.
- Tekoälyn vahvuuksia ovat väsymättömyys ja tarkkaavaisuus, nopeus, havainnointi ihmisaistien ulkopuolella, rutiinitehtävien toteutus, laskeminen (huom! Ei ChatGPT:n vahvuus), lyhytaikainen muisti ja yhtäaikaisten asioiden tekeminen.



Transforming society together

