

# Tekoäly- ja ohjelmistorobotiikkakokeilun loppuraportti - syksy 2020

ONNI-osaamistarvetiedon käsittelyn tekoälykokeilu

Helsinki



Tukea digitalisaatiokokeiluihin kaupungin työntekijöille

# ONNI-osaamistarvetiedon käsittelyn tekoälykokeilu

## Tiimi:

Sanna Lyly, Kanslia  
Päivi Mäkeläinen, Kanslia  
Nina Suursalmi, Kanslia  
Johanna Lehtimäki, Kanslia  
Liisa Lind – Jukka Orava, Kasko  
Ulla Väisänen, Sote

## Gofore:

Pasi Lehtimäki  
Tommi Vehviläinen  
Teemu Hirsimäki

# 1. Kokeilun onnistuminen

- Kokeilu onnistui pääosin eli

Antoi aavistuksen siitä, että kunhan sisällön osaajat pääsevät operoimaan käyttöliittymällä datan pohjalta muodostettuja klustereita, data on hyödynnettävissä osaamisen kehittämisen suunnittelussa.





## 2. Kokeilun eteneminen

- Tavoitteena tunnistaa automaattisesti osaamistarpeita onnistumiskeskusteluissa kerättyjen vastausten perusteella ja saada näistä organisaatiokohtainen yhteenveto sekä osayhteenvetoja.
- Alussa anonymisoitiin vastaukset poistamalla nimet sekä mahdolliset henkilötunnukset ja puhelinnumerot.
- Kokeilussa testattiin eri menetelmiä tekstiaineiston analysointiin. Tarkoituksena oli tunnistaa onnistumiskeskusteluissa esille nousseita osaamistarpeita ja ryhmitellä niitä.
- Lopulta tehtiin käyttöliittymä, jonka avulla loppukäyttäjä voi käydä läpi eri aiheklustereita, muuttaa vastausten luokittelua ja luoda raportin eri osaamisten yleisyydestä.

# 3. Kokeilun tuotokset

- Onnistumiskeskusteluiden vastaukset osaamistarpeista jaettiin lauseisiin, ja ne ryhmiteltiin automaattisesti samankaltaisuuden mukaan.
- Rakennettiin prototyyppi käyttöliittymästä, jolla loppukäyttäjä voi tarkastella, muuttaa ja uudelleennimetä mallin luomia klustereita sekä luoda uusia tai poistaa olemassaolevia klustereita.
- Työkalun tarkoitus on nopeuttaa datan läpikäyntiä, nopeuttaa helposti hahmotettavan kokonaiskuvan luomista sekä kerätä dataa, jota voidaan myöhemmin käyttää mallin parantamiseen.

- (1) Apotin IT-oppaana toimiminen  
 (2) Apotin hallitseminen  
 (3) Apotin haltuunottaminen.  
 (4) Apotin haltuunotto.  
 (5) Apotin ja erilaisten tilausjärjestelmien tunteminen  
 (6) Apotin jalkauttaminen.  
 (7) Apotin käyttö  
 (8) Apotin käyttö  
 (9) Apotin käyttö.  
 (10) Apotin käyttö.  
 (11) Apotin käyttöä.  
 (12) Apotin käyttöönottoon liittyvä osaaminen (Metropolia -oppimisympäristö ym. Apotti- oppimi  
 (13) Apotin käytön osaaminen, kun se alkaa.  
 (14) Apotin omaksuminen.  
 (15) Apotin osaaminen  
 (16) Apotti  
 (17) Apotti  
 (18) Apotti  
 (19) Apotti  
 (20) Apotti  
 (21) Apotti  
 (22) Apotti  
 (23) Apotti  
 (24) Apotti  
 (25) Apotti  
 (26) Apotti  
 (27) Apotti  
 (28) Apotti  
 (29) Apotti  
 (30) Apotti  
 (31) Apotti  
 (32) Apotti  
 (33) Apotti  
 (34) Apotti  
 (35) Apotti  
 (36) Apotti järjestelmän haltuunotto.  
 (37) Apotti -osaamisen haltuun otto.  
 (38) Apotti -potilastietojärjestelmän opettelu.  
 (39) Apotti .  
 (40) Apotti 2.  
 (41) Apotti alkaa v 2021 keväällä ja siihen liittyvät koulutukset ja perehtyminen sekä käyttöönotto.  
 (42) Apotti asiakastietojärjestelmä kevät 2021, sitä ennen Apotti Demopassi koulutus, Turvallisess  
 (43) Apotti esimerkiksi . palvelusetelijärjestelmä  
 (44) Apotti haltuun  
 (45) Apotti ja Tunstall osaamista.  
 (46) Apotti ja mahdolliset tulevat uudet tietojärjestelmät.  
 (47) Apotti ja siihen liittyvä tiedon hankinta.  
 (48) Apotti järjestelmä  
 (49) Apotti järjestelmän ymmärtämistä ja käytön osaamista ja mahdollisesti jatkossakin toiveena  
 (50) Apotti käyttöönotto, siihen liittyvät koulutukset helmi-toukokuu/2021

<input type="checkbox"/>	hallinta (146)	Nimeä uudelleen	
<input type="checkbox"/>	tarvita (185)	Nimeä uudelleen	
<input type="checkbox"/>	apotti (244)	Nimeä uudelleen	
<input type="checkbox"/>	haavanhoito (49)	Nimeä uudelleen	
<input type="checkbox"/>	tieto (137)	Nimeä uudelleen	
<input type="checkbox"/>	koulutus (141)	Nimeä uudelleen	
<input type="checkbox"/>	työntekijä (204)	Nimeä uudelleen	
<input type="checkbox"/>	enempi (103)	Nimeä uudelleen	
<input type="checkbox"/>	tieto (190)	Nimeä uudelleen	
<input type="checkbox"/>	pyrkii (163)	Nimeä uudelleen	
<input type="checkbox"/>	hoito (117)	Nimeä uudelleen	
<input type="checkbox"/>	tieto (169)	Nimeä uudelleen	
<input type="checkbox"/>	asiakas (90)	Nimeä uudelleen	
<input type="checkbox"/>	järjestelmä (128)	Nimeä uudelleen	
<input type="checkbox"/>	johtaa (73)	Nimeä uudelleen	
<input type="checkbox"/>	tietotekninen (112)	Nimeä uudelleen	
<input type="checkbox"/>	lapsi (112)	Nimeä uudelleen	
<input type="checkbox"/>	rai (244)	Nimeä uudelleen	
<input type="checkbox"/>	koulutus (178)	Nimeä uudelleen	
<input type="checkbox"/>	2020 (29)	Nimeä uudelleen	
<input type="checkbox"/>	tiimi (78)	Nimeä uudelleen	
<input type="checkbox"/>	palvelu (123)	Nimeä uudelleen	
<input type="checkbox"/>	tietotekninen (118)	Nimeä uudelleen	
<input type="checkbox"/>	enempi (101)	Nimeä uudelleen	
<input type="checkbox"/>	mielenterveys (111)	Nimeä uudelleen	
<input type="checkbox"/>	kirjata (79)	Nimeä uudelleen	
<input type="checkbox"/>	lääkehoito (149)	Nimeä uudelleen	
<input type="checkbox"/>	koulutus (112)	Nimeä uudelleen	
<input checked="" type="checkbox"/>	apotti (138)	Nimeä uudelleen	
<input type="checkbox"/>	atk (62)	Nimeä uudelleen	
<input type="checkbox"/>	asiakas (13)	Nimeä uudelleen	

# 4. Opit tekoälyn mahdollisuuksista

- Kaupungin eri palveluissa tallentuu valtavat määrät erilaista vapaatekstimuotoista dataa. Tekoälyn mahdollisuuksia tämän datan hyödyntämisessä on syytä pohtia jatkuvasti.
- Tekoälyä voi hyödyntää, mutta ihmisen tuki tässä sovelluksessa on tarpeen. Tästä syystä päädyttiinkin asiantuntijan ryhmittelyä tukevan työkalun toteuttamiseen.
- Jos organisaatio käyttäisi työkalua säännöllisesti, työkalu voisi oppia tekemään ryhmittelyn automaattisemmin.

# 5. Opit tekoälyn kehittämisestä

- Datan anonymisoinnista teki haastaavaa se, että monella järjestelmällä on ihmisen nimi.
- Teksiaineisto oli haastava, osaamistarpeita kuvaavat tekstit sisällöltään hyvin erilaisia ja eritasoisia.
- ONNI-datalla opetetut kielimallit eivät toimineet yhtä hyvin tässä sovelluksessa kuin ulkopuolisilla, suurilla tekstiaineistoilla opetetut kielimallit.
- Sanatasoinen analyysi ei toiminut yhtä hyvin kuin esim. sisällön lausetasoinen ryhmittely. Lausetason ryhmittely huomioi vastauksen kontekstin paremmin.
- Parhaiten toimi monikielinen transformer-kielimalli, joka huomioi lauseen kontekstin.
- Tekoälyn hyödyntämisessä on tärkeää, että käyttöliittymä suunnitellaan tukemaan sekä käyttäjää että tekoälymallin jatkuvaa oppimista.



# 6. Opit kokeilemisesta

- Viikkopalaverit yhteistyökumppanin kanssa tuntuivat toimivan.
- Organisaatiosta kannattaa yrittää saada mukaan eri toimialojen ihmisiä.
- Kokeilijaryhmän kannattaa tarttua yhteistyökumppanin tarjoamiin mahdollisuuksiin lisätä omaa osaamistaan kokeilun kohteena olevasta aihepiiristä.
- Sopimusasiat olisi hyvä saada aiemmin sovittua
- Hyvä, ettei tavoitetta kiinnitetä liian tiukasti - jää tilaa kokeilulle

# 7. Opit resursoinnista

- Koska keskeinen toimija tässä kokeilussa oli yhteistyökumppani, ajankäyttö ei tuottanut haasteita kaupungilla.
- Tehtävä oli haastavampi kuin alunperin kuviteltiin ja työmäärä oli suunniteltua suurempi.
  - Suurin osa ajasta kului riittävän hyvän tekoälymenetelmän löytämiseen.
  - Hyvä, että oli useamman tekijän tiimi ideoimassa ja suunnittelemassa.
- Jos käyttöliittymän prototyyppiä olisi päästy kokeilemaan itse, aikaa olisi saatu kulumaan reippaasti.

# 8. Kokeilun tekninen ympäristö

- Data haettiin ONNI-järjestelmästä xls-formaatissa
- Vastaukset anonymisointiin kehittäjän koneella python-skriptillä
- Data vietiin Azuren Machine Learning studioon
- Menetelmiä kehitettiin ML studiossa jupyter notebookeissa
  - Kokeiltiin mm. Tf-idf, fasttext, transformer-kielimallit, k-means klusterointi, voikko
- Käyttöliittymän prototyyppi toteutettiin jupyter notebookissa

# 9. Kokeilun data

- Data haettiin ONNI-järjestelmästä xls-formaatissa
  - riveillä oli vapaatekstiä vastauksena kysymykseen:  
”Mitä työssäni tarvittavaa osaamista kehitän?”
- Data anonymisoitiin kehittäjän koneella python-skriptillä.
- Datan formaatin ja puuttuvien kenttien ymmärtäminen vei jonkin verran aikaa.
- Vastaukset olivat monimuotoisia: osa oli hyvin spesifejä ja osa epämääräisiä.

# 10. Jatkopäätökset ja -ideat

- Kokeilu esitellään Stadin HR-joryssa ja mahdollisesti kaupungin laajennetussa HR-joryssa.
- Jatkoa aletaan suunnitella Stadin HR-joryn jälkeen.
- Jatkokokeilun rahoitus on mietittävä Stadin HR:ssä.
- Tuotantokelpoisen ratkaisun kehittämiseen tulee saada mukaan toimialojen laaja edustus.
- Goforen näkemys:
  - Prototyypityökalu vaikutti varsin lupaavalta, kokeilun pohjalta voisi lähteä rakentamaan tukityökalua vastausten käsittelyyn.
  - Tuotantokelpoisen työkalun käyttöliittymä olisi hyvä suunnitella siten, että se palvelee sekä käyttäjää että tekoälyn jatkuvaa oppimista.