

TRAFICOM

Liikenne- ja viestintävirasto

TraficomApuri – Generatiivisen tekoälyn pilotti

Esittely ALVN-ryhmälle
28.08.2025

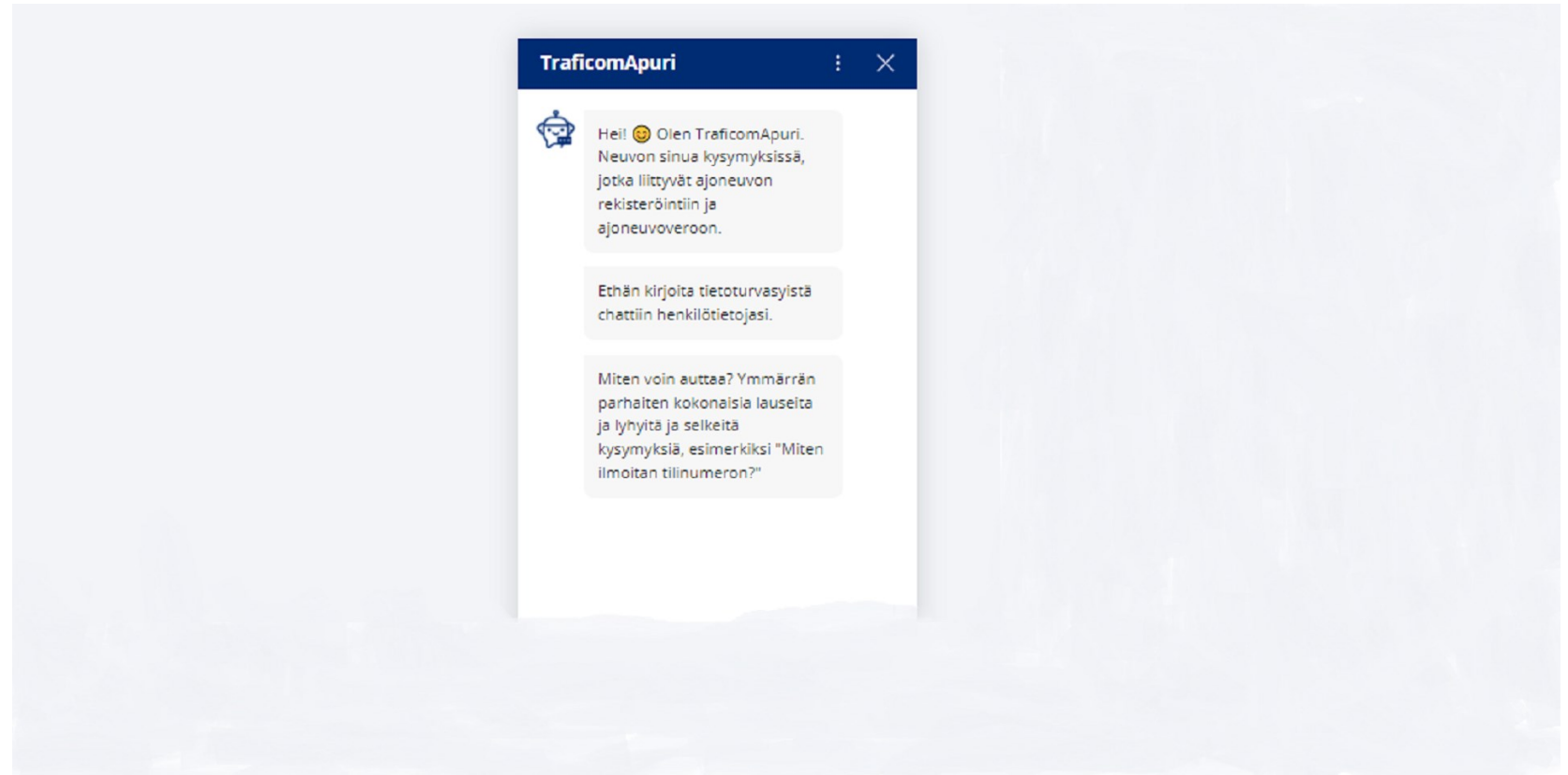


Sisällys

- ▶ Käytännön toteutus
- ▶ Pilotin fokusalueita ja keskeisiä tutkimuskysymyksiä
- ▶ Pilotin toteutus eri testausryhmissä
- ▶ Palautteet
- ▶ Haasteet ja riskit
- ▶ Yhteenveto

Käytännön toteutus

- ▶ Kyseessä oli käyttöliittymältään ns. perinteinen Chatbot.
 - ▶ Vastaukset luotiin kuitenkin Generatiivista tekoälyä käyttäen.
 - ▶ Lähteenä mm. Traficomin julkiset verkkosivut ja muu soveltuva aineisto
- ▶ Kokeilu oli rajattu Autoilijan palveluiden rekisteröinti- ja kuljettajantutkintoasioihin.
- ▶ Pilotti ei ollut käytössä asiakasrajapinnassa



Pilotin fokusalueita ja keskeisiä tutkimuskysymyksiä

1. Tarkoituksena oli testata, voitaisiinko LLM pohjainen tekoälybotti saada toimimaan niin, että se voitaisiin viedä asiakasrajapintaan, Traficom.fi:hin chatbotiksi

- ▶ Pystyykö tekoälybotti toimimaan autonomisesti, valvomattomasti? (ei esim. hallusinoi hallitsemattomasti, ei solvaa asiakasta, ei anna johdatella, vastaa luotettavasti ja oikein, pysyy annetuissa reunaehdoissa, esim. Traficom.fi – sivuston sisällöissä)
- ▶ Ovatko vastaukset käyttäjälle ymmärrettäviä ja selkeässä muodossa?
- ▶ Ovatko vastaukset luotettavia ja mitä tekoälybotti vastaa silloin, kun se ei osaa vastata kysymykseen?
- ▶ Ohjaako tekoäly asiakasta selkeästi eteenpäin asiointiprosessissa?
- ▶ Millaisia mahdollisia ongelmia ja rajoitteita tekoälyn käyttöön liittyy ja miten näihin voitaisiin vastata?

2. Tarkoituksena oli kokeilla helpommin rakennettavaa ja ylläpidettävää teknologista ratkaisua huomioimaan Traficomien laajojen palvelusisältöjen asettamat vaatimukset.

3. Pilotin tavoitteena oli kartoittaa generatiivisen tekoälyn toiminnallisuuden tietoturvaa sekä varmistaa, että järjestelmän luotettavuus ja käyttäjätietojen hallinta vastaavat Traficomien vaatimuksia – tämä toteutettiin omana kokonaisuutenaan.

Pilotin toteutus eri testausryhmissä



Traficomien ulkoistetun neuvontapalvelun (Barona) työntekijät testasivat TraficomApuria sillä ajatuksella, että se toimisi heidän työnsä tueksi tarkoitettuna tukiälynä.

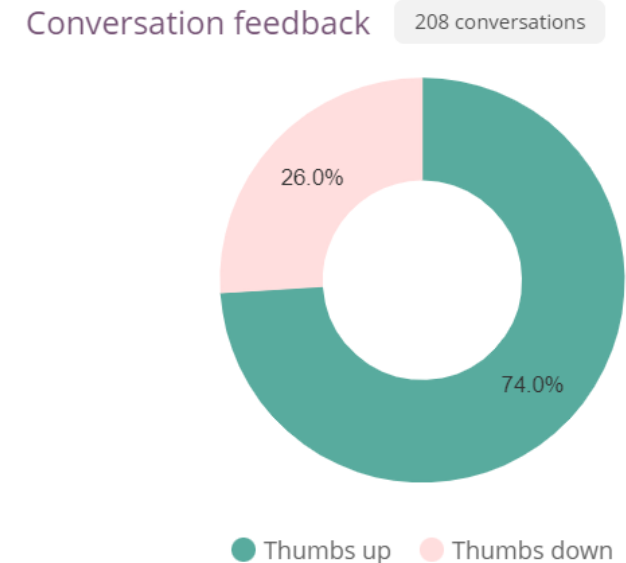
Traficomien työntekijöillä oli mahdollisuus ilmoittautua fokusryhmään, jonka osallistujat saivat pidemmän testausajan, tarkemman ohjeistuksen, ja joilta kerättiin laajempi palaute erillisen verkkokyselyn muodossa.

Kaikilla Traficomien työntekijöillä oli mahdollisuus testata TraficomApuria.

Lukuihin sisältyy myös fokusryhmän ja asiakaspalvelun keskustelut ja palautteet.

Palaute keskusteluista oli pääosin positiivista

- ▶ **74% palautteessa oli positiivista**” eli peukku ylös palautetta
 - ▶ Myös joissain ”alapeukutetuissa” keskusteluissa ei tarkemman analyysin pohjalta nähty merkittävää vikaa. Moni alapeukutettu keskustelu tuli *Out-of-Scope* asioista, eli Apurilta oli kysytty esim. asioita joihin sen ei kuulunutkaan vastata tai eri kielillä, mikä ei ollut mukana pilotissa
- ▶ Raportin palautepoimintojen perusteella positiivista palautetta annettiin esim. **vastausten selkeydestä ja ymmärrettävyydestä, ystävällisyydestä, ja siitä, että bottia ei yrityksistä huolimatta oltu saatu sekaisin**
- ▶ Myös TraficomApurin **suorituskyky oli analyysin perusteella hyvä**
 - ▶ TraficomApurin ratkaisuaaste oli 77 %. Ratkaisemattomien kysymysten määrään vaikutti suuresti testaajien halu harhauttaa chattibottia ja katsoa miten botti reagoi yllättäviin aihepiiriin ulkopuolisiin kysymyksiin



Haasteista ja riskeistä

- ▶ Vaikka yleispalaute myös luotettavuuden ja vastausten oikeellisuuden osalta oli hyvä, epävarmuus luotettavuuden suhteen sekä ajoittaiset väärät vastaukset nousivat esiin osassa vastauksia
- ▶ Myös mahdollisina haasteina tai riskeinä käyttöönotolle nähtiin erityisesti luotettavuuteen ja mahdollisiin vääriin vastauksiin liittyvät kysymykset ja vastuut

Väärät vastaukset voivat pahimmassa tapauksessa ohjata asiakkaan toimimaan vastoin lakia

Pitää olla erityisen varmaa että botti ei anna väärää informaatiota. Tästä löytyy hyvä esimerkki ulkomailta jossa botin luoma sisältö on tulkittu vastaavaksi kuin websivujen muu sisältö ja botin vastaus on ollut virheellinen, jolloin yritykselle on syntynyt korvausvastuu (Air Canada).

Asiakkaat voivat tarkoituksella tai vahingossa tulkita ohjeita väärin. Tulee laittaa disclaimer, että kyseessä ei ole viranomaispäätös tms (kunnes saadaan sellaisia tekevät botit hommiin ja lait ajan tasalle). Ja edellytetään/toivotaan myös asiakkailta, että mikäli vaikuttaa siltä, että tekoäly on tuottanut väärän tai sääntelystä poikkeavan vastauksen, tulisi tästä ilmoittaa ensi tilassa meille.

Yhteenveto



Yhteenveto

- ▶ Yleinen käyttökokemuspalaute kaikilta testiryhmiltä oli, että TraficomApuri oli helppokäyttöinen, nopea ja ystävällinen
- ▶ Apuri ei varsinaisesti hallusinoanut, eikä sitä erilaisista yrityksistä huolimatta saatu sekaisin tai toimimaan väärin
- ▶ Keskeisimmät haasteet liittyivät
 - ▶ Ajoittaisiin virheellisiin vastauksiin/ epäluotettavuuteen tai sanojen vääriin tulkintoihin.
 - ▶ Turhan usein asiakaspalveluun ohjaamiseen
 - ▶ Sekä vastausten ajoittaiseen liialliseen yleisluontoisuuteen. Apurin toivottiin kysyvän tarkennuksia, ohjaavan lukemaan lisää sekä tarjoavan paremmin lähdelinkkejä
- ▶ Käyttäjien kokemuksissa Apurin toiminnasta ja vastausten johdonmukaisuudesta oli myös vaihtelua
 - ▶ Joillekin Apuri oli esim. puhunut ruotsia, toisille ei.
 - ▶ Välillä Apuri näytti vastaavan hyvin ja välillä heikommin, vaikka kyse samasta aihepiiristä/kysymyksestä

Yhteenveto

- ▶ On myös hyvä huomioida, että käyttäjät ovat omalla AI-matkallaan hyvin erilaisissa vaiheissa. Tottuneet käyttäjät odottavat yhtä sujuvaa toimintaa, kuin mihin he ovat tottuneet esim. ChatGPT:n kanssa. Toisille käyttäjille generatiivinen tekoäly on vielä hyvin uutta eivätkä he esim. välttämättä ymmärrä toimivansa ihmismäisesti vastaavan botin kanssa.
- ▶ Käyttäjät toivoivat tekoälybotin jatkokehitystä ja oikeaa käyttöönottoa ja sillä nähtiin olevan asiakasneuvonnan tarvetta vähentävää vaikutusta
- ▶ Kokeilun tulokset olivat niin lupaavia, että kehitystyötä ja testaamista kannattaa jatkaa
 - ▶ Keskeiset kehitysteemat liittyvät erityisesti vastausten luotettavuuden sekä eteenpäin ohjaamisen edelleen kehittämiseen
 - ▶ Lisäksi kokeilun aikana korostui myös monikielisyyden tarve/vaatus

Pilotin vaatimukset ja onnistumiskriteerit täyttyivät

Opit generatiivisen tekoälyn toiminnallisuuksista

- ▶ Chattibotti toimi hyvin ja se oli toteutettavissa lyhyessä ajassa. Pidemmällä käyttöönottoprojektilla voidaan luoda helposti tuotantokelpoinen chattibotti asiakasrajapintaan.
- ▶ Generatiivinen malli vaatii huomattavasti enemmän testausta mitä strukturoitu malli.
- ▶ Mikäli tietolähteinä käytetään nettisivuja, ne on hyvä optimoida vastaamaan käyttöä tai pohtia muita ratkaisuja tietosisällölle

Tietoturvallinen toteutus

- ▶ Toteutus todettiin Traficomien tietosuoja ja -turva vaatimusten mukaiseksi sisäistä testausta varten
- ▶ GenAi sisällössä hyödynnettiin Traficomien tietolähteitä tietopohjana, mikä ukaan mahdollisti käyttöönoton ja ylläpidon tehokkaasti ja vaivattomasti.
- ▶ Chattibotin ratkaisukyvykkyys ja käyttäjäkokemus oli hyvällä tasolla.

Kiitos!