



VALTIOVARAINMINISTERIÖ

PuheKamu loppuraportti

Maahanmuuttovirasto

VM/2330/02.02.03.09/2018

Versio 1.0

27.8.2019

Sisällys

Sisällys	2
Dokumentin versiohistoria	2
1. Yhteenveto	3
2. Kokeilun toteutuminen	3
2.1. Kokeilun tiedot	3
2.2. Kokeilun rahoitus, kustannukset ja henkilötyöpäivät	4
2.3. Hankintakäytännöt	5
2.4. Riskienhallinta	5
2.5. Kokeilun tavoitellut hyödyt ja niiden toteutuminen	6
3. Kokeilun päättäminen	7
3.1. Kokeilun opit	7
3.2. Kokeilun kokemusten jakaminen	8
3.3. Kokeilun hyödyntäminen	8

Dokumentin versiohistoria

Versio	Päiväys	Laatija	Muutoksen kuvaus
0.1	24.7.2019	Susanne Miessner Arto Airaksinen	
0.2	31.7.2019	Vesa Hagström	Raportin läpikäynti ja täsmennykset
1.0	27.8.2019	Vesa Hagström	Talous yms. täsmennyksiä

1. Yhteenveto

Tämä dokumentti on uuden toimintamallin tai teknologiaratkaisun toiminnan todentamiseen tähtäävän PuheKamu -kokeilun loppuraportti.

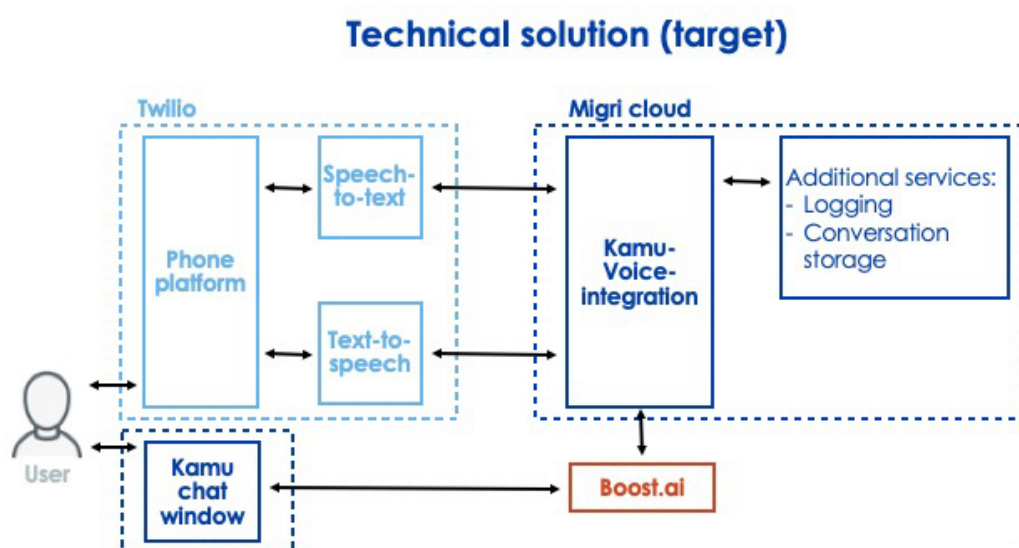
2. Kokeilun toteutuminen

2.1. Kokeilun tiedot

Kokeilussa yhdistettiin Migrin käytössä oleva keskusteleva tekoälyratkaisu (~Kamu chatbot) STT/TTS (Speech-to-text / Text-to-speech) teknologiaan opiaksemme puhemuotoisesta tekoälyratkaisusta. Kokeilun aluksi tehdyn prestudyn tuloksena kokeilussa hyödynnettäväksi teknologiaksi valittiin Twilio ja äänikanavaksi puhelinliittymä.

Twilio on ns. out of the box ratkaisu, eli käytimme kokeilussa valmista tuotetta äänen prosessointiin. Partnerimme Accenture rakensi liittymän nykyinen tekoälyratkaisumme Kamun (joka käyttää pohjanaan Boost.ai:n teknologiaa) ja Twilion välille. Twilion tekstiksi prosessoima puhe siis yhdistettiin nykyiseen tekoälyratkaisuun, ja Twilio vastaavasti konvertoi Kamun vastauksen puheeksi.

Kokeilun puitteissa käydyt keskustelut tallennetaan tekstimuodossa.



Kokeilussa oli tarkoitus oppia puheesta tekstiksi -teknologian kyvykkyyksistä ja maturiteettitasosta osana bottipohjaista asiakaspalvelua. Tarkoitus oli selvittää sosisiko puhekäyttöliittymä sellaisenaan samalle tekstipohjaiselle sisällölle.

Puhekäyttöliittymää testattiin maahanmuuttaja-asiakkaiden toimesta, jotta saatiin autenttinen kuva sekä puhekäyttöliittymän kyvykkyyksistä että kyseisen palvelun valmiustasosta tuotantokäyttöön. Asiakastestaaajia oli yhteensä 9 joista 8 englannin kielellä ja yksi suomen kielellä. Suomen kielen osalta myös 7 Maahanmuuttoviraston henkilöä testasi puhekäyttöliittymää.

- Kokeilun rajallisuuden vuoksi oletusarvoksi englannin kielen osalta asetettiin amerikanenglanti eikä käytetty eri kielikonfiguraatioita vastaamaan henkilön oikeaa alkuperää. Tällä tavoin tarkoitus oli saada kokemusta standardoidusta kielimallista ja siitä, tarvitaanko kielimallien kustomointia asiakkaiden alkuperän mukaisesti.
- Asiakastestauksista saadut kokemukset ovat teknisen toteutuksen kokeilun lisäksi tärkeimpiä kokeilusta saatuja tuloksia.

2.2. Kokeilun rahoitus, kustannukset ja henkilötyöpäivät

Kokeilun suunnitellut (käyttö- ja kirjausoikeuspäätöksen mukaiset) ja toteutuneet kustannukset euroina ovat eriteltyinä omaan ja ostettuun työhön sekä muihin kustannuksiin seuraavat:

Kustannus	Suunniteltu €	Toteutunut €
Oma työ (nykyresursseilla tehtävä työ)	30.000	3.817,85
Oma työ (kokeiluun erikseen palkattavien resursien työ)	0	0
Palvelujen ostot	43.800	24.470,40
Muut kustannukset	0	0
Kokonaiskustannus	73.800	28.288,25

Kustannukset eriteltyinä rahoituslähteittäin euroina ovat seuraavat:

Rahoituslähde	Suunniteltu €	Toteutunut €
28.70.22 Hallinnon palveluiden digitalisoinnin tuki	50.000	28.288,25
<viraston oma momentti>	23.800	0
Kokonaiskustannus	73.800	28.288,25

Kokeilun kustannusten yalitus johtuu pääosin seuraavista tekijöistä:

- Puhealustan integrointi nykyiseen chatbot tekoalustaan oli teknisesti helpompi operaatio kuin arvioimme. Koska kokeilua ei sellaisenaan voi ottaa tuotantokäyttöön ei tarvittu myöskään tuotantokäyttöön tähtääviä teknisiä tehtäviä. Työtä myös tarvittiin vähemmän, koska tuotantokäyttöön tarvittava työ jäi tekemättä.
- Oma työtä tarvittiin merkittävästi vähemmän, koska kokeilua ei sellaisenaan voi ottaa tuotantokäyttöön. Mikäli olisimme kokeilussa päätyneet tuotantokäyttömahdollisuuteen, olisi omaa työtä tarvittu juuri näihin tehtäviin, palvelun sisällön muokkaamiseen ja käyttöönottehtäviin merkittävästi enemmän.

- Puhekomponentin (Twilio) osalta käytettiin testiversiota.

Oman, kokeiluun erikseen palkatun henkilöstön toteutunut kustannus euroina ja henkilötyöpäivinä:

€	htp
0	0

2.3. Hankintakäytännöt

Kokeilussa hankittiin suoramarkintana seuraavat palvelut arvioituna hankinnan kustannuksina:

- Accenture Oy: 38.800 euroa ALV 0%
- Twilio Oy: 5.000 euroa ALV 0%

Hankinnat toteutettiin suoramarkintoina alittaen kansallisen kynnyksarvon.

Accenturen rooli oli integroida puhelinkäyttöliittymä ja puhekomponentti osaksi olemassa olevaa chatbot tekoälyalustaa. Twilio tarjosi Speech-To-Text ja Text-To-Speech palvelut sekä puhelualustan transaktiopohjaisena.

Toteutuneet kustannukset olivat pienempiä kuin suunnittelu- ja hankintavaiheessa arvioidut (ks. kohta 2.2. ”kustannusten alitus”).

2.4. Riskienhallinta

Kokeilussa ei nähty merkittäviä riskejä olemassa olevaan chatbot -palveluun liittyen, sillä kokeilu toteutettiin tuotantokannasta eriytettyssä ympäristössä.

Osa riskeinä koettavista asioista on keskeinen osa itse kokeilua. Näitä ei ole käsitelty riskeinä. Myöskään se, päättykö kokeilu lopulta tuotantoon, ei ole riski, vaan oleellinen osa kokeilua.

Kokeilun riskien tilanne kokeilun päättyessä:

Riski	Lopullinen tila	Toimenpiteet	Toimenpiteiden vaikutus
Riittämätön suomen kielen asiakastestaus.	Ei toteutunut	Käytettiin Migrin henkilöstöä täydentämään suomen kielen testausta.	Poisti riskin
Twilio Speech-To-Text (STT) kyvykkyys ja maturiteetti.	N/A	Arvioimme tulisiko kokeilla toista STT komponenttia, mutta valittiin tuote/palvelu käyttää samoja edistyneimpiä kirjastoja	Osa kokeilussa selvitettyjä asioita eli ei siinä mielessä riski.

2.5. Kokeilun tavoitellut hyödyt ja niiden toteutuminen

Kuvaa alla olevaan taulukkoon kehitettävän prosessin vaikuttavuus- ja asiakashyötypotentiaali hakemuksen mukaan ja arvioi sen toteutumista kokeilun jälkeen:

Arvio kehitettävän prosessin vaikuttavuus- ja asiakashyötypotentiaalista		
Tavoiteltava yhteiskunnallinen vaikuttavuus	Hyötyjen realisoituminen hakemuksen mukaan (keskeiset toimenpiteet ja vastuutahot vaikuttavuushyötyjen realisoinemiseksi)	Arvio hyötyjen realisoitumisen toteutumisesta, jos kokeilussa rakennettu muutos otetaan tuotantoon
Puhemuotoisen käyttöliittymän tuominen perinteisen lomake- ja tekstimuotoisen käyttöliittymän rinnalle yleisenä isona teknologiatrendinä (tukee Aurora.AI visioita)	Migrin hyötyjen osalta puhekomponentti liittyy laajempaan kokonaisuuteen kuin pelkkään asiakaspalvelutapahtumaan. Keskeinen toimenpide on se, että odotetaan vielä teknologian, puheen kielimallien ja NLP maturiteettitason nostoa sekä toisaalta kuvataan puhemuotoiseen käyttöliittymään vaadittavat tekniset ja sisällölliset kehityskohteet osana kokonaiskonseptia.	On huomattavasti Migriä laajempi kokonaisuus. Miten tulevaisuudessa kokonaisuudessaan älykkäitä palveluja käytetään puheella. Hyötyjä tästä "early-adopter" lähestymisestä muodostuu kaikille organisaatioille, jotka voivat hyödyntää kokeilun tuloksia osana omaa palvelukehitystä (=oikopolkuja etenemiseen).
Asiakashyötypotentiaali: yhdenvertaisuus, saavutettavuus - luku- ja kirjoitustaidottomat asiakassegmentit. Mahdollisuus myös puheella laajentaa kielivalikoimaa virallisista palvelukielistä.	Keskeinen toimenpide on se, että odotetaan vielä teknologian, puheen kielimallien ja NLP maturiteettitason nostoa sekä toisaalta kuvataan puhemuotoiseen käyttöliittymään vaadittavat tekniset ja sisällölliset kehityskohteet osana kokonaiskonseptia.	Asiakaspalvelun ja asiakaskokemuksen parantuminen.

Kuvaa alla olevaan taulukkoon kehitettävän prosessin vaikuttavuus- ja asiakashyötypotentiaali hakemuksen mukaan ja arvioi sen toteutumista kokeilun jälkeen:

Arvio kehitettävän prosessin tuottavuuspotentiaalista		
Taloudelliset hyödyt	Hyötyjen realisoituminen (keskeiset toimenpiteet ja vastuutahot vaikuttavuushyötyjen realisoinemiseksi)	Arvio hyötyjen realisoitumisen toteutumisesta, jos kokeilussa rakennettu muutos otetaan tuotantoon
Voidaan tarjota asiakaspalvelua robotisoidusta myös puhe- ja kirjoitustaidottomille asiakassegmenteille. Optiona laajentaa myös kielivalikoimaa puhepalvelussa.	Keskeinen toimenpide on se, että odotetaan vielä teknologian, puheen kielimallien ja NLP maturiteettitason nostoa sekä toisaalta kuvataan puhemuotoiseen käyttöliittymään vaadittavat tekniset ja sisällölliset kehityskohteet osana kokonaiskonseptia. Puhemuotoisen palvelun käyttöönotto vaatii myös asiakaspalvelutoiminnan kehittämistä, viestintää, koulutusta sekä tukipalvelun huomioimista.	Asiakaspalvelun tehostumisen neuvontatyyppisissä palveluissa puheella on arviolta 1-2 htv. Jatkovaiheen puheella asionnin osalta arvioita on vielä mahdoton antaa. Samoin on puheen soveltamisessa muihin puhetilanteisiin (haastattelut, puhuttelut).

	Kielivalikoiman laajentaminen vaatii lisäksi muita puheeseen liittyviä palveluja (mm. auto-maattinen kääntäminen).	
--	--	--

3. Kokeilun päättäminen

3.1. Kokeilun opit

Kokeilusta ja kokeilun opeista on olemassa laajemmat power point -materiaalit. Ne on tallennettu hankeraportointiin ja ovat jaettavissa oppeina ja kokemuksina muille organisaatioille. Alla tiivistettynä keskeisimpiä oppeja.

Teknologianäkökulma:

- Valmiin tuotteen ja palvelun implementointi (Twilio) sekä käyttöönotto onnistui nopeammin kuin etukäteen arviointiin. Tekninen implementaatio on lisäksi tuotantokäyttökelpoinen. Teknologian käyttöönotto itsessään eli ole kokeilun onnistumisen kulmakivi eikä este tulevaisuuden edistyneimmille ja uusille kokeiluille.

Puheteknologian maturiteetti:

- Text-To-Speech (TTS) eli tekstistä puheeksi muunnokset ja maturiteetti: Puheteknologia on näiltä osin riittävän hyvä asiakaspalvelun tuotantokäyttöön. Suomen kielen osalta on pieniä haasteita esimerkiksi englanninkielisten sanojen kanssa (sanasta on vain englanninkielinen käytäntöön vakiintunut versio) sekä esimerkiksi numeroiden taivutusten kanssa. Englannin kielen osalta tekstistä puheeksi osuudessa ei havaittu lainkaan ongelmia tai haasteita.
- Speech-To-Text (STT) eli puheesta tekstiksi muunnokset ja maturiteetti: Tällä osa-alueella havittiin ongelmia ja haasteita, jotka vielä hidastavat ja estävät puhekäyttöliittymän tuotantokäyttöä. Puheesta tekstiksi muunnokset eivät ole riittävän luotettavia ja lisäksi on paljon vaihtelua eri käyttäjien (puheen) osalta. Yllättävä havainto oli, että suomen kieli toimii paremmin kuin englannin kieli, mikä tosin voi johtua myös testihenkilöryhmistä.

Jotta puhekäyttöliittymä voitaisiin ottaa tuotantokäyttöön, seuraavilla osa-alueilla on tehtävä jatkokehitystä ja -testausta:

Kehitystyö:

- Teknologiaan liittyvä kehitystyö: Vaikka teknologia itsessään toimii riittävän luotettavasti, tarvitsee tuotepohjainen ratkaisu jotain kustomointia sekä satunnaisien ongelmatilanteiden selvittelyä. Tällaisia osa-alueita ovat mm. satunnaiset hiljaiset (vastaamattomat) tilanteet, puheen nopeus, erityiset puhekäyttöliittymän tarvitsemat äänikomennot ("Lopeta", "Toista", "Hidasta", ...)
- Sisältöön liittyvä kehitystyö: Tekstimuotoiseen chatbot -ratkaisuun tuotettu sisältö ei kaikilta osin sovellu sellaisenaan puhemuotoiseen ratkaisuun.

Selkeimpiä näistä ovat tekstimuotoisen palvelun ”Klikkaa” tai ”Lue lisää” toiminnot. Lisäksi annettujen vastausten sisältöä ja kieliasua on muutettava joiltain osin paremmin puhemuotoon soveltuvaksi (esim. liian pitkät tekstimuotoiset vastaukset).

Kattavampi ja laajempi käyttäjätutkimus sekä siihen liittyvä suunnittelu / konseptointi:

- Laajempi käyttäjätestaus parametroiduilla ominaisuuksilla (ikä, sukupuoli, äidinkieli, murre / alue)
- Erilaisen synteettisen puheen vaikutus käyttäjäkokemukseen ja luottamukseen.
- Robotisoidun puhekäyttöliittymän ja puhelinpalvelua suorittavien asiantuntijoiden vertailu.

Muuta:

- Juridiset, organisatoriset ja vastuukysymykset annettaessa asiakasneuvontaa. Nykyisellä puhekäyttöliittymän maturiteetilla on parannettunakin odotettavissa palvelutilanteisiin liittyviä odottamattomia tilanteita.

3.2. Kokeilun kokemusten jakaminen

Migri osallistuu Aurora.Ai yhteistyöhön eri tasoilla ja jaamme tulokset Aurora.Ai:n laajan verkoston kautta. Migri digikehitykseen kohdistuu paljon mielenkiintoa kansallisesti ja kansainvälisesti ja jaamme tämänkin kokeilun tuloksia aktiivisesti eri yhteistyötapahtumissa. Lisäksi hyödynnämme toimivaa multibot yhteistyötä (”Starting up Smoothly”) Verohallinnon ja PRH:n kanssa ja siihen työhön liittyy paljon kiinnostusta.

Kokeilun tuloksia ja oppeja jaetaan myös soveltuvien osien Vaken suomenkielisen tekoälyn kehittämisohjelmassa.

Palvelumuotoilutomintomme ”Inland” on jo jakanut ja jakaa kokeilun kokemuksia blogikirjoituksissaan. Esimerkki:

<https://medium.com/inland/voice-of-kamu-experimenting-with-chatbot-voice-technology-at-migri-122d17110a53>

Tulosten jakamisesta vastaavat kokeilussa mukana olleet asiantuntijat ja vastuhenkilöt sekä palvelumuotoilutoiminto ”Inland”.

Kokeilusta ja kokeilun opeista on olemassa laajemmat power point -materiaalit, joita voidaan toimittaa niitä tarvitseville organisaatioille.

3.3. Kokeilun hyödyntäminen

Kokeilun tuloksia ei sellaisenaan voi ottaa tuotantokäyttöön lähinnä Voice-To-Text maturitettitason haasteista johtuen. Kokeilun tulokset jätetään kuitenkin testikäyttöön ja niitä voidaan hyödyntää sekä esittely- ja kokeilutilanteissa että saumattomasti mahdollisissa erikseen päätettävissä jatkovaiheissa.