

Chatbotit ja niiden koulutusdata

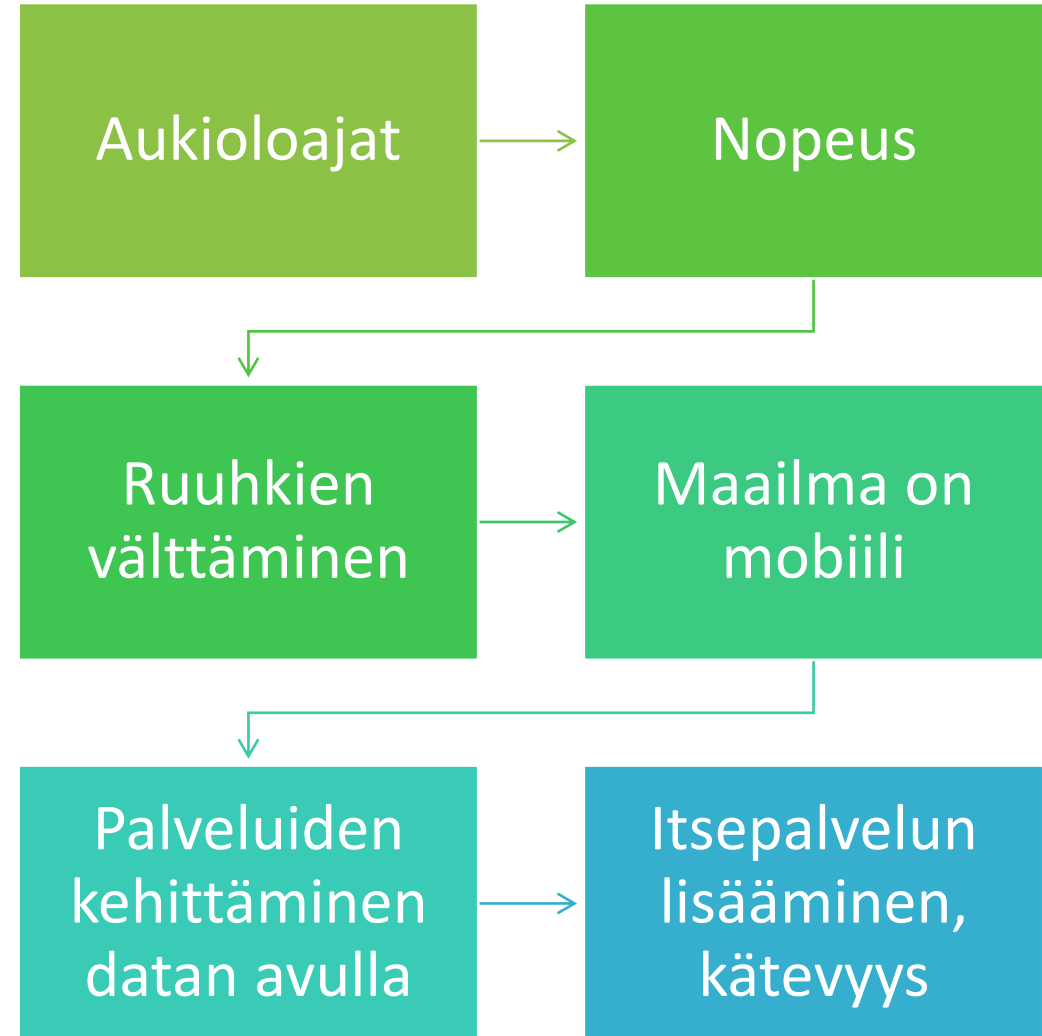
Helsinki

Laura Kerola
Erityissuunnittelija
Tekoäly ja optimointi /
Kaupunginkanslia

Chatbotit - Keskusteleva tekoäly

- Chatbotit ovat vakiintunut osa nykyaikaisia asiakaspalveluita
- Chatbotteja on eri tasoisia ja iso osa on näiden hybridejä:
 - **Nappibotit**
 - Nopeita ja edullisia ottaa käyttöön, rajoittuneita, kaikki ennalta määritettyä, ei varsinaista tekoälyä
 - **Avoimen dialogin eli keskustelevat botit**
 - Kysytään vapaamuotoisesti, mutta vaativampia ottaa käyttöön ja opetusdata merkittävässä roolissa
 - **Toiminnalliset avustajat**
 - Sisältävät itsepalvelutoimintoja esim. mahdollisuus tilata matkalippu botin kautta
 - **Generatiivista tekoälyä hyödyntävät**
 - Suurten kielimallien pohjalta keskustelevat botit jotka voivat vastaamisen lisäksi luoda kuvia, ohjelmakoodia, jne.

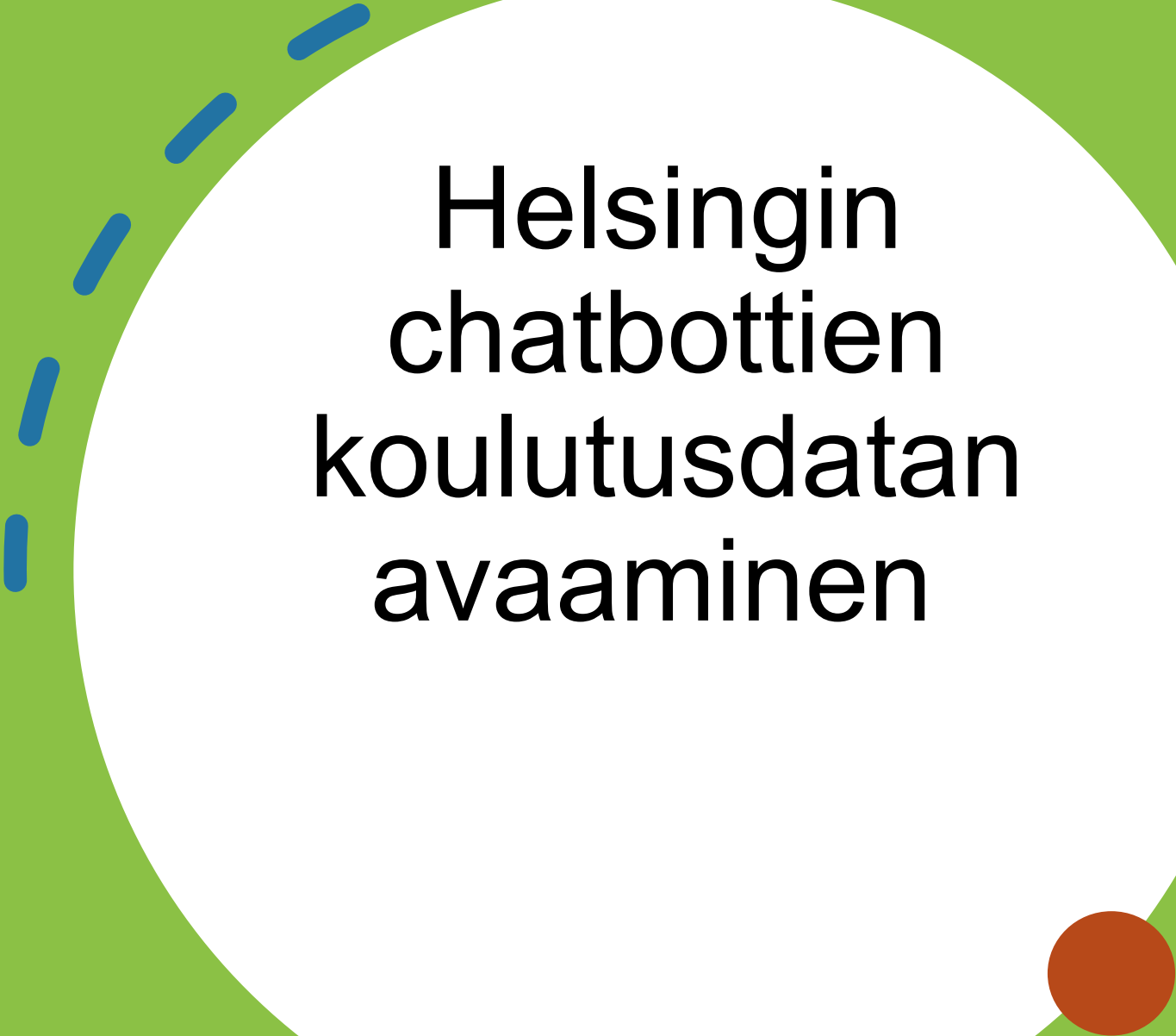
Miksi botteja ylipäättään käytetään?





Koulutusdata

- Hyvän ja hyödyllisen chatbotin kehittäminen vaatii paljon panostusta koulutusdataan
- Bottikuiskaajilla on tärkeä rooli
- Paras koulutusdata saadaan vasta julkaisun jälkeen

A decorative blue dashed line curves along the top-left edge of the white circle, and a solid orange circle is positioned at the bottom-right edge of the white circle.

Helsingin chatbottien koulutusdatan avaaminen

Watson chatbotit kaupungilla

Ulkoilubotti Urho

Pysäköintibotti

International
house of
Helsingin
Chatbot Into

Vuokra-asunnon
haun botti

Sotebotti Hester

Kesäsetelibotti

Taloushallinta-
Palveluiden Talbotti

IT-tuen Helbotti*

*eläköitynyt

Koulutusdatan avaamisen hyödyt

Läpinäkyvyys (Tekoälyn eettiset periaatteet)

Tarjotaan valmis pohja hyödynnettäväksi

Inspiraatiota ja ideoita muille tekijöille



Chatbottien koulutusdatat HRI:ssä

Ensimmäinen datasetti avattiin
vuonna 2022 (Neuvola-robotti
NeRo)

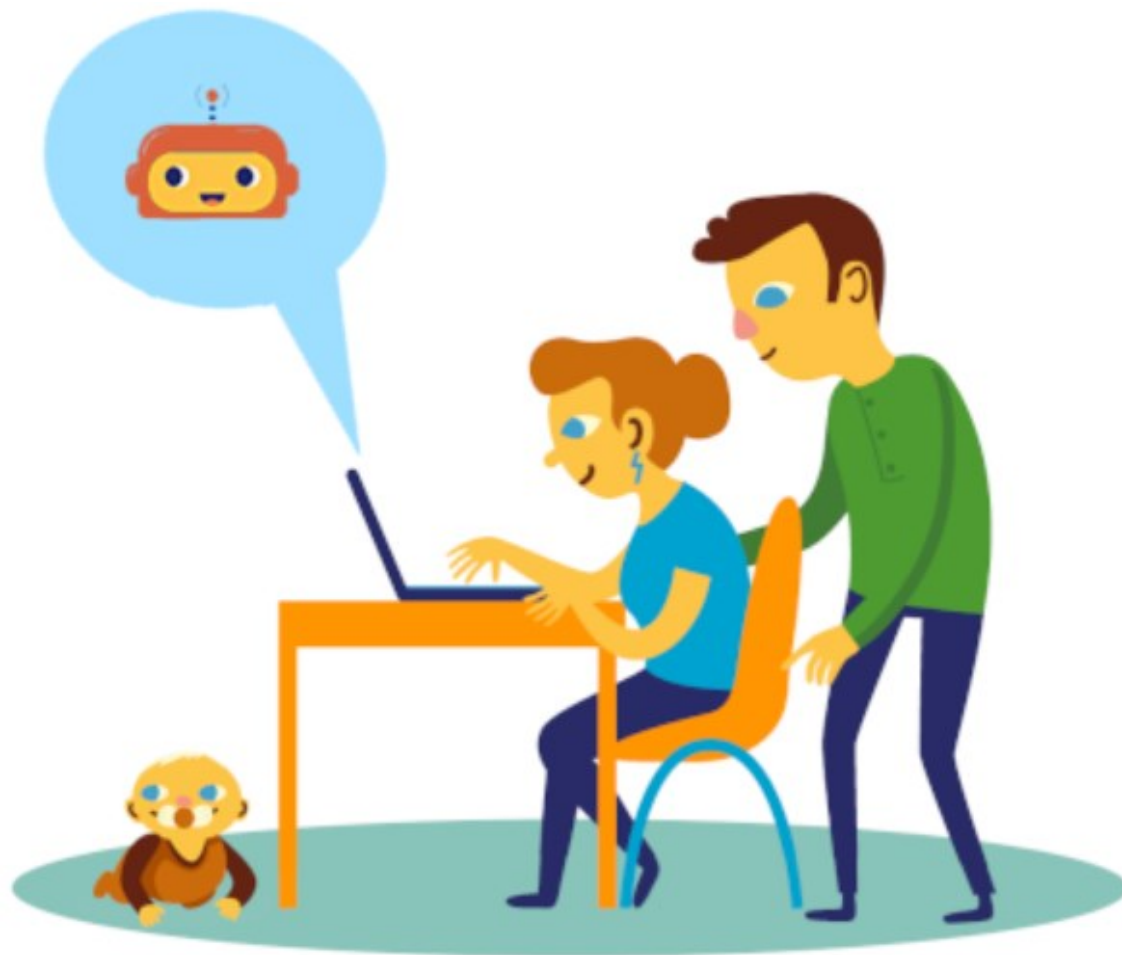
Valittiin vuoden 2022
parhaaksi data-avaukseksi

Rohkaisi jatkamaan datan
avaamista myös muiden
chatbotien osalta

Chatbottien koulutusdata HRI:ssä

Jokaisen chatbotin opetusdatasta on julkaistu .xlsx-muodossa seuraavat osat:

- **Vastaukset:**
 - jokaisen vastauksen yksilöivä ID ja itse vastausteksti
- **Intentit:**
 - Intentin yksilöivä ID ja esimerkkikysymykset tunnistamista varten
- **Entiteetit:**
 - Yksilöivä ID ja variaatiot/synonyymit/taivutusmuodot



Datan avaamisessa huomioitavaa

Järjestelmäkohtaisuudet
(esim.
toimintotunnisteet)

Dialogipuu ei sisälly
dataan

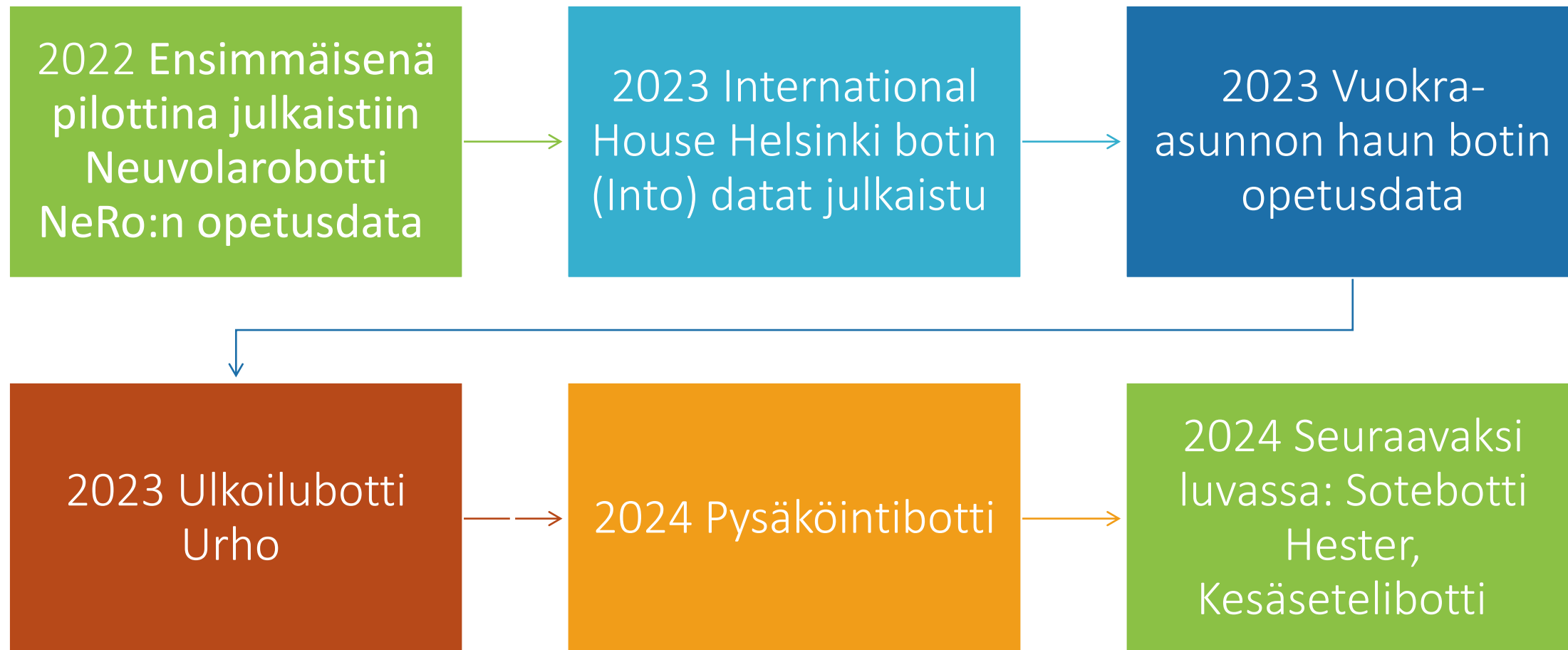
Bottikuiskaajat mukaan!

Kieliversiot ja niiden
muotoilu

Avatun datan
päivittäminen?

Tietosuoja!

Mitä tehty ja mitä luvassa



Mistä chatbottien koulutusdatat löytää?

- Hakemalla HRI:n etusivun haulla ”chatbot”
- Excel-muodossa, jokaisesta tiedostosta löytyy eri välilehdiltä intentit, entiteetit ja vastaukset
- Linkki chatissa
- [Helsingin kaupungin chatbotien koulutusdatat - Helsinki Region Infoshare \(hri.fi\)](https://hri.fi)

Data ja resurssit



Vuokra-asunnon haun chatbotin koulutusdata



Ulkoilubotti Urhon koulutusdatat



NeRo neuvolarobotin koulutusdatat



Chatbot Into koulutusdatat



Pysäköintibotin koulutusdata



KIITOS!

Laura.kerola@hel.fi

Erityissuunnittelija

Janne.kantsila@hel.fi

Johtava asiantuntija