

Chatbottien taustaa ja tulevaisuudennäkymiä

Sakari Koivunen 11.5.2022

keskustelu²

TAIVUTUS ▲

yks. nom.	keskustelu
yks. gen.	keskustelun
yks. part.	keskustelua
yks. ill.	keskusteluun
mon. nom.	keskustelut
mon. gen.	keskustelujen, keskusteluiden, keskusteluitten
mon. part.	keskusteluja, keskusteluita
mon. ill.	keskusteluihin

keskusteleminen, mielipiteiden vaihto, neuvottelu, pohdinta, väittely.

Vilkas pöytäkeskustelu.

Päästä keskustelun alkuun.

On virinnyt, sukeutunut laaja keskustelu [= julkinen keskustelu, debatti, polemiikki] verouudistuksesta.

Alustavat keskustelut [= neuvottelut, kaavailut] hallituksen muodostajasta.

Budjetin lähetekeskustelu lähetekäsittely.

Welcome to

EEEEEE	LL	IIII	ZZZZZZ	AAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LL	II	ZZZ	AAAAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LLLLLL	IIII	ZZZZZZ	AA AA

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.

This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU:

NLP

- NLP, *natural language processing*
- Luonnollinen kieli? Millainen kieli olisi ei-luonnollista tai luonnotonta?

Ihminen vs kone

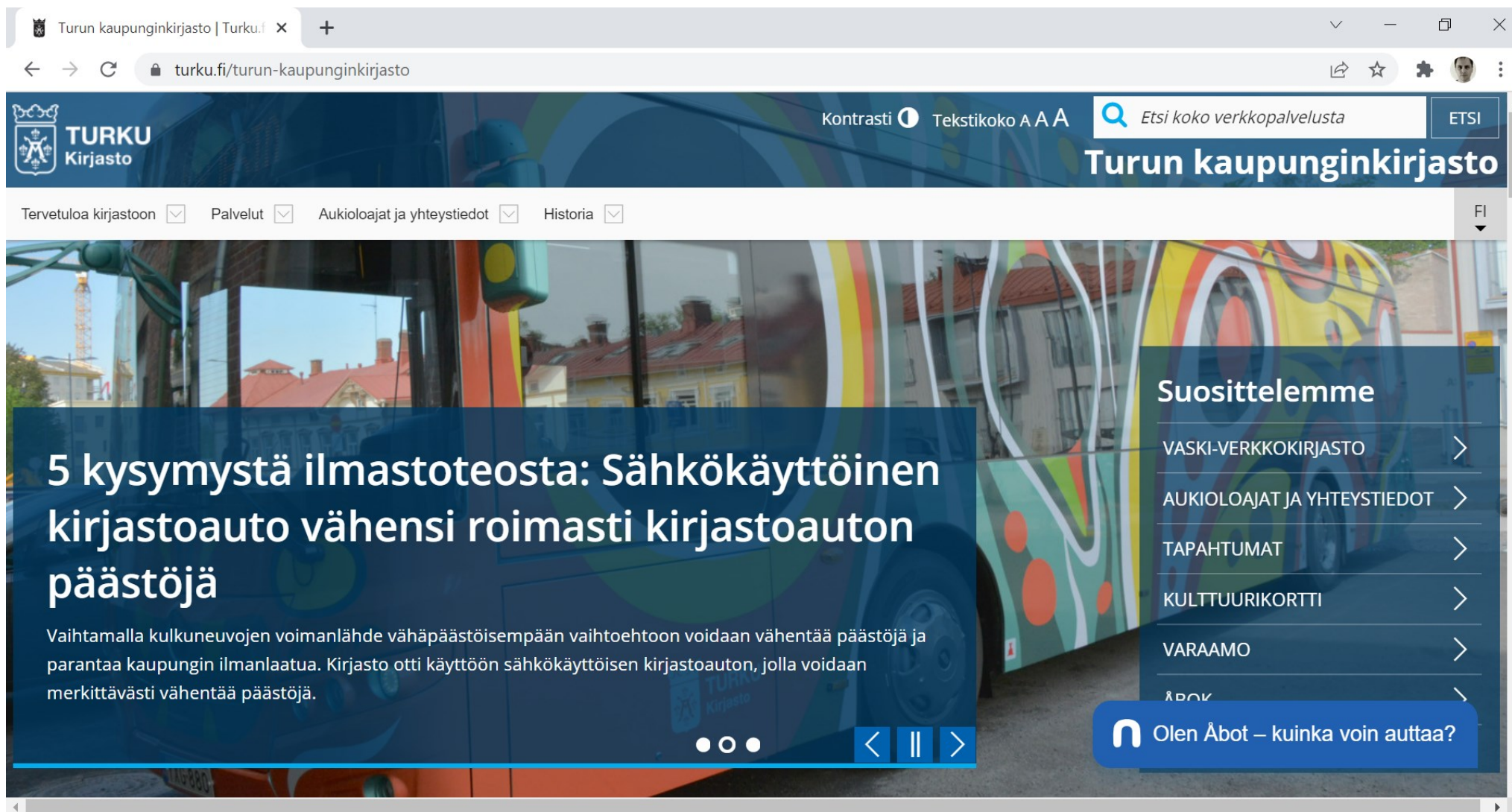
- Kysymys: *“Voiko jossain Turun kirjastossa käyttää 3D-tulostinta?”*
- Vastaus: *“Pääkirjastossa ainakin onnistuu”*

Ihminen vs kone

- Kysymys: <https://api.kirjastot.fi/v3/library?city.name=turku&service.name=3d-tulostimet&lang=fi>
- Vastaus: {

```
"total": 1,  
"type": "organisation",  
"items": [  
  {  
    "parent": 85661,  
    "web_library": null,  
    "address": {  
      "area": null,  
      "zipcode": "20100",  
      "box_number": null,  
      "city": "Turku",  
      "street": "Linnankatu 2",  
      "coordinates": {  
        "lon": "22.27086440",  
        "lat": "60.45046340"  
      },  
      "info": null
```

Kompromissi: ihminen + kone



The screenshot shows the homepage of the Turku City Library (Turun kaupunginkirjasto). The header includes the library's logo, navigation links for 'Tervetuloa kirjastoon', 'Palvelut', 'Aukioloajat ja yhteystiedot', and 'Historia', a search bar with the text 'Etsi koko verkkopalvelusta', and a language selector set to 'FI'. The main content area features a large banner image of a colorful, modern library bus. Overlaid on the left side of the banner is a text box with the headline '5 kysymystä ilmastoteosta: Sähkökäyttöinen kirjastoauto vähensi roimasti kirjastoauton päästöjä' and a subtext explaining that replacing fossil fuel vehicles with electric ones can reduce emissions and improve air quality. On the right side of the banner is a 'Suosittelemme' (We recommend) section with a list of links: 'VASKI-VERKKOKIRJASTO', 'AUKIOLOAJAT JA YHTEYSTIEDOT', 'TAPAHTUMAT', 'KULTTUURIKORTTI', 'VARAAMO', and 'ÄPÖK'. At the bottom right, there is a blue button with the text 'Olen Äbot – kuinka voin auttaa?'.

Turun kaupunginkirjasto | Turku.fi

turku.fi/turun-kaupunginkirjasto

Kontrasti Tekstikoko A A A

Etsi koko verkkopalvelusta ETSI

TURKU Kirjasto

Tervetuloa kirjastoon Palvelut Aukioloajat ja yhteystiedot Historia FI

5 kysymystä ilmastoteosta: Sähkökäyttöinen kirjastoauto vähensi roimasti kirjastoauton päästöjä

Vaihtamalla kulkuneuvojen voimanlähde vähäpäästöisempään vaihtoehtoon voidaan vähentää päästöjä ja parantaa kaupungin ilmanlaatua. Kirjasto otti käyttöön sähkökäyttöisen kirjastoauton, jolla voidaan merkittävästi vähentää päästöjä.

Suosittellemme

- VASKI-VERKKOKIRJASTO >
- AUKIOLOAJAT JA YHTEYSTIEDOT >
- TAPAHTUMAT >
- KULTTUURIKORTTI >
- VARAAMO >
- ÄPÖK >

Olen Äbot – kuinka voin auttaa?

Kone tykkää rakenteista

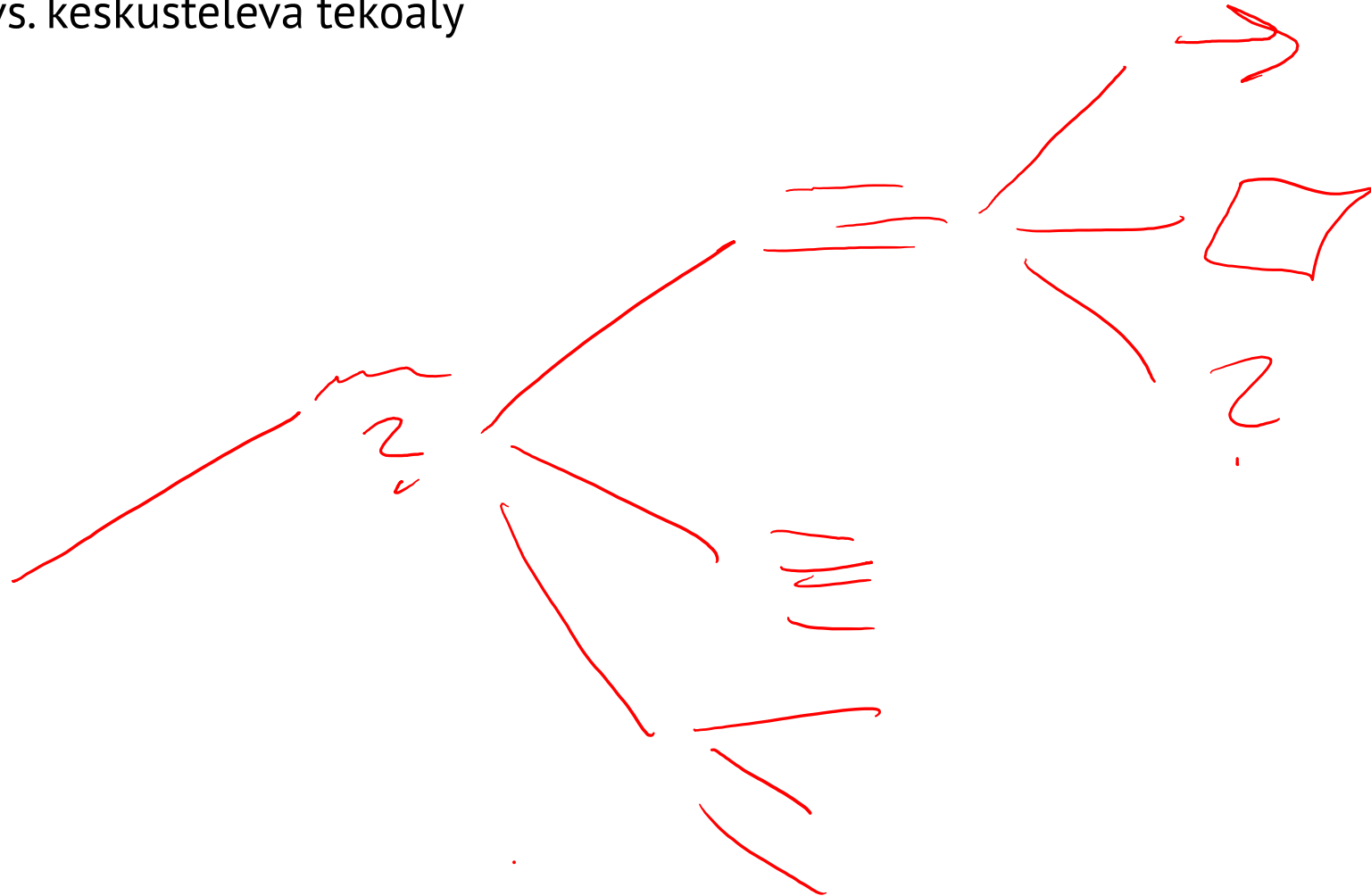
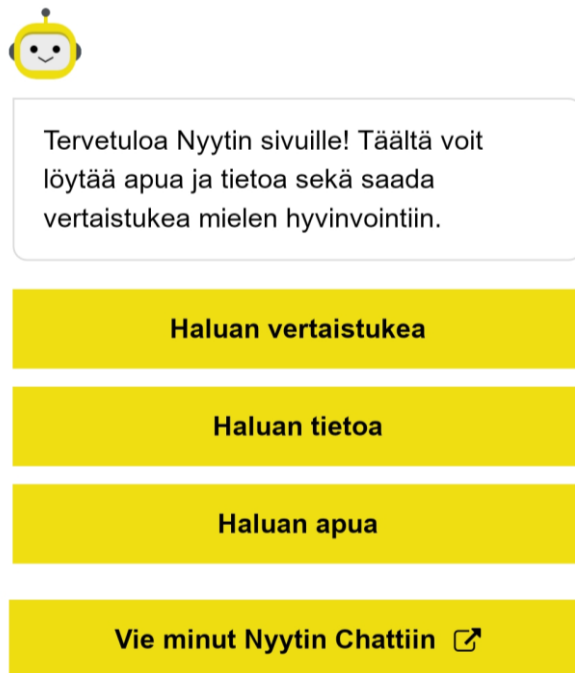
- Rakenteellinen (*structured*) vs. rakenteeton (*unstructured*) tieto
 - Ihmisen viestintä on rakenteetonta, vaikka siinä sinänsä onkin tietty kielioppi ja rakenne
 - Rakenteellinen data on koneelle helppoa käsitellä, rakenteeton ei
 - Miten rakenteeton data saadaan rakenteiseksi niin, että kone saa siitä tolkkua?
- NLP voidaan jakaa eri toimintoihin
 - NLU, *natural language understanding*
Miten tunnistetaan kysymyksestä “Voiko jossain Turun kirjastossa käyttää 3D-tulostinta?”
avainsanat ja merkitys niin, että voidaan muodostaa koneelle kysely
<https://api.kirjastot.fi/v3/library?city.name=turku&service.name=3d-tulostimet&lang=fi>
 - NLG, *natural language generation*
Miten koneen koodimaisesta vastauksesta muodostetaan ihmiselle helpommin ymmärrettävä suomenkielinen vastaus?
- (muitakin toimintoja toki on, mutta tyydytään nyt näihin kahteen)

Puhuttu vs. kirjoitettu kieli

- Yleisesti: kirjoitettu kieli on koneelle puhetta helpompaa
- Ihmiseltä koneelle:
 - Puhuttu kieli pitää ensin muuttaa tekstimuotoon koneen käsiteltäväksi
 - Puheen tunnistaminen on haastava tehtävä, vaikka sekin nykyään sujuu koneilta melko hyvin
- Koneelta ihmiselle:
 - Koneen on (melko) helppo tuottaa kirjoitettua tekstiä
 - Puheen tuottaminen on merkittävästi hankalampaa, mutta nykyään useimmilla kielillä koneen tekemä puhesynteesi alkaa olla jo ihan mukavaa kuunneltavaa
- Ihmisen ja koneen keskustelu puhuttuna on merkittävästi visaisempi pulma kuin sama keskustelu tekstimuodossa
- Chatbotit toimivat lähes aina kirjoitetun kielen avulla

Chatbotit

- Sääntöpohjainen klikkailubotti vs. keskusteleva tekoäly (esimerkkikuvat: nyyti.fi)



Sääntöpohjainen klikkailubotti

- Hyvä rajatuissa tilanteissa
- Teknisesti helppo rakentaa, käteviä työkaluja tarjolla
- Ohjaa käyttäjää melkein luontevasti, mutta keskustelu ei kuitenkaan ole keskustelua vaan klikkailua
- Nettisivuille sovitettu versio puhelinpalveluista tutusta valikkoklikkailusta: *paina 1 jos asiasi koskee huonevarausta, paina 2 jos...*
- Joustavuutta ja “älyä” voi rakentaa lisää esimerkiksi alussa kuvattujen rajapintakyselyiden kautta
- Mahdollista tarjota myös vaihtoehto jutella ihmisen kanssa: joku bottikeskustelun haara voi päättyä ihmischattiin, ja tässä vaiheessa botti on jo karsinut “väärät” ihmiset pois. Hälyn määrä ihmiselle tulevissa kontakteissa on näin pienempi.
- Analytiikka siitä, mitä valintoja käyttäjät ovat tehneet/mitkä keskusteluhaarat näyttävät olevan suosittuja ja missä kohtaa keskustelu useimmiten katkeaa

Keskusteleva tekoäly

- Vapaaseen tekstiin pohjautuva vuorovaikutus, oikea chat koneen kanssa
- Ihmiselle luontainen toimintatapa on jäsentymätön epäselvä säheltäminen, ei loogisesti vuokaaviossa kulkeva käsikirjoitus. Siksi nappibotit tuntuvat usein jäykiltä, ja samasta syystä oikeasti luontevan botin tulisi olla aidosti keskusteleva ja tilannetajuinen.
- Ideaalitilanteessa osaa vastata eri kysymyksiin, myös tarkentavilla vastakysymyksillä ja jatkaa keskustelua luontevasti tilanteiden välillä (Mitä tekoälyä käsitteleviä kirjoja teillä on? Mitkä niistä on hyllyssä heti noudettavissa? Tai oikeastaan riittää että ovat noudettavissa ensi tiistaina. Varaa nuo kaksi minulle ja laita meiliä kun ovat noudettavissa.)
- Vaikeaa saada toimimaan oikeasti hyvin (vrt. Applen, Amazonin ja Googlen kehitysbudjetti ja miten helppoa on silti päätyä tilanteeseen, että Siri, Alexa tai Google ei oikein ymmärrä mitä tarkoittit)
- Esim. Turun kaupungin Åbot <https://covid.neuvo.ai/turku?language=fi>

Hanketuloksia/esimerkki kielen prosessoinnista

- Kirjastomaailman yksi ilmiö on, että tietomäärä on valtava ja aineistoja ei voi yksi ihminen mitenkään koko laajudessa analysoida, vertailla ja ymmärtää
- Kone on ahkera käymään läpi suuriakin aineistoja
- Mitä voidaan oppia, jos sopivalla kielimallilla varustettu kone kahlaa jonkin aineiston läpi oppimistarkoituksessa?

Älykäs ennakointi –hanke 2020–2022

Louhinta yksi työpaketti. HeadAI Theseus-louhinnan teknisenä toteuttajana.

Aineistona Theseuksessa julkaistut opinnäytetyöt (tiivistelmät) vuosilta 2010-2021. Kaikki alat ja ammattikorkeakoulut.

Taustalla oletus: *tehdyt opinnäytetyöt kertovat, mitä asioita*

- a) yritykset kehittävät,*
- b) valmistuvien opiskelijoiden odotetaan osaavan sekä*
- c) opintosisällöissä pitäisi jollain tasolla olla esillä.*

Voiko tällaista ajatusta jossain tilanteessa käyttää esimerkiksi botin kouluttamiseen?

Kestävää kasvua ja työtä -ohjelma

Vipuvoimaa
EU:lta
2014–2020



Euroopan unioni
Euroopan sosiaalirahasto

Pari pohdintakysymystä kotitehtäväksi

- Onko jatkossa enää perinteisiä nettisivuja, vai onko keskusteleva tekoäly käyttöliittymä kaikkeen?
- Miten tekoälyn käymiä keskusteluja valvotaan, ohjataan, kehitetään ja kritisoidaan?
- Jos järjestelmä oppii, miten varmistutaan siitä, että oppimisen lähde on luotettava ja ei-murrettava? Mitä riskejä on, jos botin oppimiseen vaikutetaan jotenkin kielteisellä tavalla tarkoituksellisesti? Esimerkiksi botit viranomaisten sivuilla, tunnistatko kurjia skenaarioita?
- Minkä osan sinun työstäsi älykäs chatbot voi tehdä nyt? Entä 15 vuoden päästä?
- Miten oma työsi muuttuu, kun kollegana on chatbot, joka kytkeytyy keskeisiin tietojärjestelmiin ja antaa sinulle luonnollisen käyttöliittymän kaikkeen relevanttiin tietoon?

Kiitos!

sakari.koivunen@turkuamk.fi

