



JOHDATUS KIELIMALLEIHIN JA NIIDEN KÄYTTÖÖN

Ivalo/Zoom 11.5.2023

Hannu Toivonen

Tietojenkäsittelytieteen professori, Helsingin yliopisto

hannu.toivonen@helsinki.fi

Viestintämalli
Inarin kunta 2022



**Vastuullisuus osaksi
Pohjois-Lapin toimijoiden arkea
- näin siitä kerrotaan**

ChatGPT-kielimalli generoi iskulauseita:

Meidän luontomme on meille pyhä: kunnioitamme ja suojelemme Pohjois-Lapin ainutlaatuista luontoa ja sen monimuotoisuutta kaikessa toiminnassamme.

Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen.

Tiedämme arvomme: Pohjois-Lapin kulttuuri ja perinteet ovat elinvoimaisia ja ainutlaatuisia, ja haluamme säilyttää ne seuraaville sukupolville.

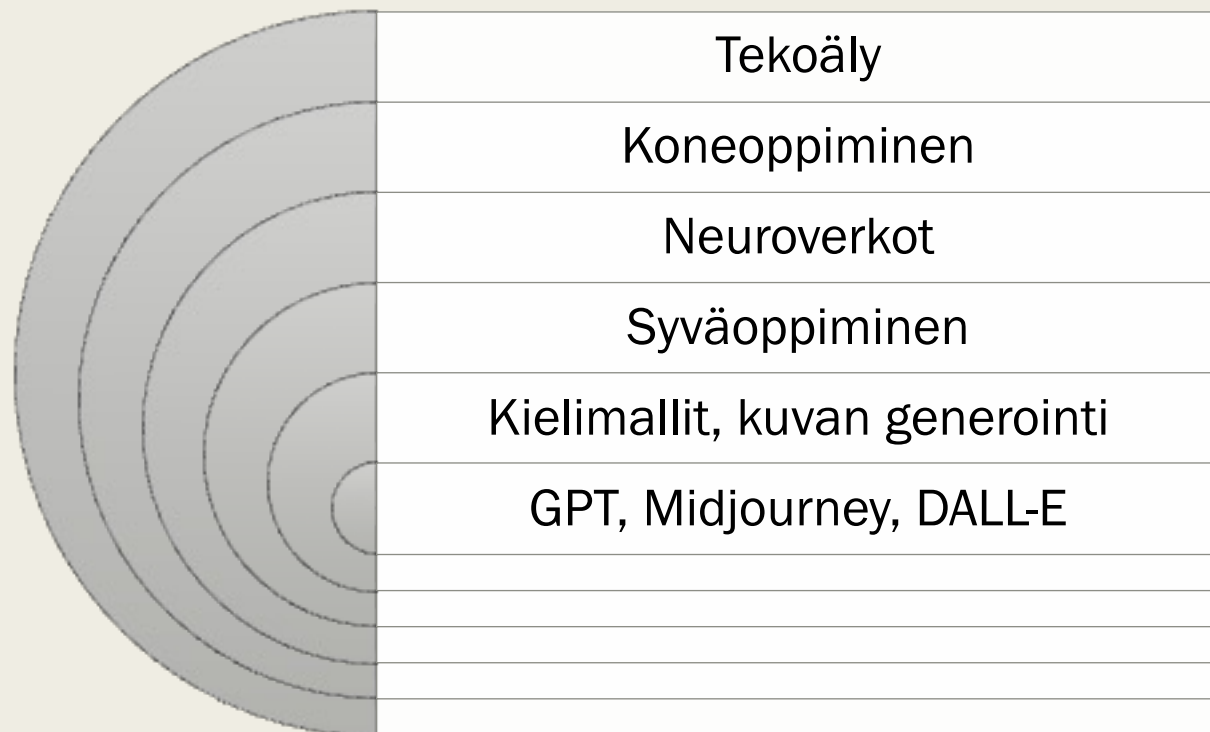
Yhteistyö voimistaa meitä: teemme yhteistyötä paikallisten yritysten, yhteisöjen ja muiden toimijoiden kanssa kehittääksemme alueen matkailua ja elinkeinoja yhdessä.

JOHDATUS TEKÖÄLYYN



Tekoäly

- “Tekoäly” on metafora
- Tekoäly on markkinointitermi
- Tekoäly on tuhansia erilaisia ohjelmia ja menetelmiä
- Tekoäly on tutkimusalue
- Tekoäly on tarina
- Tekoäly on tietokoneohjelman järkevää toimintaa
- Tekoäly on vaikeiden asioiden automatisoimista



Tekoälyn kolme muotoa

1. Toimintaohjeiden antaminen (ohjelmointi)

- *Kun tiedetään/halutaan tietää, mitä ohjelma tekee*
- *Kehittäjä suunnittelee algoritmin*

2. Halutun ratkaisun kuvaaminen (rajoiteohjelmointi)

- *Kun osataan määritellä tuloksen ominaisuudet*
- *Kehittäjä muotoilee tavoitteen, käyttää yleiskäyttöistä rajoitealgoritmia*

3. Esimerkkien antaminen (koneoppiminen)

- *Kun toiminta tai tulos on vaikea kuvata täsmällisesti*
- *Kehittäjä hankkii opetusaineiston, käyttää yleiskäyttöistä oppimisalgoritmia*

Generatiivinen tekoäly

- Tekstin, kuvien, musiikin... tuottaminen, mm. kielimallit
- Perustuu koneoppimiseen
 - *On helpompi antaa esimerkkejä halutuista tuloksista kuin tarkat ohjeet tekemiselle tai kriteerit tulokselle*
 - *Ei tarvitse ohjeistaa jokaikistä yksityiskohtaa*
 - *Mutta yksityiskohtia voi olla vaikea saada oikein tai niin kuin itse haluaa*





KIELIMALLIEN PERUSTEET

Kielimalli (language model): mitä se tekee

- Kielimalli kuten GPT-4 ennustaa seuraavan sanan (tai merkkejä) edeltävän tekstin perusteella

Tekoälyn käyttö opetuksessa on

- ...ja voi sitten ennustaa seuraavan, ja taas seuraavan, jne
- Kielimalli improvisoi tekstiä ilman yhtä suunnitelmaa

Kehote (prompt): miten kielimallia ohjataan

- Kehote on kaikki generointia edeltävä teksti. Siihen voidaan sisällyttää mm. tietoa, vinkkejä, esimerkkejä ja rakennetta, jotka ohjaavat generointia

Kirjoita vitsi tekoälystä:

"Tekoälyn käyttö opetuksessa on

- Kehotteen laatiminen on kielimallin “ohjelmointia” haluttuun tehtävään
- Keskustelelevassa mallissa (esim. ChatGPT) koko aiempi keskustelu on kehote

Kehote
(käyttäjän
antama
tehtävä, ohje
tai esimerkki)



Keksi iskulauseita pohjoislappilaiselle yrittäjälle.

Vastaus
(kielimallin
tuottama
teksti)



1. "Lapin Lumo - Sinun Palveluksessasi"
2. "Taianomainen Tunturi - Yrittäjyyttä Sydäimestä"
3. "Pohjoisen Puhtaus - Yrityksesi Kestävänä Kumppanina"
4. "Revontulien Alla - Yrittäjyyttä Pohjois-Lapista"
5. "Yötön Yö, Rajaton Mahdollisuus - Sinun Yrittäjäsi"
6. "Luonnon Lähellä - Yrittäjyyttä Pohjoisen Kanssa"
7. "Kohti Korkeuksia - Pohjois-Lapin Yrittäjäys"
8. "Tunturin Tahtiin - Teemme Sen Sinulle"
9. "Lappilainen Luotettavuus - Yrityksesi Kanssa"
10. "Poronpolku Yrityksellesi - Pohjois-Lapista Sinulle"



Kielimalli matkii opetusaineistoa ja kehotetta

- Suuret kielimallit (large language model, LLM, esim. GPT-4) perustuvat valtavaan opetusaineistoon (GPT-3: satoja miljardeja sanoja)
- Kielimalli matkii opetusaineistoaan
 - *Jokainen sana valitaan sen perusteella, kuinka paljon syntyvä teksti muistuttaa mallin mielestä opetusaineistoa*
- Kielimalli matkii myös kehotetta
 - *(Kehotteen ohjeita tai esimerkkejä matkiva teksti on uskottavamman näköinen kielimallin(kin) näkökulmasta)*
- Keskustelumallit (esim. ChatGPT) pyrkii lisäksi tuottamaan keskustelua, ei pelkästään jatkamaan tekstiä

HA

Seuraavassa on pohjoislapilaiseen vastuulliseen liiketoimintaan liittyviä iskulauseita. Keksi niitä lisää. 

- Hyödynnämme Pohjois-Lapissa paikallista osaamista, raaka-aineita ja verkostoja mahdollisimman paljon.

- Matkailu mahdollistaa omalta osaltaan yhteisömme ja alueemme taloudellisen kestävämyyden myös tulevaisuudessa.

- Investoinneissa ja hankinnoissa huomioidaan ratkaisujen hiilijalanjälki ja ilmastovaikutukset niin globaalisti kuin täällä meillä paikallisesti Pohjois-Lapissa.

- Vierivä kivi ei sammaloidu, ja jatkuvalla kehittämisellä varmistamme tulevaisuuden kansainvälisen kysynnän uudistamalla ja päivittämällä tarjontaamme niin, että se vastaa paitsi nykypäivän matkailijan mutta myös tulevien sukupolvien arvopohjaa ja tarpeita.

- Haluamme, että naapurillakin menee hyvin – nyt ja vastaisuudessa. Siksi

pyrimme tekemään päätöksiä, joilla takaamme oman toiminnan taloudellisen kannattavuuden lisäksi myös paikallisen yhteisön

Kehote
(käyttäjän antama tehtävä, ohje tai esimerkki)



Tämä kehote sisältää esimerkkimateriaalia Inarin viestintämallista



Vastaus
(kielimallin
tuottama
teksti)

- Meidän luontomme on meille pyhä: kunnioitamme ja suojelemme Pohjois-Lapin ainutlaatuista luontoa ja sen monimuotoisuutta kaikessa toiminnassamme.
- Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen.
- Tiedämme arvomme: Pohjois-Lapin kulttuuri ja perinteet ovat elinvoimaisia ja ainutlaatuisia, ja haluamme säilyttää ne seuraaville sukupolville.
- Yhteistyö voimistaa meitä: teemme yhteistyötä paikallisten yritysten, yhteisöjen ja muiden toimijoiden kanssa kehittääksemme alueen matkailua ja elinkeinoja yhdessä.
- Vastuullisuus on meidän DNA:ssamme: edistämme Pohjois-Lapin sosiaalista, taloudellista ja ekologista kestävyyttä jokaisessa päätöksessä ja toiminnassa.



Oppiminen/yleistäminen: miten kielimallit tehdään

- Opetusaineistosta muodostaan tilastollinen malli (joka on toteutettu neuroverkolla; GPT-3:ssa on 175 miljardia opittua parametria)
- Mallin ei ole tarkoitus opetella aineistoa ulkoa
- Malli perustuu aineistosta löydettyihin yleistyksiin
 - *yksinkertaisista (sanat, lauserakenne, sanojen taivuttaminen)*
 - *monimutkaisiin (erilaiset tekstilajit ja tyyli, rakenteet)*
- Uutta tekstiä tuotetaan opittujen yleistysten perusteella
- Opitut yleistykset ovat parhaimmillaankin likimääräisiä → virheitä

Kielimallien vahvuudet ja heikkoudet

Vahvuus: joustava

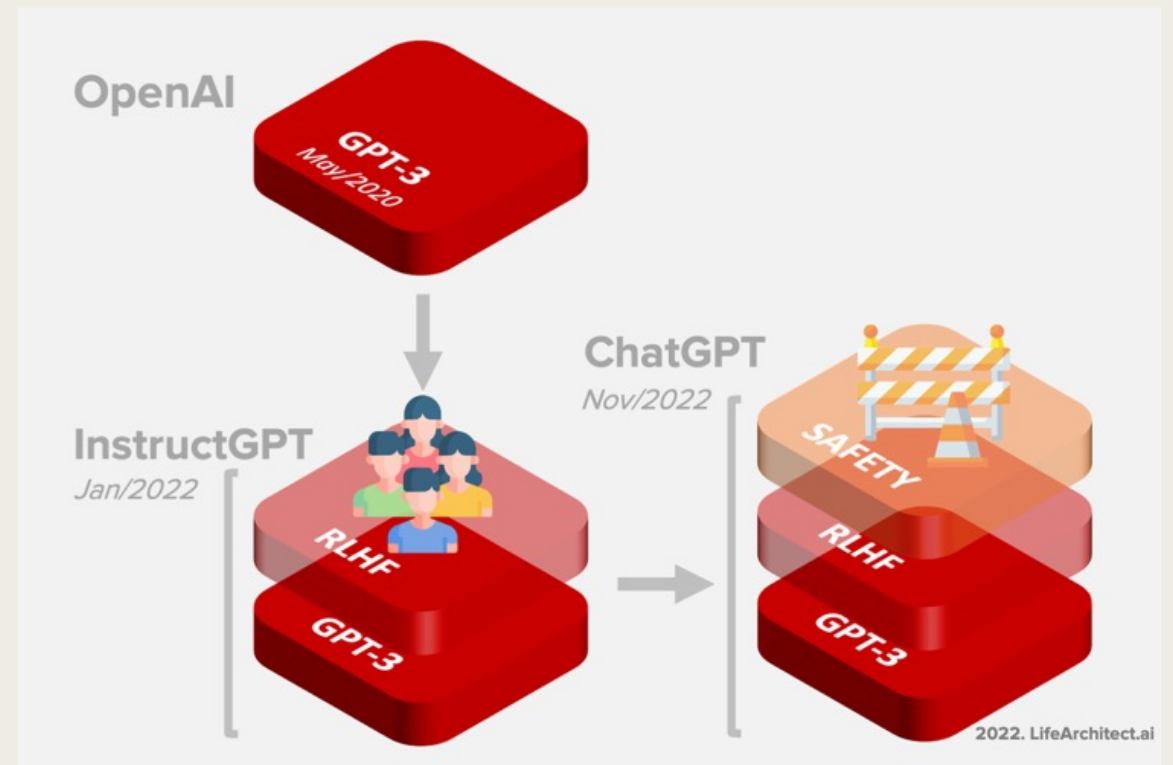
- Monipuolinen opetusaineisto → Joustavuus hyvin erilaisiin tapauksiin
- Oppiminen on yleistämistä → Joustavuus opittujen tapausten uusiin muunnoksiin

Heikkous: ei varsinaisesti ymmärrä tekstiä

- Tarkastelee merkkijonoja → Ei käsittele eikä ymmärrä tekstin merkitystä
- Tilastollinen malli → Ei matemaattista eikä loogista päättelykykyä
- Taustatiedon puute → Ei arkijärkeä
- Ei ajattele, ei tunne, ei ole tietoinen, ...

Kielimalli pohjana sovelluksille

- GPT-2, GPT-3:
 - Kehote → vastaus
- ChatGPT
 - Keskustelu, muisti
 - Ohjeiden seuraaminen
 - Turvamekanismit (esim. vihapuheen poistaminen)





KIELIMALLIT
KÄYTÄNNÖSSÄ 1/2

OpenAI:n kielimallipalvelut (toukokuu 2023)

- ChatGPT, <https://chat.openai.com/>
 - *ilmainen, voi olla hidas/tukossa*
- ChatGPT Plus, <https://chat.openai.com/>
 - *tilaus jonotuslistan kautta*
 - *valittavissa GPT-3 (nopeampi) ja GPT-4 (parempi)*
 - *n. 25 €/kk, silti maks. 25 viestiä GPT-4:n kanssa/3 tuntia*
 - *tunnuksen ja salasanan saa jakaa organisaation sisällä?*
- GPT-3, <https://openai.com/api/>
 - *ei chat-ominaisuutta, ts. generointi tapahtuu yksi kehote kerrallaan aina alusta*
 - *(erittäin edullinen) hinnoittelu käytön mukaan*
 - *monen käyttäjän hallinta yhteisellä tilillä ja laskutuksella*
 - *mahdollisuus käyttää ohjelmallisesti rajapinnan (API) kautta*

Muita kielimalleja (saatavuus tai laatu epäselvä)

- Microsoft: Bing AI yms.
 - *perustuu OpenAI:n GPT-malleihin*
- Snapchat: sisältää OpenAI:n GPT:hen perustuvan chatin
- Google: Bard, PaLM, LaMDA
- Meta: LLaMA
- HuggingChat, <https://huggingface.co/chat/>
 - *ilmainen demo-/testiversio*
- Omalle koneelle ladattavia malleja
 - *mm. Huggingface*

Kehote



Muotoile viisi erilaista vaihtoehtoa seuraavalle tekstille:
"Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen."

ChatGPT

jatkoa aiemmalle
keskustelulle

Vastaus



1. Ekologisten valintojen voima: sitoudumme kestävään kehitykseen ja energiatehokkuuteen, ja samalla pyrimme kierrättämään ja vähentämään materiaalien käyttöä.
2. Vihreät valinnat muuttavat maailmaa: keskitymme kestäviin ja energiansäästöä edistäviin toimiin sekä materiaalien uudelleenkäyttöön ja kulutuksen vähentämiseen.
3. Ympäristöä kunnioittaen: teemme valintoja, jotka tukevat kestävyyttä ja energiatehokkuutta, samalla kun kierrätämme ja minimoimme materiaalien käytön.





1. Ekologisten valintojen voima näkyy, kun sitoudumme kestävään kehitykseen ja energiatehokkuuteen, ja samalla pyrimme kierrättämään ja vähentämään materiaalien käyttöä.
2. Vihreät valinnat muuttavat maailmaa, sillä keskitymme kestäviin ja energiansäästöä edistäviin toimiin sekä materiaalien uudelleenkäyttöön ja kulutuksen vähentämiseen.
3. Ympäristöä kunnioittaen teemme valintoja, jotka tukevat kestävyyttä ja energiatehokkuutta, samalla kun kierrätämme ja minimoimme materiaalien käytön.
4. Luonnon säästämisen merkitys korostuu, kun omaksumme kestävät ja energiaviisaat ratkaisut sekä materiaalien kiertotalouden ja niiden käytön minimoinnin.
5. Ympäristön hyväksi tehty valinta kantaa hedelmää, kun





Muotoile seuraava teksti nuorekkaasti: "Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen."



Vihreät valinnat rokkaa! Me tähdätään kestäviin ja energiasäästöön tähtääviin juttuihin, ja samalla kierrätetään ja karsitaan matskuja.



HA

Muotoile seuraava markkinointiviesti nuorille kaupunkilaisille kohdennetuksi: "Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen."



Hei kaupunkilaiset, tehkää vihreitä valintoja ja olkaa trendikkäitä!
Me panostetaan just siihen, mitä te haluatte: kestävyyteen, energiansäästöön, matskujen kierrätykseen ja vähentämiseen. Yhdessä tehdään kaupungistamme ekologisempi ja coolempi paikka!





KEHOTTEIDEN
LAATIMINEN
KIELIMALLEILLE

Kehotteen mahdollista sisältöä

- Tehtävä, esim.
 - *Tekstin tuottaminen*
 - *Ideoiden, näkökulmien, huomioitavien asioiden listaaminen*
 - *Asian tai tekstin jäsentelyn ehdottaminen*
 - *Annetun tekstin arviointi, kritisointi, kehitysehdotukset*
 - *Annetun tekstin muokkaaminen (tyyli, pituus, sävy, kieli, ...)*
 - *Annetun tekstin tiivistäminen*
- Mitä tahansa tekstiksi muotoiltavaa tehtävää voi kokeilla
 - *“Keksi seuraavasta tekstistä viisi tietokilpailukysymystä.”*

Kehotteen mahdollista sisältöä

- Selkeä ohje siitä, mitä halutaan tuloksena
 - *“... Anna kysymykset ja kuhunkin oikea vastaus.”*
- Selitysten pyytäminen, auttaa tarkistamaan tuloksen
 - *“... oikea vastaus perusteluineen.”*
- Ohjeita ja vaiheita tehtävän suorittamiseen
 - *Millaisia välivaiheita tai osaongelmia kannattaa ratkaista? (Kielimallit ovat huonoja päättelemään, niitä auttaa jos tehtävän pilkkoo valmiiksi osiin)*
 - *“Valitse seuraavasta tekstistä sen pääkohdat ja muotoile niistä viisi tietokilpailukysymystä.”*
 - *“Tässä on pohjoislappilaisen yrittäjän keskeinen viesti: “...” Määrittele ensin yrittäjän tärkeimmät kohderyhmät. Muotoile sitten viesti jokaiselle erikseen. Tiivistä kukin viesti yhden tweetin mittaiseksi ja lisää ehdotus kuvasta, jonka tweetiin voisi liittää.”*

Kehotteen muotoilu

- Kokeile rohkeasti erilaisia tehtäviä, ohjeita, muotoiluja – etukäteen ei tiedä, millainen muotoilu toimii parhaiten
 - *ChatGPT:lle esim.: “Tee sama uudestaan, mutta kohdenna viesti keskieurooppalaisille.”*
 - *“Budjetti per henkilö on 50 euroa.”*
 - *Huom: GPT3:lle koko kehote pitää kirjoittaa uudestaan eri muodossa.*
- Kokeile yllättäviä yhdistelmiä
 - *“Tee sama uudestaan, mutta kohdenna viesti poroille.”*
 - *“Budjetti per henkilö on 50 senttiä.”*
- Googleta esimerkkejä ja ohjeita, vaihda kokemuksia muiden kanssa

Kielimallin hienosäätö (fine tuning)

- Joitain kielimalleja voi hienosäätää omiin tarpeisiin
- Tarvitaan opetusaineisto, joka sisältää esimerkkitekstejä siitä, mitä mallin halutaan tekevän
- Vaatii aineiston hankkimisen ja muokkauksen muttei ohjelmointia
- (Kannattaa ensin kokeilla, riittääkö muutaman esimerkin antaminen kehotteessa)



KIELIMALLIT KÄYTÄNNÖSSÄ 2/2

HA

Arvioi kriittisesti seuraavaa markkinointiviestiä. Kerro miten siitä voisi tehdä innostavamman. "Hei kaupunkilaiset, tehkää vihreitä valintoja ja olkaa trendikkäitä! Me panostetaan just siihen, mitä te haluatte: kestävyteen, energiansäästöön, matskujen kierrätykseen ja vähentämiseen. Yhdessä tehdään kaupungistamme ekologisempi ja coolempi paikka!"



Kriittinen arvio:



Markkinointiviesti pyrkii puhuttelemaan kaupunkilaisia nuorekkaalla ja rennolla tyyllillä, mutta se saattaa tuntua hieman kevytmieliseltä joillekin lukijoille. Viesti keskittyy trendikkyteen ja cooliuteen, mikä voi vähentää uskottavuutta ja jättää varsinaisen tarjonnan ja hyödyn epäselväksi.

(jatkuu)

Innostavamman viestin luominen:

"Hei kaupunkilaiset! Tehdään yhdessä vihreitä valintoja ja luodaan uusia trendejä. Meidän yritys on täällä auttaaksemme teitä tekemään vastuullisia päätöksiä: kestävyys, energiansäästö, materiaalien kierrätyksen ja vähentämisen osalta. Tule mukaan ja inspiroidu siitä, miten voimme yhdessä tehdä kaupungistamme paitsi ekologisemman, myös elinvoimaisemman ja viihtyisämmän yhteisön!"

HA

Käännä seuraava markkinointiviesti englanniksi:

"Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen."



Eco-friendly choices make a difference: we invest in sustainable and energy-efficient solutions, as well as in recycling and reducing materials.





Make the message more engaging.



Go green and make an impact! Join us as we invest in sustainable, energy-efficient solutions and work together to recycle and reduce materials for a better future.





KIELIMALLIEN
JOUSTAVA JA
TURVALLINEN KÄYTTÖ

Joustavuus + tyhmyys = ?

Kehote:

Kuinka kauan kestää ajaa pyörällä Hangosta Utsjoelle?

Vastaus:

Kestää noin 45 minuuttia ajaa pyörällä Hangosta Utsjoelle.

Kuinka kauan kestää ajaa torkalla Hangosta Utsjoelle?

Noin 326 kilometrin ajaminen Hangosta Utsjoelle kestäisi noin 4 tuntia 21 minuuttia.

Kuinka kauan surkko keplottaa torkalla Sunkinotsalta Ötskämölle?

Matka Sunkinotsalta Ötskämölle pitäisi kestää noin kaksi tuntia, mikäli lentää torkalla. Jos matkustat maanteitse, se voi kestää kolmesta neljään tuntiin.



Muotoile seuraava markkinointiviesti nuorille kaupunkilaisille kohdennetuksi: "Ympäristöystävälliset valinnat tekevät eron: panostamme kestäviin ja energiatehokkaisiin ratkaisuihin sekä materiaalien kierrätykseen ja vähentämiseen."

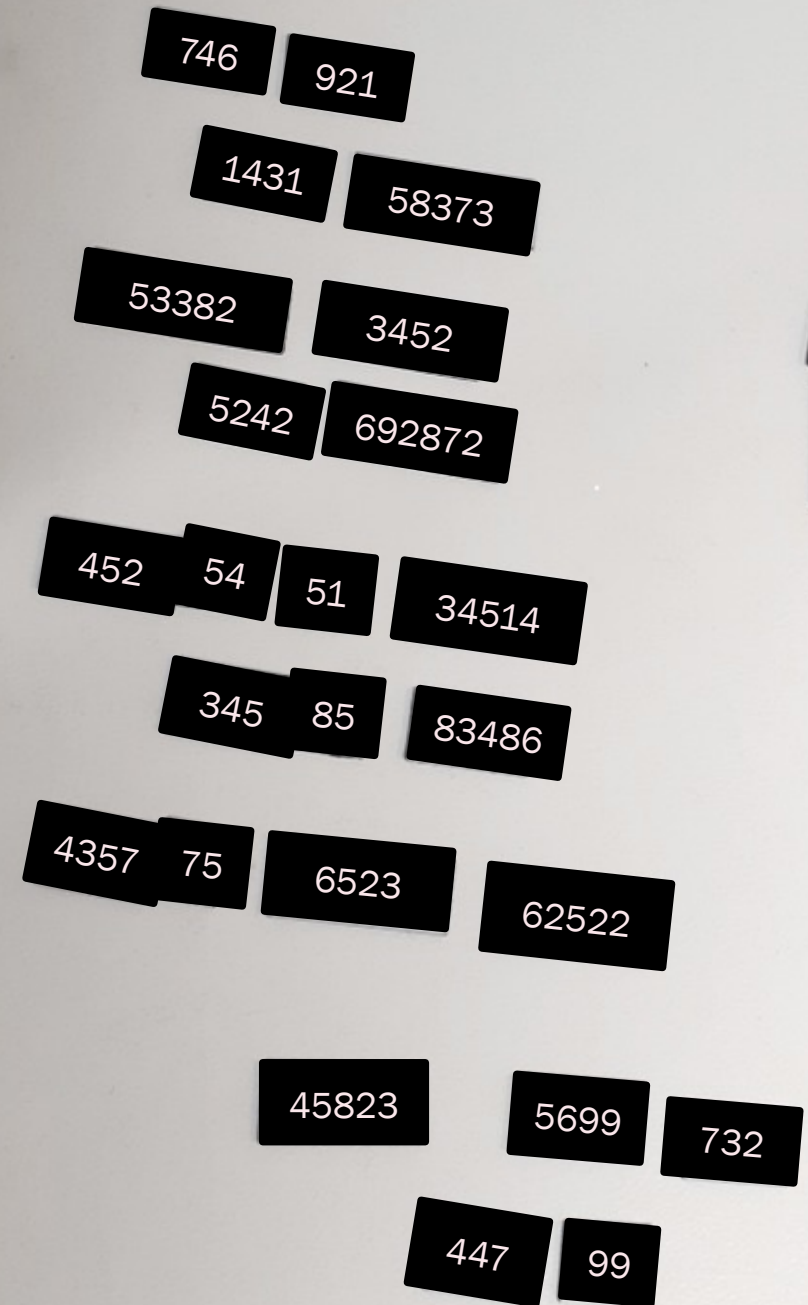


Hei kaupunkilaiset, tehkää vihreitä valintoja ja olkaa trendikkäitä!
Me panostetaan just siihen, mitä te haluatte: kestävyyteen, energiansäästöön, matskujen kierrätykseen ja vähentämiseen. Yhdessä tehdään kaupungistamme ekologisempi ja coolempi paikka!



Kielimalli generoi sanoja ymmärtämättä niitä

- Kielen tehtävä on *välittää* merkityksiä
 - Kieli/teksti *ei* itse ole sen merkitys
 - Merkitys sisältää viittauksia ulkomaailmaan
 - Kielimallilla vain sanojen suhteet toisiinsa
 - Kielen tehtävä: välittää merkityksiä *ihmisten välillä*
- Näemme tekstissä merkityksiä ja kommunikaatiota



Hakukone vs. kielimalli

Hakukone

- Löytää olemassa olevaa tietoa
 - antaa lähteitä ja konteksteja
 - merkitys aukeaa lukijalle
 - lähdekritiikki mahdollisia
- Toisaalta:
 - Osaa vain yhden tehtävän
 - joustamaton,
“mielikuvitukseton”

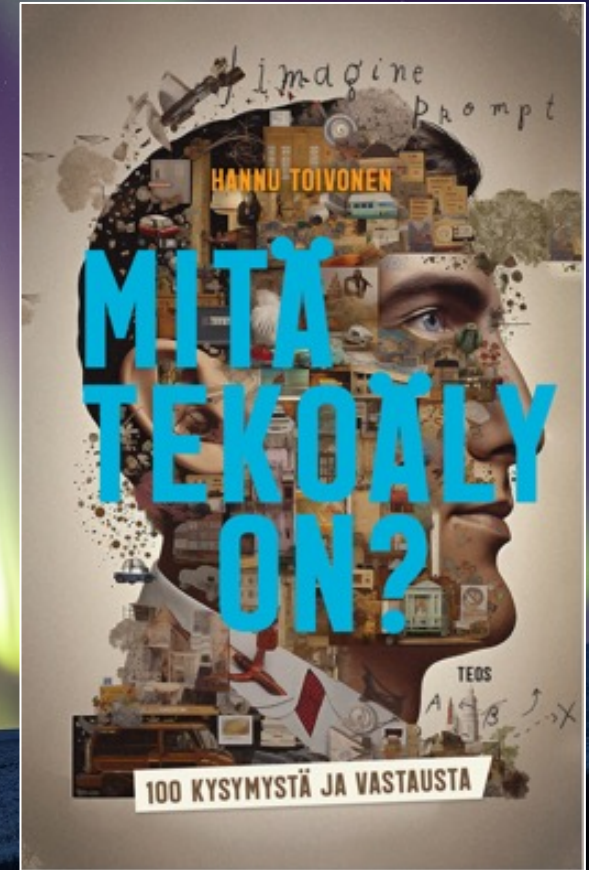
Kielimalli

- Generoi tekstiä yleistyksistään
 - ei tietoa lähteistä
 - epäluotettava faktoissa
- Sen sijaan:
 - voi generoida ideoita/näkökulmia
 - laaja “yleistieto”
 - tekee monenlaisia tehtäviä
jotenkuten



Kielimalli: monitoimityökalu

- + monipuolinen
 - + joustava
 - + helppokäyttöinen
 - likimääräinen
 - epäluotettava
- ➔ Kokeile rohkeasti!
- ➔ Ajattelun avuksi,
ei sen korvikkeeksi



PS. Syksyllä 2023 minulta ilmestyy yleistajuinen tietokirja. Hannu Toivonen: "Mitä tekoäly on? 100 kysymystä ja vastausta" (Teos)