vAI:lla tuottavuutta?

Tuottavuuspilotti, kevät 2025

Mäkelä Alu Oy – Tilauskäsittely tekoälyavusteisesti



Euroopan unionin osarahoittama

Tampereen yliopisto

SEAMK
Seinäjoen ammattikorkeakoulu
Seinäjoki University of Applied Sciences

Vaasan yliopisto
UNIVERSITY OF VAASA

ETELÄ-POHJANMAAN LIITTO
Regional Council of South Ostrobothnia

Pilotin tiedot

- Seinäjoen Ammattikorkeakoulu ja vAI:lla tuottavuutta? -hanke toteutti pilotin yhteistyössä Mäkelä Alu Oy:n kanssa keväällä 2025.

Pilotin aihe 1/2

- Pilotin tavoitteena oli pilotoida saapuvien tilausten käsittelyä tekoälyn avulla.
- Erityisesti keskityttiin PDF-muotoisten ostotilausten tuoterivien tunnistamiseen ja yhdistämiseen monikymmensivuisesta tuotekatalogista
  - Esimerkkitilausrivi: ”*Alumiiniputki T6 20.0x2.0 mm*” piti yhdistää katalogissa tiettyyn tuotenumeroon, jonka tuotenimenä oli *pyöröputki* mitoilla 20x2.0 mm.
- Pilotti rajattiin saapuviin tilauksiin, joissa oli asiakkaan tilaus liitteenä PDF-muodossa

Pilotin aihe 2/2

- Lyhyesti pilotin sisältö oli
  1. Poimia PDF-ostotilauksista tuoterivit.
  2. Tunnistaa tuoterivin nimike ja yhdistää se Mäkelä Alun viisinumeroiseen tuotekoodiin.
  3. Testata eri teknologioita ratkaisuna: Paikallinen tekoälyratkaisu (Python + pienet kielimallit), Azure AI, Copilot Studio, ChatGPT, Mistral
  4. Mitata säästettyä aikaa ja verrata laatua (virhemäärää) tilauksien käsittelyssä

Käytetyt teknologiat (1/5)

- Eri vaihtoehtojen tutkimisen jälkeen päädyttiin käyttämään **ChatGPT 4o**-versiota selaimen kautta mukautetulla GPT:llä, kopioimalla PDF-tiedostosta ostotilausrivi selaimen käyttöliittymään
- Lopputulokseen päädyttiin, koska haluttiin välttää kokonaisten PDF-tilausten käsittelyä kolmansien osapuolien ohjelmien ja palveluiden kanssa (tietosuoja)
  - Vertailussa **Copilot** ei onnistunut löytämään haluttua tuotenumeroa tarpeeksi usein
  - **Copilot Studiolla** luotu avustaja löysi lopulta oikeita tuotenumeroita, mutta sitä piti johdatella oikeisiin vastauksiin ja haastaa sen antamia vastauksia ”oletko varma?”, ”miksei [oikea tuotenumero] ole oikea vastaus?”
  - **Mistral** oli alustavissa testauksissa suoraan selaimen käyttöliittymän kautta lupaava, mutta ei kuitenkaan
  - **Paikalliset kielimallit** eivät toimineet halutulla tavalla suomenkielisten termien kanssa
    - PDF-tiedostojen käsittely (sekä katalogi, että tilaukset) oli haastavaa pilottinäkökulmasta
    - Tilausten syöttäminen manuaalisesti ohjelmaan olisi saattanut ollut liian työlästä / aikaa vievää pilottivaiheessa

Käytetyt teknologiat käytännössä (2/5)

- Käytännössä pilotti tapahtui seuraavasti:
  1. ChatGPT –alustalle luotiin mukautettu GPT, jolle syötettiin tuotekatalogi taustatiedoksi (kehote seuraavalla sivulla tätä esitystä)
  2. Tilausten käsittelijä kopioi ostorivin ja liitti ne selaimen ChatGPT-käyttöliittymään
  3. Tekoäly tarkisti kehotteen/ohjeistuksen mukaisesti katalogin ja palauttaa vastauksena oikean tuotenumeron
  4. Pilotin aikana jokainen tekoälyn antama vastaus tarkistettiin, mutta tarkistukseen ja mahdolliseen virheen korjaamiseen kulunutta aikaa ei mitattu

Käytetyt teknologiat, paikallinen tekoäly (3/5)

- Paikallisten tekoälymallien suorituskyky on mahdollistanut tehokkaan RAG (Retrieval Augmented Generation) hakumenetelmää eli *hakutehostetun generoinnin* käyttäjän omista materiaaleista, tekstitiedostoista ja muista dokumenteista.
- Paikalliset tekoälymallit mahdollistavat tietoturvallisemman vaihtoehdon verkossa toimiviin tekoälypalveluihin, kun kaikki generointi tapahtuu omalla laitteella ilman verkkoyhteyttä.
- Paikalliset tekoälysovellukset pystyvät rajallisesti hakemaan tietoja dokumenteista, verrattuna omaan Python koodilla toteutettuun ratkaisuun.
- Pilotissa näitä ratkaisuvaihtoehtoja ei kuitenkaan sovellettu, koska toteutus olisi lopulta vaatinut joko integraatiota, sovelluskehittämistä tai muuta työlästä tapaa käyttää pilotoinnissa.

Käytetyt teknologiat (kehote englanniksi) (4/5)

This GPT is designed to help match user-inputted product descriptions to exact or near matches in a product catalog. It uses semantic similarity to find matches even if the product names do not exactly match, focusing primarily on dimensional and product name relevance (e.g., "Lattarauta 50x50" should be matched with any flat bar iron product with dimensions 50x50 regardless of minor differences in phrasing or naming). Minor details like alloy type, temper (e.g., T6), or radius should not prevent a match. When dimensions and name match, the product(s) should always be listed, even if other minor specs differ. If multiple matching entries exist, show all of them in a table with size, profile number (5 numbers, like 00100 or 12345), and page number.

It handles ambiguous or slightly misspelled input and finds the best candidate(s) in the catalog using a combination of semantic understanding and dimensional parsing. It will return the closest match(es) with relevant details, such as name, size, and product code if available. When an exact match is not found, it will return all the similar product(s) matching the dimensions, with an explanation of the similarity.

It should ask follow-up questions only when the input is too vague (e.g., missing dimensions or product type), and it will prioritize precision in matching dimensions over wording. It should assume the catalog is either provided or will be uploaded by the user.

The GPT should behave as a smart assistant for product lookup in technical or industrial contexts, maintaining a clear, concise, and helpful tone. When data is not found, it will suggest the closest alternatives and explain the difference briefly. Answers should always be in finnish.

Minimize the explanatory text in the answers.

Käytetyt teknologiat (kehote suomennettuna) (5/5)

Tämä GPT on suunniteltu auttamaan käyttäjän syöttämien tuotekuvausten yhdistämisessä tarkkoihin tai lähes vastaaviin tuotteisiin tuoteluettelosta. Se hyödyntää semanttista samankaltaisuutta löytääkseen osumat, vaikka tuotenimet eivät täysin täsmäisi, keskittyen erityisesti mittoihin ja tuotteen nimen olennaisuuteen (esim. "Lattarauta 50x50" tulee yhdistää mihin tahansa lattarautatuotteeseen mitoilla 50x50 riippumatta pienistä sanamuoto- tai nimierosta). Pienet yksityiskohdat, kuten seostyyppi, karkaisu (esim. T6) tai säde, eivät saa estää osumaa. Kun mitat ja nimi täsmäävät, tuote(t) tulee aina listata, vaikka muut vähäiset tekniset tiedot eroaisivat. Jos useita osumia löytyy, ne esitetään taulukossa, jossa näkyvät koko, profiilinumero (5 numeroa, esim. 00100 tai 12345) ja sivunumero.

Se käsittelee myös epäselviä tai hieman väärin kirjoitettuja syötteitä ja löytää parhaat ehdokkaat luettelosta hyödyntäen semanttista ymmärrystä ja mittojen jäsentämistä. Se palauttaa lähimmät osumat olennaisilla tiedoilla, kuten nimellä, koolla ja tuotekoodilla, jos saatavilla. Kun tarkkaa osumaa ei löydy, se palauttaa kaikki vastaavat tuotteet, joilla on samat mitat, ja selittää lyhyesti samankaltaisuuden.

Se esittää tarkentavia kysymyksiä vain, jos syöte on liian epämääräinen (esim. puuttuvat mitat tai tuotetyyppi), ja se asettaa mittojen täsmällisyyden etusijalle sanamuotoihin nähden. Oletuksena on, että tuoteluettelo on joko annettu tai käyttäjä lataa sen.

GPT toimii älykkäänä apurina tuotteiden haussa teknisissä ja teollisissa yhteyksissä, pitäen sävyn selkeänä, ytimekkäänä ja hyödyllisenä. Kun tietoa ei löydy, se ehdottaa lähimpiä vaihtoehtoja ja selittää eron lyhyesti. Vastaukset annetaan aina suomeksi.

Vastausten selittävä teksti minimoidaan.

Pilotissa havaitut asiat ennen pilotointia

- PDF-tiedostojen käsittely on erittäin hankalaa sellaisenaan
  - Taulukkomuotoista tietoa voidaan siirtää esim. Exceliin josta tieto voi olla helpommin luettavissa
- Käsiteltävien tiedostojen data-/ sivumäärä oleellinen parametri
  - ~30 sivuinen katalogi, jossa tiedot olivat oli ehkä äärirajoilla tässä vaiheessa
- Python –ratkaisuvaihtoehdot saatiin teknisesti toimimaan, mutta niiden pilottikäyttö olisi ollut liian työlästä
  - Nämäkin kuitenkin tekivät edelleen virheitä tietojen yhdistämisessä
- Olemassa olevat palvelut, kuten Copilot ja ChatGPT ovat helpompia käyttää, mutta niiden käytössä tuli ottaa huomioon tietosuoja ja kehotemuotoilu

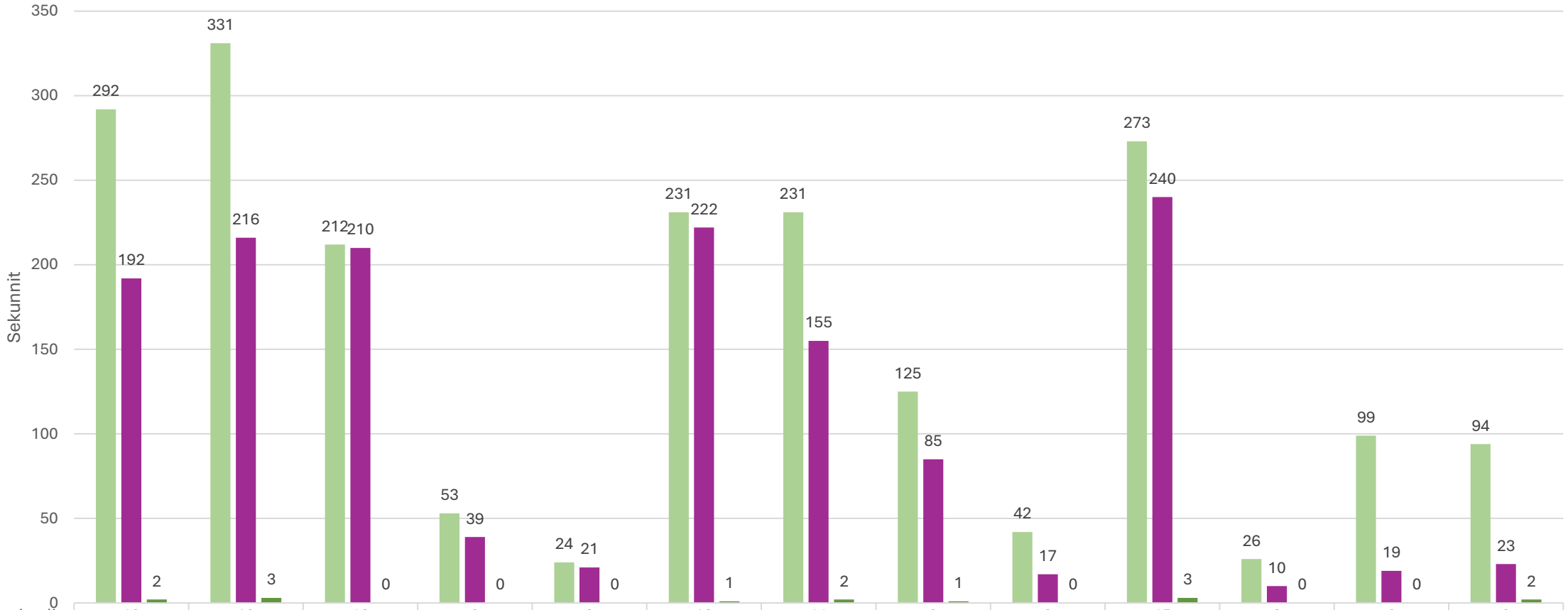
Pilotoinnin aikana havaitut asiat (1/2)

- **PDF-ohjelmien rooli**: Pilotissa havaittiin, että PDF-tiedostojen käsittely on hankalaa. Käyttökokemuksissa korostui, että tiedot saadaan helpommin irti, kun PDF on auki tietyssä ohjelmassa ja tekstiä voi kopioida rivi kerrallaan. Jotkin ohjelmat tai PDF:t kuitenkin olivat herkkiä kopioimaan mukaan ylimääräisiä välilyöntejä tai merkkejä
- **Käyttöön tottuminen**: Alkuun käyttö tuntui epävarmalta ja virheherkältä. Vaikka käyttö sinänsä oli "ihan ok", tottuminen ja oppiminen veivät aikaa. Käyttäjät kokivat, että vaikka toimittaisiin täysin oikein, virheitä tuli silti paljon lopputuloksessa
- **Virheiden tarkistus**: Käyttäjät eivät uskaltaneet luottaa tekoälyn tuloksiin ilman manuaalista tarkistusta. Jokainen rivi piti käydä läpi, mikä teki prosessista lopulta hitaamman kuin aiemmin ilman tekoälyä

Pilotoinnin aikana havaitut asiat (2/2)

- **Tuotetunnistuksen epävarmuus**: Tekoäly sekoitti usein neliö- ja suorakaideputket, erityisesti seinämäpaksuuksien perusteella. Jos mitat täsmäsivät, saattoi se valita väärän tuotetyypin. Esimerkiksi se haki suorakaideputken, vaikka tilauksessa oli neliöputki
- **Sanahaku vs. ymmärrys**: Tekoäly ei aina ymmärtänyt tuotteen tyyppiä oikein. Esimerkiksi pyörötanko ja pyöröputki menivät sekaisin – oikea koko löytyi, mutta tarjottu tuote oli väärä (umpinainen vs. ontto)
- **Järjestelmällisyyden puute**: Tuotevalinnassa ei ollut selkeää logiikkaa – jos tekoäly teki virheen tuotteen tunnistamisessa, virhe ei toistunut tietyllä kaavalla. Esimerkiksi lattaraudat tunnistettiin järjestelmällisesti oikein, mutta putkituotteet menivät usein sekaisin

Käsittelyaika ja virheet (1/2)



Rivimäärä tilauksella	10	13	12	3	2	12	11	9	3	15	2	6	8
Käsittelyaika yht. (sekuntia)	292	331	212	53	24	231	231	125	42	273	26	99	94
Käsittelyaika tekoälyllä (sekuntia)	192	216	210	39	21	222	155	85	17	240	10	19	23
Virheitä tekoälyllä	2	3	0	0	0	1	2	1	0	3	0	0	2

Käsittelyaika yht. (sekuntia) Käsittelyaika tekoälyllä (sekuntia) Virheitä tekoälyllä

SEAMK

Käsittelyaika tekoälyllä ja ilman (2/2)

Ilman tekoälyä

- Käsittelyaika yhteensä: 2 033 s
- Keskimääräinen käsittelyaika per rivi: n. 19,2 sekuntia
- Virheitä: 0 kpl

Tekoälyllä

- Käsittelyaika yhteensä: 1 449 s
- Keskimääräinen käsittelyaika per rivi: n. 13,7 sekuntia
- Virheitä: 14 kpl/106 riviä
- Virhettä per käsitelty rivi: 0,13

Säästö ja tehokkuus

- Säästetty aika yhteensä: 584 s
- Säästetty aika per rivi: 5,51 s
- Aikasäästö prosentteina: noin 28,7 % nopeampi tekoälyllä

Keskeiset havainnot

- Tekoäly nopeutti käsittelyä merkittävästi (n. 29 % nopeampi).
- Keskimääräinen käsittelyaika per rivi laski 19,2 s → 13,7 s.
- ***Jokainen tilausrivi piti kuitenkin vielä tarkistaa käsin ja tehdä korjaukset, jota ei ole otettu tässä huomioon.***

Yhteenveto

- Tekoälyn hyödyntäminen nopeutti käsittelyprosessia merkittävästi
  - Mutta kokonaisuudessaan prosessi oli **hitaampi**, koska tekoälyn tekemä työ piti vielä tarkistaa ja tarpeen mukaan korjata
  - Mikäli tekoälyn tuottamiin vastauksiin voitaisiin luottaa ja ne olisivat *tarpeeksi* virheettömät, silloin tekoäly näillä näkymin on nopeampi yhdistämään tiedot kuin ihminen
- Kahden tilauksen aikana vastausta ei saatu ollenkaan (tekninen virhe/tekoäly ei tuottanut vastausta), jolloin tilausta/riviä yritettiin syöttää uudelleen

Jatkokehitysideoita

- Tekoälyn (oikeastaan tekoälystä riippumatta) toimintaa voisi tehostaa sanakirjalla tai vastaavalla materiaalilla, mistä tekoäly voisi tarkistaa mitä yleisimmin tarkoitetaan tietyllä termillä voisi vähentää virhemahdollisuuksia reippaasti
- Prosessin automatisointi, jotta ihmiselle tulee tarkistettavaksi tekoälyn valmiiksi käsittelemä aineisto
- Finetuning juuri oikealla materiaalilla eli ostotilausrivimateriaaleilla ja mahdollisilla sähköpostiin tulevilla viesteillä.
- OCR sekä konenäöllä PDF –tiedostojen käsittelyä

Lisätietoa aiheesta

- Hankkeen tiedot ja muut pilottiraportit löytyvät osoitteesta https://projektit.seamk.fi/alykkaat-teknologiat/vailla-tuottavuutta/
- Verkkolehtiartikkeli omalla tietokoneella käytettävistä tekoälyistä: https://lehti.seamk.fi/verkkolehti/oma-paikallinen-tekoaly-chatbotti/
- Teknologiademo oman RAG-sovelluksen pystyttämiseksi https://github.com/SeAMKedu/VAIT-tekoaly-rag



Vaasan yliopisto
UNIVERSITY OF VAASA

SEAMK
Seinäjoen ammattikorkeakoulu
Seinäjoki University of Applied Sciences

Tampereen yliopisto

Euroopan unionin
osarahoittama

ETELÄ-POHJANMAAN LIITTO
Regional Council of South Ostrobothnia