

# LOPPURAPORTTI

## Laskujen oikeellisuuden tarkastaminen ja niistä saatavan datan hyötykäyttäminen toiminnan ohjaamisessa sekä kehittämisessä

### Tiimi

Jari Virtala, Kaupunkiympäristön toimiala  
Veijo Nerg, Kaupunkiympäristön toimiala  
Anni Lehto, CGI  
Tuomo Pursiainen, CGI  
Sami Minkkinen, CGI

Helsinki



*“Laskujen automaattisen tarkistuksen ulottaminen sisältöön toisi merkittävää säästöä ja tehokkuutta nykytilaan verrattuna. Tarvittava sisältö tulee jatkossa saada suoraan toimittajalta. Tämän jälkeen automaation ja tekoälyn käyttö on suoraviivaista”*

# Sisällysluettelo

1. **Tiivistelmä**
2. **Kokeilun tavoitteet**
3. **Kokeilun keskeiset opit**
4. **Kokeilun eteneminen**
5. **Kokeilun tuotokset**
6. **Opit kokeiltavan ratkaisun tai toimintatavan mahdollisuuksista**
7. **Opit asiakkaiden tai palvelun käyttäjien tarpeista**
8. **Opit ratkaisun kehittämisestä teknisesti**
9. **Opit kokeilemisestä yleensä**
10. **Opit kokeiluprojektin arjen pyörittämisestä**
11. **Kokeilun tekninen ympäristö**
12. **Kokeilun data**
13. **Jatkopäätökset ja -ideat**

# 1. Tiivistelmä

## Kokeilun tavoitteet

- Kokeilun päätavoitteena oli ymmärtää, millä tavalla kiinteistöautomaation laskunkäsittelyn prosessia voisi parantaa ja manuaalisia työvaiheita vähentää prosessikehityksellä ja tekoälyavusteisella automaatiolla.

## Kokeilun opit

- Toiminta nykyisellään sallii toimittajien toimittaa hyvin erimuotoisia ja erisisältöisiä laskuja Helsingille käsittelyyn, ja tämä on haasteellista.
- Laskumateriaalissa on koneluetavuuden ja automaation kehittämisen kannalta huomattavia puutteita (esim. rivitietojen otsikoiden ja XML-tägien puute).
- Azure AI Document Intelligence on työkaluna hyvinkin tehokas tunnistamaan tietoja PDF-tiedostoista ja laskujen kuvista.
- Tekoälyn ja sääntöpohjaisen automaation mahdollisuuksista saatiin uutta ymmärrystä ja näkemys, kuinka prosessi voitaisiin automatisoida, kunhan tarvittavat tiedot on saatu toimittajalta.

## Suosituksat jatkotoimenpiteiksi

- Seuraavissa puitesopimusneuvotteluissa olisi suositeltavaa vaatia toimittajia käyttämään laskuissa standardimuotoa ja -tietosisältöä. Vaadittu tietosisältö on tuotettu Tilat-palvelun kiinteistöautomaation osalta tässä kokeilussa.
- Toimittajan pitää hyödyntää kaikki käytettävissä olevat XML-tägit, jotta laskujen oikeellisuutta voidaan helposti tarkastaa koneellisesti. Tarvittaessa voidaan käyttää määrämuotoista tietoa esim. FREE\_TEXT -tägin alla.
- Tietojen "summausta" ei saisi käyttää, vaan tavarat/tuotteet ja työ pitää löytyä eriteltynä XML-tiedostosta.

## Laskujen oikeellisuuden tarkastaminen

Uskomme että tekoälyllä olisi mahdollista automaattitarkistaa laskuja melko hyvin ja että paremman laskudatan perusteella voidaan kehittää toimintaa

Kokeillaksemme tätä aiomme kouluttaa tekoälyn vanhalla laskudatalla ja testata sitä toisella vanhalla datasetillä

.. ja mittaamme tunnistaako tekoäly erilaiset virheet

Olemme oikeassa, jos tekoäly osaa tunnistaa 90% virheistä (false negative <10%, false positive selvästi <1%)

# 2. Kokeilun tavoitteet

## Ongelman kuvaus

Helsingin kaupungin Tilat-palvelu käsittelee isännöitsijöiden tekemien tilausten laskuja Baswaren P2P-järjestelmässä. Tilauksia tehdään Helsingin (puitesopimusten alaisilta) alihankkijoilta. Usein laskut tulevat käsittelyyn puutteellisin tiedoin (esim. työn kuvaus, tilausnumero, sopimustuntihinta ja hintaluokat voivat puuttua), ja siksi niiden täsmäminen on haasteellista ja käsitteleminen vaatii huomattavan määrän manuaalista työtä.

Yhteenvetona: Helsinki haluaa sujuvoittaa laskujen tarkastamisen ja hyväksymisen prosessia, vähentää manuaalisen työn määrää ja aloittaa toiminnan kehittämisen (esim. puitesopimustoimittajien hintavertailu).

## Rajaus

Kokeilu rajattiin koskemaan ainoastaan Kaupunkiympäristön toimialan Tilat-palvelun Kiinteistöautomaatioon liittyviä laskuja.

# 2. Kokeilun tavoitteet

## Oletukset

Ennen teknistä kokeilua oletettiin, että tekoäly voi merkittävästi tehostaa laskujen käsittelyä ja tarkastusta, mutta se on riippuvainen laskujen sisältämistä tiedoista. Sääntöpohjaista automaatiota varten laskujen tulisi olla samansisältöisiä ja -muotoisia.

## Tavoitteet

1. Ymmärtää nykyinen Kiinteistöautomaation laskunkäsittelyn prosessi, mm.
  - a. mitä toimijoita prosessiin osallistuu?
  - b. mistä tehtävistä prosessi koostuu?
  - c. mitä tietoja laskuilta tarkastetaan manuaalisesti?
2. Ymmärtää, millä tavalla prosessia voisi parantaa esim. tekoälyavusteisella automaatiolla.
3. Kartoittaa, mitä tietoja laskumateriaalissa (XML, PDF) on ja mitä puuttuu.
4. Koostaa optimitietosisältö laskuille niin, että siinä on mukana kaikki Helsingin kaupungin tarvitsemat tiedot rakenteellisessa (= helposti koneluettavassa) muodossa.
5. Kokeilla Azuren Document Intelligence -työkalua laskujen tarkastamiseen ja selvittää, onko manuaalisesta laskuntarkastuksesta mahdollista luopua tulevaisuudessa.

# 3. Kokeilun keskeiset opit

## Kokeilu tuki alkuoletuksia vahvasti.

- Laskujen tietojen puutteellisuuden osalta saatiin vahvistus oletuksille.
- Tekoälyn ja sääntöpohjaisen automaation mahdollisuuksista saatiin uutta ymmärrystä ja näkemys, kuinka prosessi voitaisiin automatisoida, kunhan tarvittavat tiedot on saatu toimittajalta laskulla.

## Kokeilun alussa asetettuihin tavoitteisiin päästiin.

- Saavutettiin tarvittava ymmärrys nykyprosessista ja määriteltiin, mikä olisi toiminnan tavoitetila.
- Saavutettiin tarvittava ymmärrys nykylaskujen tietosisällöstä ja laskujen muotoon ja sisältöön liittyvistä haasteista ja puutteista.
- Tekoälykokeilu oli mahdollista toteuttaa onnistuneesti Helsingin kaupungin Azuressa.

# 4. Kokeilun eteneminen

Kokeilu jaettiin viiteen vaiheeseen, joiden sisällöt on listattu lyhyesti alla.



# 4. Kokeilun eteneminen

Kokeilu kesti ajallisesti noin kolme kuukautta (syyskuun loppu – joulukuun alku) sisältäen seuraavat Teamsillä toteutetut tapaamiset ja työpajat sekä projektitiimin tekemät työt:

- 27.9.2024: **Aloituspalaveri**
- 4.10.2024: **Nykyprosessin kuvaaminen**
- *Taustalla: Baswaren tuen kontaktointi, keskustelu Talpan kanssa*
- 31.10.2024: **Talpan ja P2P-järjestelmän rooli laskujen käsittelyssä**, keskustelu
- 1.11.2024: **Laskun optimitietosisällön laatiminen**, työpaja
- *Taustalla: Teknisen ympäristön pystyttäminen ja käyttöoikeuksien anominen, XML-tiedostojen analysointi, Azure AI Document Intelligence –kokeilu*
- 27.11.2024: **Välikatselmointiin valmistautuminen**
- 29.11.2024: **Välikatselmointi**
- 10.12.2024: **Lopputapaaminen**

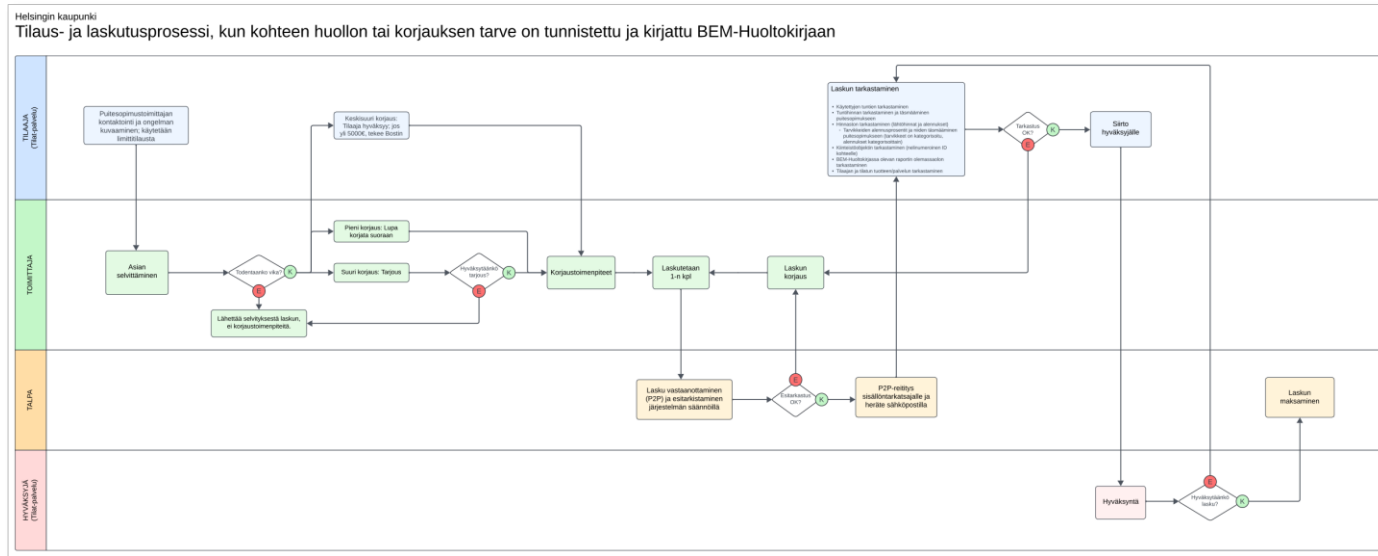
Huom.! Alkuperäisen suunnitelman mukaisesti kokeiluun oli tarkoitus ottaa mukaan kaksi työkalua: M365Copilot ja Azure AI Document Intelligence. Copilotia ei kuitenkaan lopuksi käytetty vaan laskujen xml-tiedostojen analysointi tehtiin käsin. Tuotantoversiossa käytettäisiin esim. Azure OpenAI Services –palvelua generatiivista tekoälyä vaativiin palveluihin.



# 5. Kokeilun tuotokset

Laskunkäsittelyn nykyprosessi (työstetty Lucidchartissa, lopputuotos PDF-tiedosto)

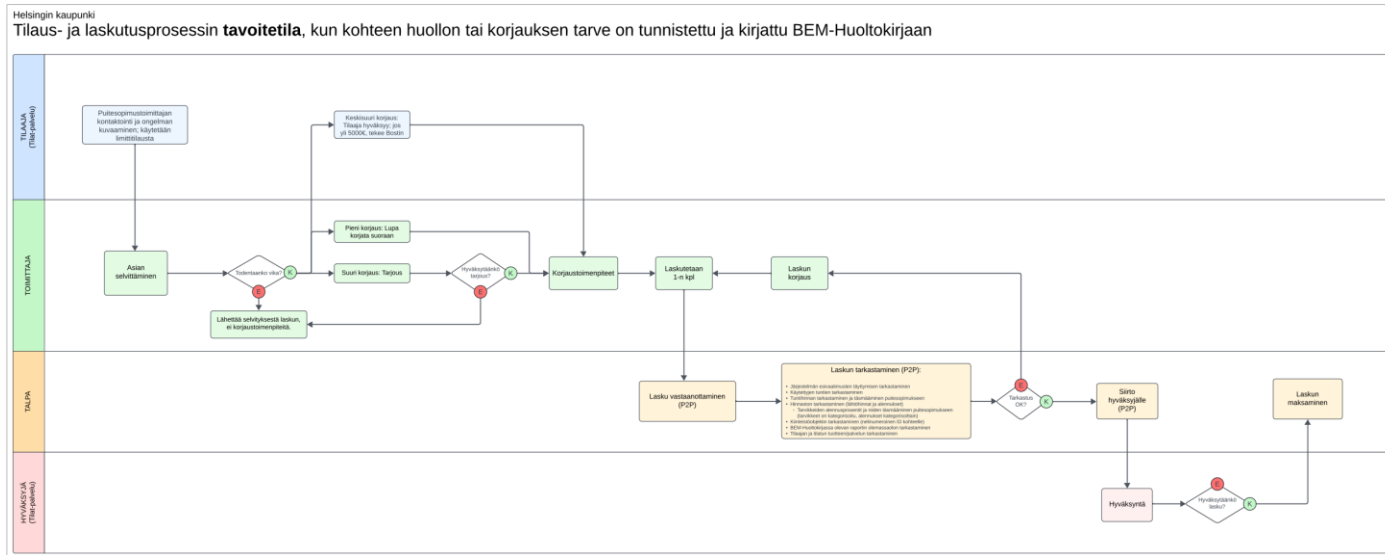
Kuvattiin nykyprosessi kiinteistöautomaatioon liittyvän palvelun ja/tai tuotteen tilauksesta laskun maksamiseen saakka. Tunnistettiin toimijoiksi Tilat-palvelun tilaaja, puitesopimustoimittaja, kaupungin talouspalvelut (Talpa) ja Tilat-palvelun laskun hyväksyjä. Laskunkäsittelyyn käytetään Baswaren P2P-järjestelmää.



# 5. Kokeilun tuotokset

## Laskunkäsittelyn tavoitetilan prosessi (työstetty Lucidchartissa, lopputuotos PDF-tiedosto)

Kuvattiin tavoitetilan prosessi kiinteistöautomaatioon liittyvän palvelun ja/tai tuotteen tilauksesta laskun maksamiseen saakka. Toimijoina samat tahot, mutta laskun tarkastaminen on kokonaisuudessaan toteutettu koneellisesti P2P:n sisällä. Tilauksen tekijän ei enää tarvitse tutkia laskumateriaaleja ja verrata tietoja puitesopimuksiin manuaalisesti. Välivaiheissa voi olla, että tarvitaan ihminen tekemään tarkistuksia epäselvissä tapauksissa.



# 5. Kokeilun tuotokset

## Laskun optimaalinen tietosisältö (Word-tiedosto)

Laadittiin lista laskun optimaalisesta tietosisällöstä eli niistä tiedoista, jotka Helsingin kaupunki tarvitsee sujuvaa laskunkäsittelyn prosessia varten. Listatut tiedot toimitetaan tiedoksi toimittajille sekä osaksi uusia puitesopimuksia.

Dokumentista löytyy tiedon nimet ja usein selite, mahdollisesti esimerkki ja muita huomioita.

Viimeisenä listalla oleva “huoltoraportti” on laskun liitteenä vaadittu erillinen dokumentti tehdystä huoltotyöstä. Tämä viedään kohteen huoltokirjalle.

Tilat-palvelun kiinteistöautomaation laskutus:

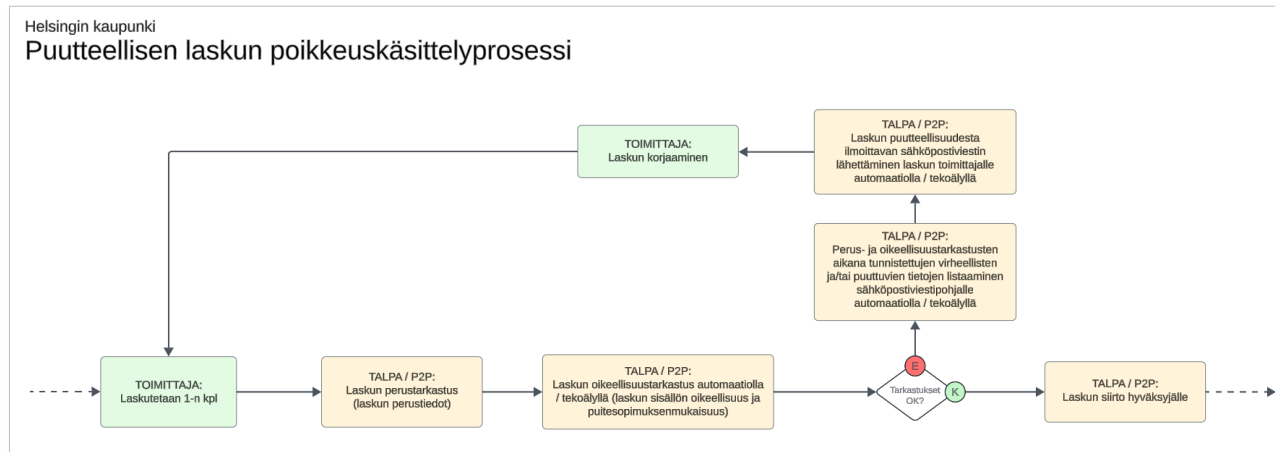
### Laskun optimaalinen tietosisältö

- Laskun antamispäivä = päivä, jolloin lasku on laadittu
- Juokseva tunniste = laskunumero
- Laskuttajan ALV-tunniste (y-tunnus)
- Laskuttajan nimi
- Laskuttajan osoite
- Laskun saajan nimi = esim. "KYMP/RYA/TILAT-PALVELU"; kuten mainittu sopimuksella
- Laskun saajan osoite = paikka, mihin ostolaskut pitää toimittaa
- Tilaaajan yhteyshenkilön nimi = kuka tilauksen on tehnyt; kuka on hyväksynyt, että toimittaja tulee tekemään tilatun työn; "tilaaajan viite"
- ALV-kanta
- Maksuehto
- Veron peruste verokannoittain ja yksikköhinta ilman veroa
- Veroton hinta ja arvonlisäveron osuus euromääräisenä
- Puitesopimuksen tunniste (Clouidiassa ja intrassa) tai erillisen tilauksen tunniste
- Alennukset = puitesopimuksella sovitut alennukset
- Hyvitykset = esim. väärrälle toimialalle osoitetut laskut hyvitetään
- Työkohteen nimi = korjauksen kohteen nimi, "työmaa"
- Mitä tilattu, mikä on ongelma, mitä halutaan tehtävän
- Tarvittaessa rakennuksen nimi = esim. Auran sairaala-alueella on 30 rakennusta, tärkeää tietää mihin rakennukseen korjaus kohdistuu
- Kiinteistöobjekti = nelinumeroinen kiinteistön tunniste
- Työkohteen osoite = kiinteistön virallinen osoite
- Tilausnumero = juokseva tilauksen tunnus, Talpa, samaa tilausta voidaan laskuttaa usealla laskulla
- Tilauspäivämäärä = milloin toimittajalle on soitettu ja työ tilattu
- Kuvaus töistä = vapaa kuvaus työstä; haluttaisiin, että sisältää pvm, tehty työ rivitasolla, kuka tilannut, eur (sovittu tuntihinta on mainittu puitesopimuksella per työlaji)
- Työntekopvm = päivä, jolloin työ on toteutettu
- Työn tekemiseen käytetyt tunnit
- Tarvikkeiden ja materiaalien ostokuitit tarvittaessa
- Tarvikkeiden ja materiaalien voimassa olevan hinnaston tunniste
- Huoltoraportti: pitää olla, mutta sisältöä ei ole tarkasti määritetty, osa kelvottomia

# 5. Kokeilun tuotokset

**Kuvaus puutteellisen laskun poikkeuskäsittelystä** (työstetty Lucidchartissa, lopputuotos PDF-tiedosto)

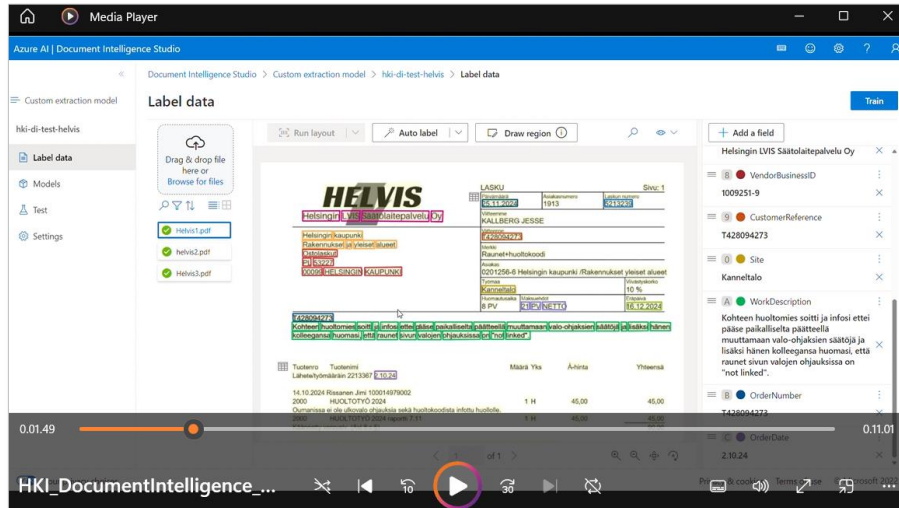
Kuvattiin tavoitetilaan liittyvä laskun poikkeuskäsittelyn prosessin osa tarkemmin. Tässä skenaariossa P2P tarkastaa laskun perustiedot ensin ja sitten automaation tai tekoälyn avulla jatkaa tarkastamaan myös laskun sisällön oikeellisuuden ja puitesopimuksenmukaisuuden. Mikäli jompikumpi tai molemmat tarkastuksista havaitsevat puutteita, lähetetään toimittajalle automaattinen sähköpostiviesti, joka koostaa sisälleen tiedot puutteista ja virheistä sekä pyytää toimittajaa korjaamaan laskun sisällön. Laskun saapuessa korjattuna viedään se saman tarkistuksen läpi.



# 5. Kokeilun tuotokset

## Azure AI Document Intelligence -demo (video)

Kokeiltiin Azure AI Document Intelligence -tekoälytyökalua laskuilta löytyvien tietojen tunnistamiseen ja keräämiseen rakenteelliseksi tiedoksi JSON-tiedostoon. Videolta näkyy, kuinka tekoälymalli on ensin manuaalisesti luotu ja kuinka se sen jälkeen osaa melko hyvällä varmuudella itsenäisesti etsiä PDF-muotoisilta laskuilta tietoja. Odotusten mukaisesti tekoäly ei pystynyt luomaan laskulta puuttuvia tietoja, vaan ne tulee saada suoraan toimittajalta.



# 6. Opit kokeiltavan ratkaisun tai toimintatavan mahdollisuuksista

Kokeilussa analysoitiin nykyprosessi sekä siihen liittyvää laskumateriaalia ja sen sisältämää rakenteellista ja manuaalisesti (silmin) havainnoitavaa tietoa. Kokeilussa käytetyt päämenetelmät olivat siis toiminta- ja data-analyysi, ja ne toimivat tarkoitukseen hyvin.

Prosessia kuvaamalla ja analysoimalla voitiin todeta, että toiminta nykyisellään sallii toimittajien toimittaa hyvin erilaisia ja erisisältöisiä laskuja Helsingille käsittelyyn, ja tämä on haasteellista. Edellä mainittu todennettiin käytännössä laskumateriaalin analyysillä: eri toimittajien laskut poikkeavat toisistaan huomattavasti sekä muodoltaan että sisällöltään. Myös liitteitä (huoltokirjat) puuttuu, vaikka ne ovat Tilat-palvelun toiminnalle tärkeitä dokumentteja.

Kokeilun aikana tuli myös ilmi, että tavalla tai toisella puutteellisia laskuja vastaanotetaan ympäri kaupungin toimintoja, ei pelkästään Tilat-palveluissa. Näkemyksemme mukaan vastaavaa vaatimusten tarkentamista toimittajien laskuihin sekä sen jälkeistä automatisointia kannattaisi miettiä myös muissa toiminnoissa.

# 6. Opit kokeiltavan ratkaisun tai toimintatavan mahdollisuuksista

Azure AI Document Intelligence -työkalu osoitti käyttötapauksessa kyvykkyytensä, mutta myös rajoitukset tilanteessa, jossa tietoa ei ole olemassa.

Laskujen tietojen puutteellisuuden osalta saatiin vahvistus oletuksille. Tekoälyn ja sääntöpohjaisen automaation mahdollisuuksista saatiin uutta ymmärrystä ja näkemys, kuinka prosessi voitaisiin automatisoida, kunhan tarvittavat tiedot on saatu toimittajalta.

Kuten usein, ongelman korjaaminen prosessin loppuvaiheessa ei ole tehokasta tai edes mahdollista, vaan ongelma pitää pyrkiä poistamaan mahdollisimman aikaisin. Tässä prosessissa oikea vaihe on sopimusten ja vaatimusten muokkaaminen vastamaan Helsingin kaupungin Tilat-palvelun kiinteistöautomatisaation tarpeita.

# 7. Opit asiakkaiden tai palvelun käyttäjien tarpeista

Kokeilun aikana oli tunnistettavissa, että laskun optimaalinen tietosisältö (eli ts. tarvittavien tietojen lista) on melko laaja ja kattaa Tilat-palvelun ohella myös Talpan tarpeita. Huomioitavaa on, että mikäli laskujen tarkastuksen automaatiota haluttaisiin kokeilla kaupungilla laajemmin, olisi kriittistä kartoittaa tietotarpeet myös laajemmin, aina osallistuvien tahojen toimintaan ja siihen vaikuttaviin tekijöihin (esim. raportoinnin vaatimuksiin tai Suomen lakiin) heijastaen.

Kokeilulla pyrittiin saamaan ymmärrystä siitä, millä tavalla nykyprosessia voisi parantaa ja manuaalisen työn määrää vähentää tekoälyavusteisella automaatiolla. Tehdyn työn ansiosta voitiin todeta, että laskun tietojen tarkastusta voidaan automatisoida huomattavasti, mutta toimivaa ratkaisua varten laskun tietosisältö ja formaatti tulisi normalisoida\* ja kaiken Helsingille tulevan laskutiedon pitäisi olla koneluettavassa muodossa.

\*Normalisoinnilla tarkoitetaan tässä yhteydessä sitä, että laskujen tietosisältö ja formaatti yhtenäistetään sovitun standardin mukaiseksi. Tämä tarkoittaa, että kaikki laskut sisältävät samat tiedot samoilla otsikoilla, samassa järjestyksessä ja samassa muodossa. Normalisointi helpottaa automaattista tietojen tarkastusta.



# 8. Opit ratkaisun kehittämistä teknisesti

Osana kokeilua analysoitiin esimerkkilaskujen sisällöt sekä laskujen kuvien että XML-tiedostojen osalta.

- XML-muotoiset tiedostot analysoitiin manuaalisesti CGI:n konsultin toimesta. XML-laskuja oli seitsemältä toimittajalta ja ne käytiin läpi verraten niitä sekä tavoitetilan listaan että keskenään tietosisältöjen ja rakenteen osalta.
- Laskujen kuvien osalta tekoälykokeilussa välineenä käytettiin Azure AI Document Intelligence -työkalua, jossa luotiin tekoälymalli perustuen kolmeen Helviksen laskuun. Tämän jälkeen luodun mallin toimivuutta testattiin seitsemällä saman toimittajan eri laskulla (eli laskuilla, joita tekoäly ei ollut ennen nähnyt). Tavoitteena oli todentaa, kuinka hyvin tekoäly osasi tunnistaa laskuilta tietoja. Tuotanto-toteutuksissa laskuja opetettaisiin useita kappaleita ja useilta toimittajilta. Lähtökohtana tulisi kuitenkin olla laskujen saaminen yhdenmukaisessa, koneluettavassa muodossa, jolloin laskujen kuvia käytettäisiin vain ihmisiä varten poikkeustilanteissa.

# 8. Opit ratkaisun kehittämistä teknisesti

**Yleiset havainnot:** Laskut ovat hyvin monessa eri formaatissa/layoutissa riippuen toimittajasta. Sarakkeet vaihtelevat toimittajien välillä merkittävästi, mikä vaikeuttaa yhtenäistä käsittelyä.

## XML-tiedostojen manuaalisen analysoinnin havainnot

- XML-tiedostoissa oli kattavasti samat tiedot kuin laskujen PDF-tiedostoissa.
- Merkittävimmät puutteet:
  - Taulukkomuotoisten rivitietojen otsikot puuttuivat, mikä estää automaation käytön rivitietojen käsittelyssä.
  - Lisätiedot: Osa laskuista sisälsi tarkastukseen vaadittavia lisätietoja vapaissa tekstikentissä, mutta osa laskuista ei sisältänyt näitä tietoja lainkaan.
- XML-muodossa ei aina hyödynnetä kaikkia XML-tägejä.
  - esim. Toimittaja 1:n laskuissa ei käytetä ORDER\_DATE -tägiä, kun taas esim. Toimittaja 2 hyödyntää sitä.
- Puutteellisimmat laskut (XML:n ohjelmallisen hyödyntämisen kannalta) ovat Toimittaja 3:n.
  - niissä on vain yksi summarivi: <ARTICLE\_ID>HUOLTOTYÖ</ARTICLE\_ID> ja <ARTICLE\_NAME>Työjono: FDXI-1165098</ARTICLE\_NAME>, ja tämä "Työjono: FDXI-1165098" on sitten liitteenä eritelty PDF:ssä myöhemmin (ja erittelystä löytyy työtä ja tavaraa), mutta erittelyä ei ole XML:ssä.  
→ Tämä malli estää käytännössä automaation hyödyntämisen tarkistuksessa tai vaatii toimittajakohtaisia räätälöintejä automaation toimintaan.

# 8. Opit ratkaisun kehittämisestä teknisesti: Azure AI Document Intelligence

## Tekoälykokeilun toteutus

- Kokeiltu väline oli Azure AI Document Intelligence -tekoälytyökalu, joka tunnistaa PDF-muotoisia laskuja ja niiden tietosisältöjä.
- Koneoppimismalli opetettiin tunnistamaan laskuja käyttämällä kolmea laskua mallin opettamiseen, ja seitsemällä laskulla testattiin mallin toimivuutta. Tuotannossa käytettäisiin useampia toimittajia ja laskupohjia opetukseen.
- Opettamisessa käytettiin myös manuaalista merkintää halutuille kentille, jossa käyttäjä valitsee tärkeät kentät, kuten eräpäivän ja laskun päivämäärän.

## Tekoälykokeilun havainnot

- Tunnistustarkkuus: Malli pystyi automaattisesti tunnistamaan tilauspäivämäärän jokaisesta laskusta, mutta vapaamuotoisen työnkuvauksen tunnistamisessa oli haasteita.
- Laskun perustietojen osalta mallin ns. confidence-taso, joka ilmaisee tunnistuksen varmuuden, oli hyvin korkea. Vapaamuotoisten tekstien osalta confidence oli matalampi.
- Tunnistetut tiedot saadaan ulos JSON-muodossa, jota on erittäin helppo hyödyntää ohjelmallisessa käsittelyssä.

## Tekoälykokeilun perusteella tehdyt suositukset

- Mikäli tietoja ei saataisi XML-muodossa, mutta ne olisivat laskun kuvassa, Document Intelligence on erittäin tehokas työkalu tietojen poimintaan tekoälyn avulla. Työkalu toimii mallikkaasti myös esimerkiksi laskujen liitteiden lukemisessa ja niiden tyyppin tunnistamisessa.
- Koska tekoälykään ei voi poimia tietoja, joita laskuissa ei ole, suositus on saada toimittajilta määrämukoista dataa laskuille nimettyihin vakiokenttiin, jolloin ne ovat suoraan XML-tiedostosta saatavilla automaattiseen käsittelyyn ja tarkistukseen.

# 8. Opit ratkaisun kehittämisestä teknisesti

## Tekniset suositukset

- 1) Toimittajan pitää hyödyntää kaikki käytettävissä olevat XML-tägit (kuten esim. ORDER\_DATE), jotta laskujen oikeellisuutta voidaan helposti tarkastaa koneellisesti.
- 2) Jos tarkoitukseen ei ole sopivaa XML-tägiä, esim. FREE\_TEXT -tägin alla tulee olla tarvittava informaatio määrämuotoisena puitesopimuksessa kuvatussa formaatissa: esim. "kiinteistöobjekti\_numero:1234", jolloin tämä tieto on helposti purettavissa automaation käyttöön
- 3) Toimittaja 3:n käyttämää "summausta" ei saisi käyttää, vaan tavarat/tuotteet ja työ pitää löytyä eriteltynä XML-tiedostosta.
  - 1) Summauksessa tiedot on yhdistetty vain koosteeksi, eikä yksittäisiä töitä ole eroteltu itse laskulla vaan erillisessä liitteessä.

# 9. Opit kokeilemisesta yleensä

## Mikä toimi hyvin?

Yhteistyö, asiantuntijoiden tunnistaminen ja saaminen työpajoihin, käytetyt kommunikointi tavat ja välineet.

## Kehitysideat

Teknisissä kokeiluissa tarvittavat alustat tulisi tunnistaa ja tilata mahdollisimman aikaisessa vaiheessa, jotta niiden pystyttämiseen jäisi mahdollisimman paljon aikaa ennen varsinaista tarvetta.

# 10. Opit kokeiluprojektin arjen pyörittämisestä

Kokeilu vei kalenteriaikaa noin kolme kuukautta, ja työtä tehtiin CGI:n konsulttien puolelta reilut 10 henkilötyöpäivää. Budjetti riitti tämäntyypiseen kokeiluun hyvin.

Suurin osa työajasta käytettiin tuotosten laatimiseen (työpajat + konsultin tekemä työ) ja teknisen kokeilun toteuttamiseen Azure AI Document Intelligencellä. Teknisen ympäristön ja työkalujen ympärillä käydyt keskustelut veivät hieman enemmän aikaa kuin alunperin oletettiin. Copilotille suunniteltu osuus kokeilusta jätettiin pois osittain siksi, että sitä ei saatu riittävän vikkelästi käyttöön, eikä sitä toisaalta ollut tarkoitus käyttää tuotantoa vastaavaan käyttötarkoitukseen vaan avustamaan datan analysoinnissa.

Tiimin yhteistyö sujui mallikkaasti Teamsin välityksellä.

# 11. Kokeilun tekninen ympäristö

## Kokeilun toteutusympäristö

- Azure Blob Storage (laskujen kuvat/pdf-tiedostot)
- Azure AI Document Intelligence (Studio) -työkalu mallin koulutukseen ja testaamiseen testidatalla. Tietojen tunnistukseen käytettiin Custom-mallia.

## Mitä teknisessä ympäristössä olisi hyvä ottaa huomioon, jos kokeilu tehtäisiin nyt uudelleen?

Pyrittäisiin tunnistamaan tarve ja tekemään tilaus tarvittaville komponenteille aiemmin, jotta voitaisiin tiivistää kalenteriaikaa. Toisaalta vaadittiin riittävä määrä työpajoja teknisen kokeilun lopullisen rajauksen tekemiseen.

# 12. Kokeilun data

## Mitä kokeilun datasta ja sen käsittelystä opittiin?

Dataa saatiin joustavasti, ja vaikka lukumääräisesti tiedostoja oli varsin pieni määrä, niistä saatiin todennettua kokeilun oletukset ja tehtyä tekniset työt.

Laskudata oli helppoa, koska valitut laskut eivät sisältäneet sensitiivistä dataa. Mikäli käsiteltäisiin aineistoja esim. terveystietojen osalta, olisi tiedon sanitointiin ja käsittelyyn jouduttu käyttämään merkittävästi enemmän resursseja. Tämä vaikuttaisi myös osaltaan tuotantokäytön käytäntöihin ja ratkaisuihin.



# 13. Jatkopäätökset ja -ideat

**Millaisia päätöksiä jatkosta on tehty ja millä perustein? (“Jatketaan” / “muutetaan” / “päätetään”)**

Tilat-palveluissa lähdetään työstämään vaatimuksia ja sopimusmalleja tuleviin kilpailutuksiin, jotta datan laatu saadaan ajan myötä paremmaksi.

**Mitä tulisi huomioida jos kokeilun pohjalta lähdetään rakentamaan tuotantokelpoista ratkaisua?**

Tuotantokäyttöä ajatellen ratkaisu todennäköisesti perustuisi pääosin sääntöpohjaiseen automaatioon, jota tuettaisiin tarvittavissa kohdissa teköälypalveluilla.

**Kokeilun skaalaamisen potentiaali ja alustava laskelma hyödyistä**

Hyötylaskelman osalta tulisi kartoittaa laajemmin muilla toimialoilla ja toiminnoissa tehtävää laskujen käsittelyä, tarkistusta ja palautuksia. Siirtymällä määrämuotoisiin laskudatoihin myös sisällön osalta päästäisiin näkemyksemme mukaan merkittäviin säästöihin läpi organisaation.