

# Tekoälykokeilun loppuraportti

Linked Events -tapahtumarajapinnan  
asiasanoituksen tekoälykokeilu

Maaliskuu 2020

Helsinki



Tukea digitalisaatiokokeiluihin  
kaupungin työntekijöille

# Linked Events -tapahtumarajapinnan asiasanoituksen tekoälykokeilu

Alexi Salonen + Riku Oja (Helsingin kaupunki)  
Jussi Martikka ja Noora Söderling (SAS)

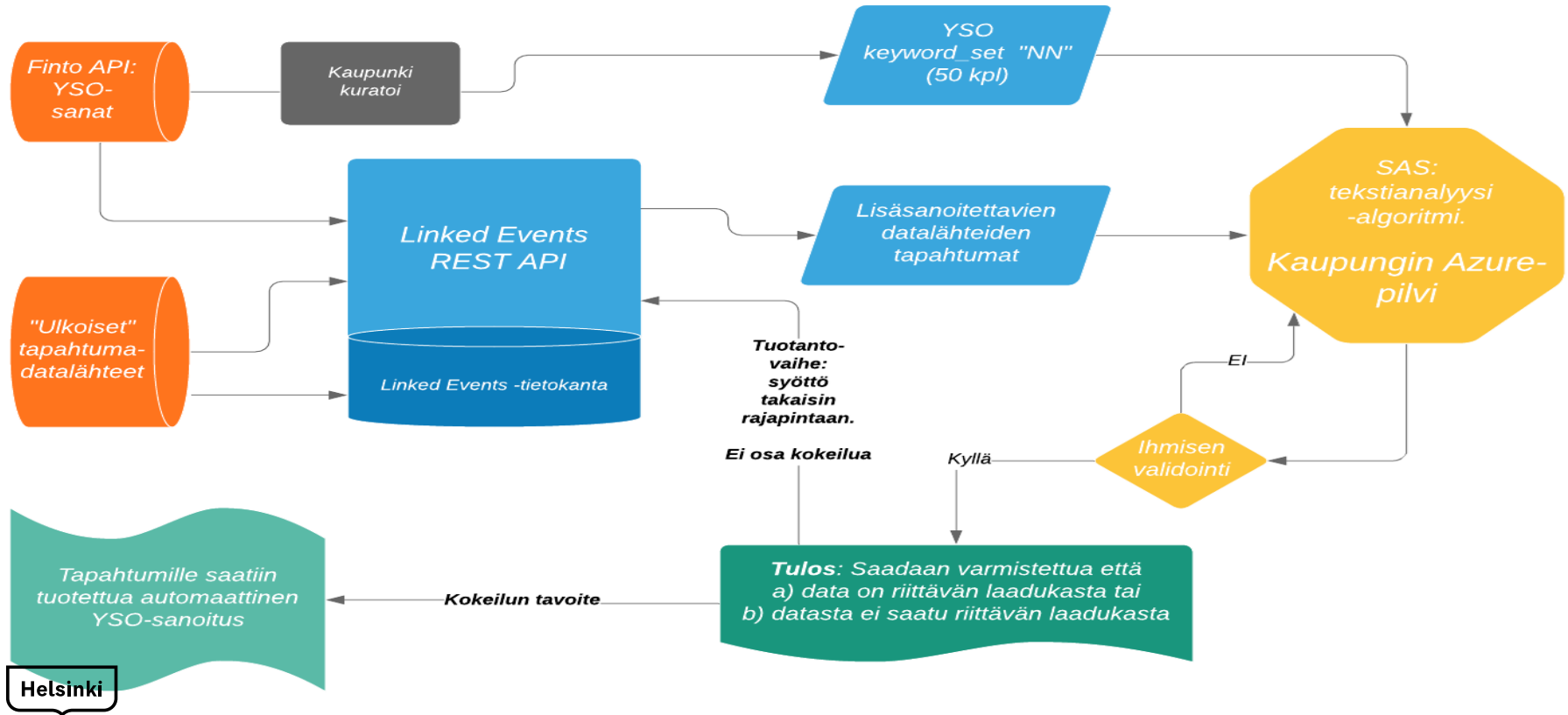
# 1. Kokeilun onnistuminen

- Kokeilun perustavoitteessa onnistuttiin: Osoitettiin että tekstianalyysin keinoin voidaan tuottaa riittävän laadukasta automaattista asiasanaluokittelua tapahtumadatalle.
- Tavoitteeksi ei asetettu tuotantoon valmista ratkaisua ja siihen pääseminen vaatisi teknistä jatkokehitystä.
- Kokeilu osoitti, että tekstianalyysipohjainen luokittelu on järkevä kehityssuunta Linked Eventsille.
- Tuotantoratkaisu on kehitettävä erikseen, mutta kokeilussa saatiin luotua selvät askelmerkit tulevalle ratkaisulle.
- Luokittelun onnistuminen antaa hyvän viitteen siitä, että vastaavaa automaattista luokittelua kannattaisi hyödyntää myös kaupungin muiden tekstiaineistojen luokitteluun.

# 2. Kokeilun eteneminen

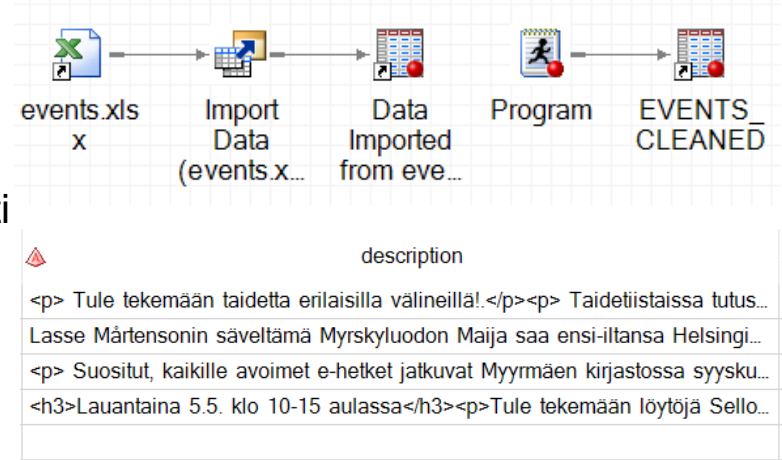
- Kokeilussa testattiin tapahtumadatan automaattista asiasanoitusta, käyttäen SAS:n Viya-tekstianalyysi- ja koneoppimistyyökalua.
- Kokeilun vaiheet pääpiirteittäin:
  - Ratkaistavan ongelman määrittely ja kokeilun rajaaminen
  - Tapahtumadatan ja asiasanasetin valinta (50 YSO-asiasanaa ja 9 YSO-kohderyhmää)
  - Datan siirtäminen analyysiympäristöön (SAS Viya)
  - Datan valmistelu
  - Taulujen yhdistäminen ja mallintamiselle olennaisten muuttujien luonti
  - Datan tutkiminen
  - Sääntöpohjaisen mallin luonti
  - Mallin rikastaminen säännöillä luokitteluiden tarkentamiseksi
  - Mallin kouluttaminen hyödyntäen tekoälyä
  - Välituloksien läpikäynti ja iterointi
  - Tulosten visualisointi ja analyysien tuottaminen
  - Tulosten tarkastelu ja jatkokehityksen ideointi

# 2a. Kokeilun eteneminen: lähtökohta



# 2b. Datan valmistelu

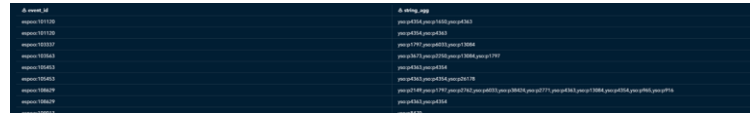
- Ensimmäinen askel oli datan valmistelu
- Rajapintadumppina haettu CSV-tiedosto muutettiin xlsx-muotoiseksi, jotta utf-8-ongelmat (ääkkösten muuttuminen sekasotkuksi) saatiin selätettyä vaivattomasti
- Xlsx-tiedosto importattiin SASiin, jossa datasta poistettiin ohjelmallisesti turhat html-tägit ym. ylimääräinen, sekä poistettiin tyhjä rivit sekä duplikaatit
- Datasta lisäksi tiputettiin turhat sarakkeet, jolloin jäljelle jäivät analyysissä keskeisimmät kolumnit eli: tapahtuman **yksilöivä id**, nimi, sekä suomenkielinen kuvaus.



id	name_fi	description_fi
helmet:139...	Enter opasta...	Tapulikaupungin kirjastossa järj...
helmet:115...	Enter ry:n tie...	Haluatko apua ja neuvoja tietok...
matko:846...	Allas-konsertit	Kauppatorin kupeessa sijaitsev...
helmet:115...	Vauvojen lor...	Kuka saa, kuka saa lorupussiin...

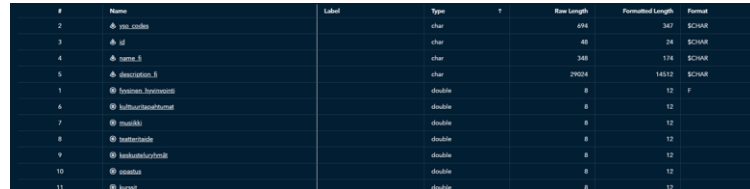
# 2b. Datan valmistelu

- Mallin kouluttamisessa hyödynnettiin historiadataa, joka vaati tapahtumien aiempien luokitteluiden yhdistämistä dataan, joka sisältää tapahtuman yksilöivän id:n, nimen ja kuvauksen
- YSO-id:eillä merkityt luokittelut eroteltiin omiin sarakkeisiin ja käännettiin YSO-koodia vastaavaan sanaan
- Luokittelusanoille annettiin arvo 0 tai 1, riippuen siitä oliko tapahtuma luokiteltu käyttäen kutakin YSO-luokkaa



id	event_desc
100001	...
100002	...
100003	...
100004	...
100005	...
100006	...
100007	...
100008	...
100009	...
100010	...

```
1 // BEGIN data step with the output table data %  
2 data (_sq_outputTable) (calllib=(_sq_outputCallib)) promote="no";  
3 /* Set the input set */  
4 set (_sq_inputTable) (calllib=(_sq_inputCallib));  
5 if index(yso_codes, "yso:p000") > 0 then kulttuuritapahtumat = 1;  
6 if index(yso_codes, "yso:p1000") > 0 then musiikki = 1;  
7 if index(yso_codes, "yso:p2000") > 0 then teatteritaido = 1;  
8 if index(yso_codes, "yso:p3000") > 0 then keskusteluryhmat = 1;  
9 if index(yso_codes, "yso:p4000") > 0 then opetus = 1;  
10 if index(yso_codes, "yso:p5000") > 0 then kurssit = 1;  
11 if index(yso_codes, "yso:p6000") > 0 then satutunnit = 1;  
12 if index(yso_codes, "yso:p7000") > 0 then kieli_ja_kielet = 1;  
13 if index(yso_codes, "yso:p8000") > 0 then kirjallisuus = 1;  
14 if index(yso_codes, "yso:p9000") > 0 then leikkiminen = 1;  
15 if index(yso_codes, "yso:p10000") > 0 then konsertit = 1;  
16 if index(yso_codes, "yso:p11000") > 0 then ...
```



Name	Label	Type	Row Length	Formatted Length	Format
yso_codes		char	494	247	SCHAR
id		char	40	24	SCHAR
name_5		char	340	174	SCHAR
description_5		char	2904	1452	SCHAR
event_name		double	8	12	F
event_desc		double	8	12	F
musiikki		double	8	12	F
teatteritaido		double	8	12	F
keskusteluryhmat		double	8	12	F
opetus		double	8	12	F
kurssit		double	8	12	F

# 2c. Mallien luominen

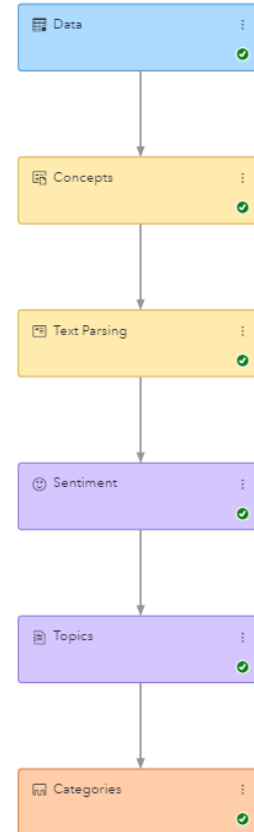
- SAS Viya:ssa on visuaalinen näkymä mallien luomiseen.
- **Malli voidaan luoda kahdella tavalla:**

## 1) Hyödynnetään tapahtumiin ihmisten tekemiä luokitteluja

- Datan perusteella malli voitaisiin opettaa luokittelemaan samalla tavalla kuin ihminen on tapahtumia aiemmin luokitellut

## 2) Sääntöpohjainen, manuaalisesti tehty luokittelu

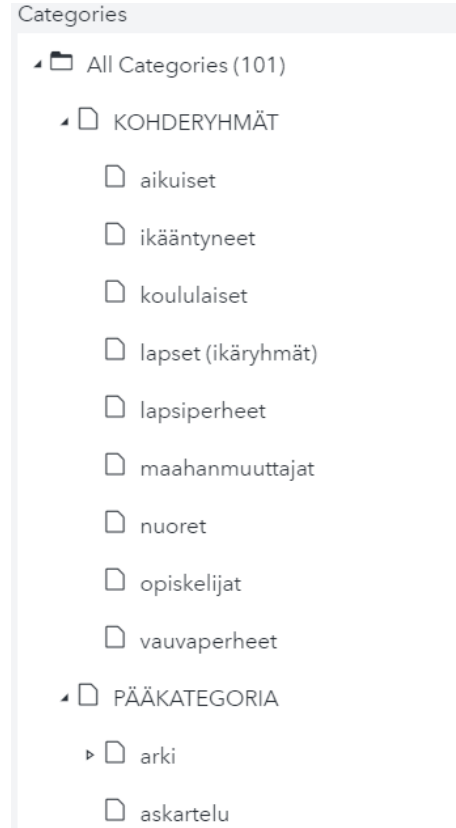
- Säännöt tehtiin käyttämällä apuna Finto.fi -sivuston YSO-sanojen määrittelyjä (sanojen synonyymit)





## 2d. Sääntöjen luominen

- Viyassa **sääntöjä** voidaan luoda joko konsepti- tai kategoriapohjaisesti. Sääntöjä voi myös yhdistää, jolloin kategoriaosuudesta saa siistimmän:
  - Concepts-osioon saa tehtyä valmiiden konseptien lisäksi omia konsepteja, joita kone osaa tunnistaa (paikat, ihmiset, etäisyys tiettyjen sanojen välillä jne.)
  - Yhdistely Categories-osion kanssa: esim. “kategoriaosumaa ei oteta mukaan tuloksiin, jos tietty paikka on mainittu”
- Kategoriat luodaan määrittämällä sanoja, joita **halutaan** tai **ei haluta** tunnistaa tuloksiin

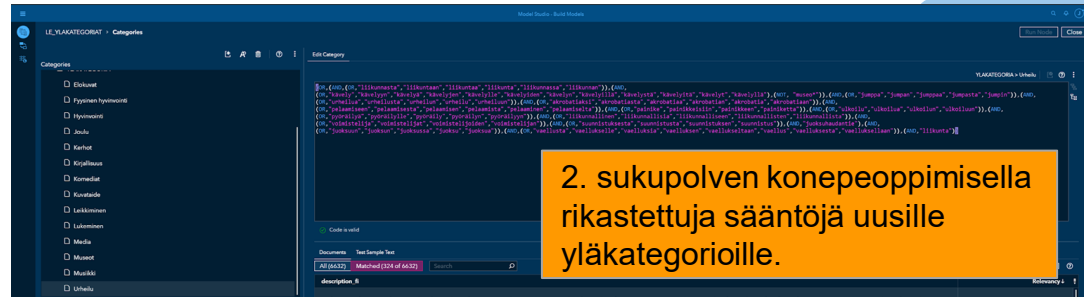
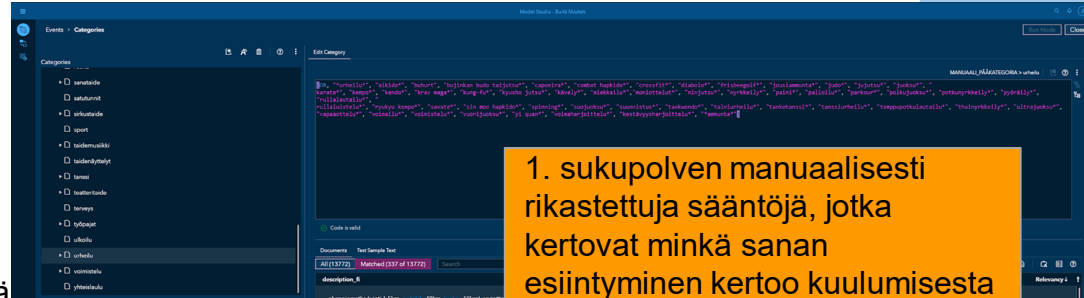


## 2e. Tekoälyn kouluttaminen: Sopivan mallin valinta

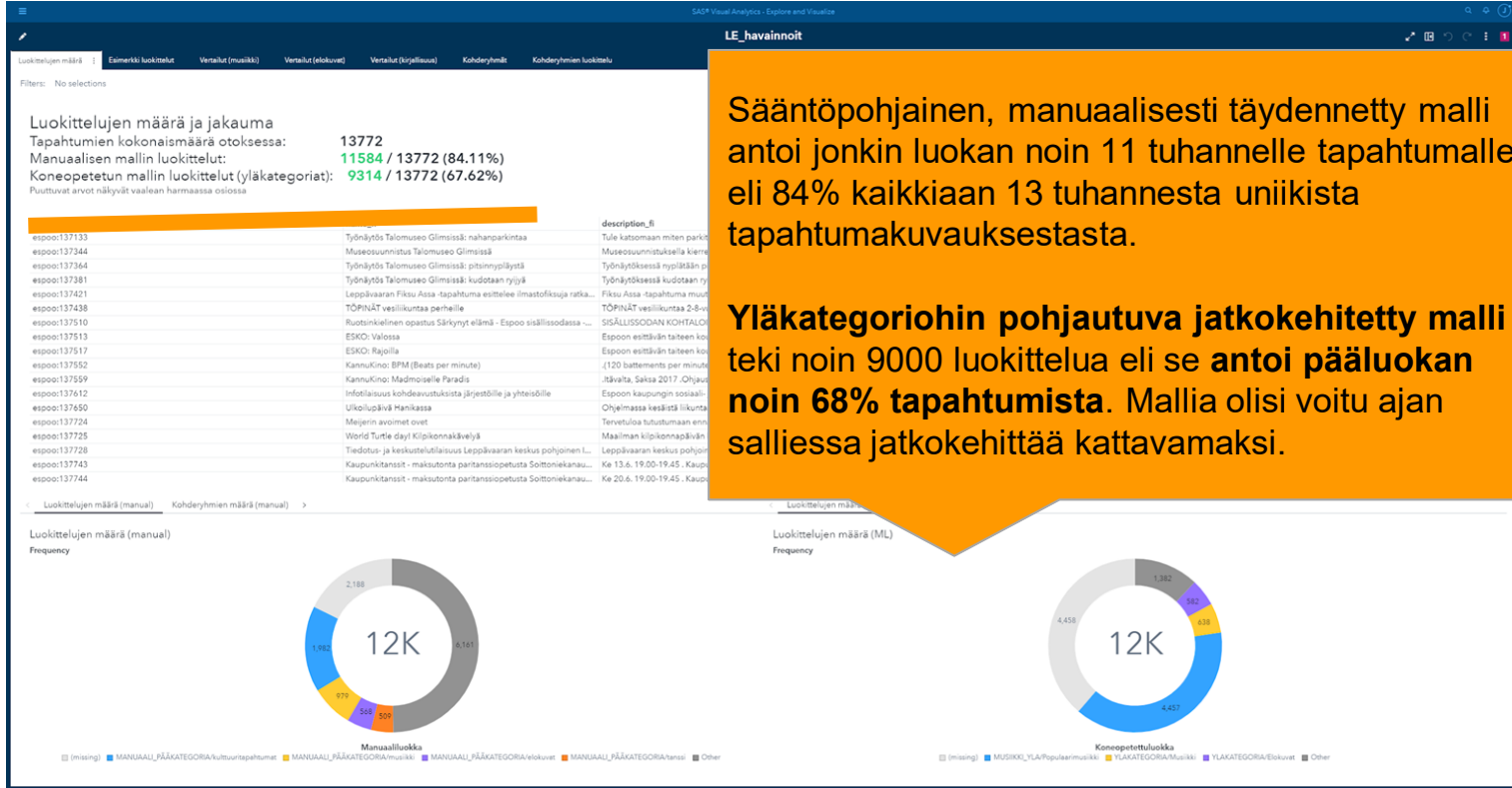
- Malli voidaan luoda antamalla algoritmille joko ihmisten tekemiä aiempia luokitteluja tai tekemällä joukko sääntöjä
  - Todettiin, että säännöt antoivat tarkempia tuloksia, koska opetusdatassa oli joitakin puutteita. Luokittelun vaihteleva laatu oli tärkeimpiä syitä kokeilun aloittamiselle, joten tämä ei ollut yllätys.
  - Mallia lähdettiin kouluttamaan / kehittämään sääntöpohjaisesti
- Koska kokeiluun valittu 50 historiallisesti käytetyimmän YSO-sanan luokittelu ei ollut looginen (päällekkäisyyksiä, eritasoisuuksia yms.), tehtiin luokille 17kpl joukko kokoavia yläluokkia, jotta tekoäly ymmärtäisi ja oppisi luokkia paremmin.

## 2f. Tekoälyn kouluttaminen: Mallin hiominen

- Mallinnuksessa käytimme kahta erilaista lähestymistapaa:
  - Historiatadan hyödyntäminen koneopetuksessa
  - **Sääntöpohjaisen mallin rikastaminen** manuaalisesti (kuva 1)
- Historiatatalla tehdyssä koneopetuksessa luokitteluissa esiintyi paljon päällekkäisyyksiä ja virheellisiä luokitteluja
- Päällekkäisten luokittelujen poistamiseksi ja tarkemman pääkategorian valitsemiseksi loimme **yläkatogorioita** (kuva 2), jolloin mallin koneopettaminen generoi vaikuttavia **sääntöjä automaattisesti** (kuva 3)
- Koneopetettua mallia voi rikastaa vielä validoimalla opittuja sääntöjä itse, luomalla lisäsääntöjä ja kouluttamalla mallia niiden pohjalta lisää → Tulosten paraneminen



# 3. Kokeilun tulokset: kokonaiskuva



# 3. Kokeilun tulokset: esimerkkivertailu 1

- Esimerkki kohderyhmäluokittelusta, jossa on vertailtu alkuperäistä luokittelua ja koneopetetun mallin tuottamaa luokittelua.

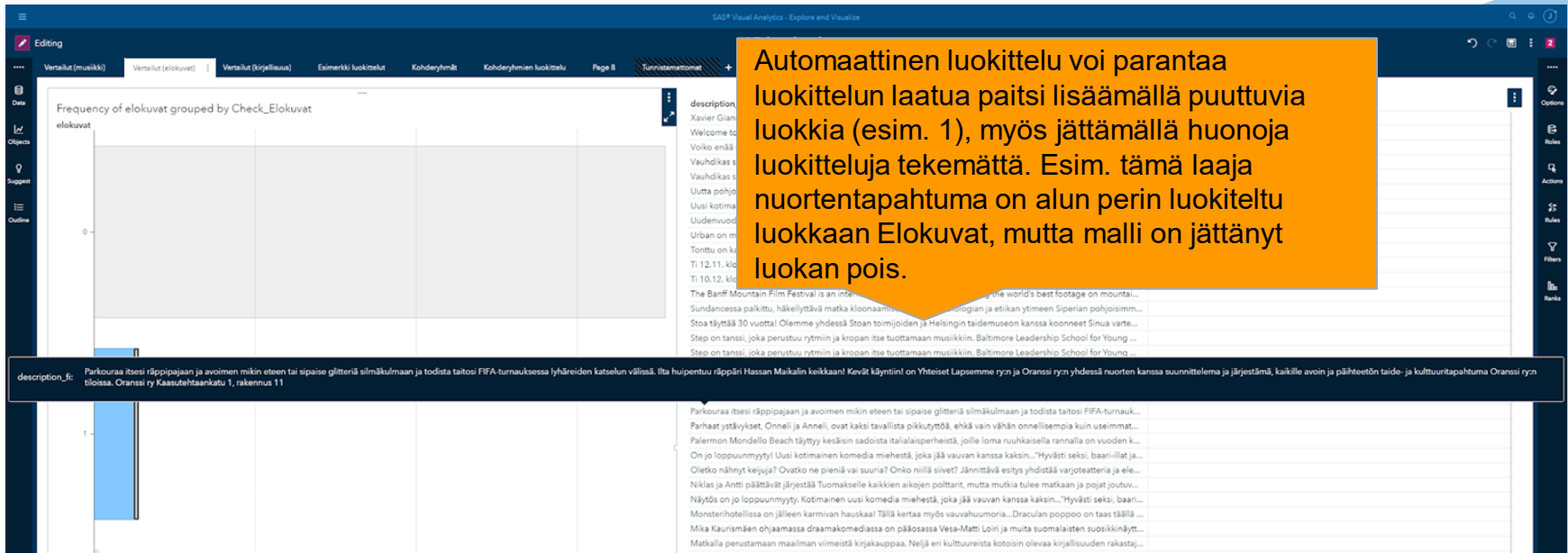
Sääntöpohjainen, synonyymeillä täydennetty ja koneoppimiseen pohjaava malli päätyi antamaan tapahtumalle kohderyhmäluokan, vaikka se puuttui alkuperäisestä datasta.

The screenshot displays a table with two columns: 'description\_fi' (original data) and 'ML/...' (machine-generated classifications). The 'description\_fi' column contains event descriptions such as 'Bailla Bailla Summer Festival 2018', 'International Clown Festival - Clown Land 2017', and 'Kotouttamisen haasteet ja työpäivät - KOTO 3 - hankkeen loppuseminaari'. The 'ML/...' column shows classifications like 'ML/Aikuiset', 'KR\_alkuiset (1)', 'ML/ikäkääntyneet', 'KR\_ikäkääntyneet (1)', 'ML/Kouluikäiset', 'KR\_kouluikäiset (1)', 'ML/Lapset (ikäryhmät)', 'KR\_lapset\_ikäryhmat (1)', 'ML/Maahanmuuttajat', 'KR\_maahanmuuttajat (1)', 'ML/Nuoret', 'KR\_nuoret (1)', and 'ML/Opiskelijat (1)', 'KR\_opiskelijat (1)'. A red arrow points to the 'ML/Maahanmuuttajat' classification for the 'Kotouttamisen haasteet...' event, which was not present in the original data.

description_fi	ML/...
Kaikkie avoimissa Walk -illoissa jaetaan asumiseen liittyvää tietoa, tutustutaan vapaaehtoistoimintaan sekä tavataan maahanmuuttajia ja toisia vapaaehtoisia. Vapaaehtoisille on tarjolla monipuolisia tehtäviä maahanmuuttajien asumisen tukena: voit tulla mukaan vaikkapa asuntonäyttökäynnille, näin voidaan tarjota työpaj...	ML/Aikuiset KR_alkuiset (1) ML/ikäkääntyneet KR_ikäkääntyneet (1) ML/Kouluikäiset KR_kouluikäiset (1) ML/Lapset (ikäryhmät) KR_lapset_ikäryhmat (1) ML/Maahanmuuttajat KR_maahanmuuttajat (1) ML/Nuoret KR_nuoret (1) ML/Opiskelijat (1) KR_opiskelijat (1)
Bailla Bailla Summer Festival 2018	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
International Clown Festival - Clown Land 2017 - Työpaja turvapaikanhakijoille	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Jan Uuspöld: Brüsseli kapsas - Brysselin kaali	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Kotouttamisen haasteet ja työpäivät - KOTO 3 - hankkeen loppuseminaari	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Maahanmuuttajan poika	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Monikielinen lukupiiri "Avattu kirja"	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Peruuntunut 3.12. Vapaaehtoistoiminta Walk-in-illat	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
Saman taivaan alla - hyväntekeväisyyskonsertti	ML/Maahanmuuttajat KR_maahanmuuttajat (1)
SuomiKamu-koulutus Espoossa	ML/Maahanmuuttajat KR_maahanmuuttajat (1)

# 3. Kokeilun tulokset: Vertailu 2

- Linked Eventsin aiempaa luokittelua verrattiin mallien kanssa, josta esimerkkinä havainto virheellisestä luokittelusta, jonka malli on osannut jättää pois



# 3. Kokeilun tulokset: vertailu 3

- Esimerkki tapahtuman luokittelusta, jossa vertaillaan sääntöpohjaisen mallin, koneopetetun mallin ja ihmisen tekemää luokittelua “Afrojazz Club”-nimiselle tapahtumalle

The screenshot shows the SAS Visual Analytics interface with a table titled 'LE\_havainnot'. The table compares classification results for the event 'Afrojazz Club'. The columns are 'name\_fi', 'description\_fi', 'Manuaalinen kategoria', 'Koneopetettu kategoria', and 'musiikki'. The 'Koneopetettu kategoria' column is expanded to show a list of categories like 'YLAKATEGORIA/Musiikki', 'MUSIIKKI\_YLA/Jazz', etc.

name_fi	description_fi	Manuaalinen kategoria	Koneopetettu kategoria	musiikki
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/teatteri/taide/teatteri	YLAKATEGORIA/Musiikki	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/teatteri/taide/teatteri	YLAKATEGORIA/Joulu	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/teatteri/taide/teatteri	MUSIIKKI_YLA/Jazz	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/tanssi	MUSIIKKI_YLA/Jazz	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/populaarimusiikki	YLAKATEGORIA/Joulu	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/populaarimusiikki	MUSIIKKI_YLA/Jazz	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/populaarimusiikki	YLAKATEGORIA/Joulu	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/joulu	YLAKATEGORIA/Musiikki	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/joulu	YLAKATEGORIA/Joulu	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/jazz	MUSIIKKI_YLA/Jazz	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/jazz	YLAKATEGORIA/Musiikki	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/jazz	YLAKATEGORIA/Joulu	1
Afrojazz Club	Afrojazz Club: Afrojazz Quintet feat Charlotta Kerbs & Emilla Sisco, Dj Lulu Pikkujoulut järjestetään Afrojazz Clubilla marraskuussa...	MANUAALI_PÄÄKATEGORIA/jazz	MUSIIKKI_YLA/Jazz	1

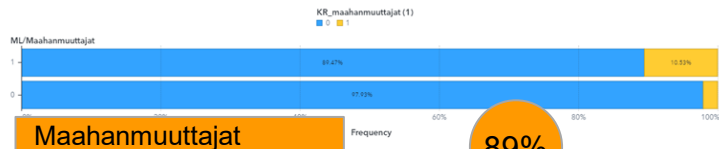
**Ihmisluokittelija** on antanut tapahtumalle luokat Teatteri, Tanssi, Joulu, Populaarimusiikki ja Jazz. **Koneopetettu malli** on antanut sille yläkategoriat Musiikki ja Joulu ja tunnistanut lisäksi, että tapahtuma sisältää musiikin alakategorian Jazz.

Tapahtuma on alun perin luokiteltu kategoriaan Teatteri (luultavasti tapahtumapaikan vuoksi: Kansallisteatterin Lavaklubi), mutta luokka on virheellinen. Tapahtumassa voi tanssia, mutta se ei ole tanssi-esitys. Enemmän ei ole aina parempi.

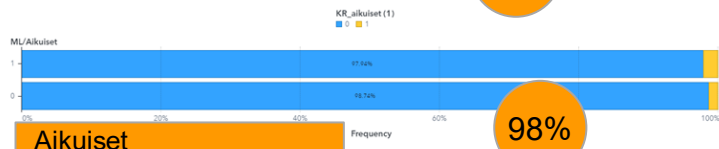
# 3. Kokeilun tulokset: osuvuus kohderyhmittäin



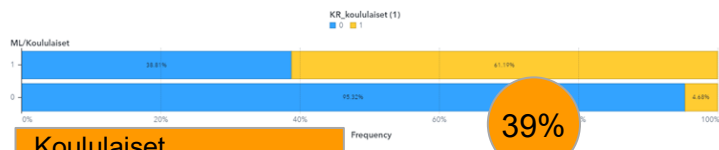
Maahanmuuttajat



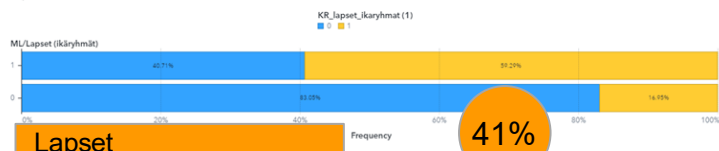
Aikuiset



Koululaiset



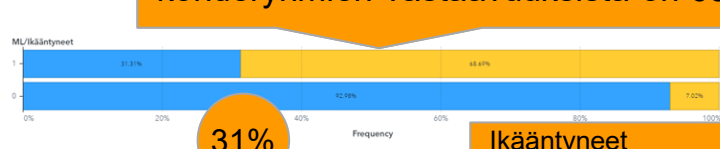
Lapset



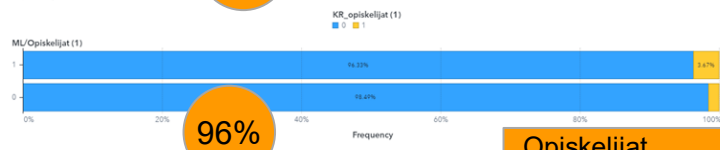
Kohderyhmien luokittelu:  
Koneoppimilla tuotettu luokittelu verrattuna alkuperäiseen luokitteluun

Koneopetettu malli päätyi samoihin kohderyhmiin kuin ihmiset välillä 97%–31%. Keskiarvo kohderyhmien vastaavuuksista on 68 prosenttia.

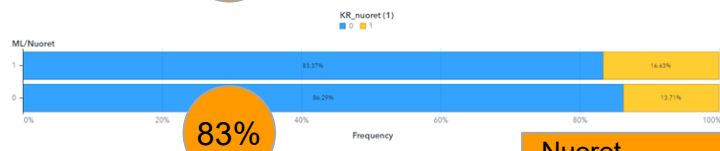
Ikääntyneet



Opiskelijat



Nuoret





# 4. Opit tekoälyn mahdollisuuksista

- Tekoälypohjaiselle asiasanoitukselle on hyvät lähtökohdat myös tuotantopalvelussa.
- Tapahtumadatan automaattisen asiasanoittamisen parissa kannattaa jatkaa myös kokeilun jälkeen.
- Työajan säästön / työn korvaamisen ohella tekoäly voi toimia myös asiasanoituksen ja täten suositusten **laadun** parantajana.
- Kun tekoälyn tueksi on laadukasta koulutusmateriaalia, vähenee manuaalisen työn tarve.
  - Testimielessä [Finton asiasanarajapintaa](#) hyödyntämällä saatiin nostettua osumien määrää (esim. “musiikki”, n. 600kpl → 1200kpl)
- Selvät, loogiset yläkategoriat parantavat tekoälyn oppimismahdollisuuksia

# 5. Opit tekoälyn kehittämisestä

- Kehittäminen on oppimista: Käytännön työn aloittaminen paljastaa tehtyjen ennakoivalintojen ja rajausten puutteet ja näyttää suuntaa sille, miten myös tavoitteita kannattaa muokata ja selkiyttää.
- Työn lähtökohtana pitää olla ongelma, ei ratkaisu: Ratkaisun etsiminen terävöittää ongelman määrittelyä ja tällä tavoin ei sitouduta liian aikaisin ennako-oletuksista lähteviin tekotapoihin.
- Ratkaisun etsimisessä voi selventyä myös mikä sitä tukisi: kokeilun aikana saadaan uusia ideoita ja suuntia kehitykselle.
- Tekoälykokeilu on myös hyvä keino selventää tavoitteita.
- Kehittämisprosessi vahvasti hypoteesin, joka toimi kokeilun lähtökohtana: olemassaoleva luokitteludata ei ollut riittävän laadukasta, jotta se olisi toiminut koneopettamisen pohjana  
→ Luokitteludatan laatu ei siis ole riittävä myöskään sinällään.

# 5. Opit tekoälyn kehittämisestä

- Algoritmin valinta perustui tekoälykokeilutilaisuudessa toimittajalta saatuun suositukseen ja toimittajan työvoiman saatavuuteen. Valintaprosessi ei sinällään ollut systemaattinen.
- Kokeilussa ei tullut ilmi varsinaisesti eettisiksi luokiteltavia kysymyksiä.
- Kouluttamisessa oli hyvä vertailla vaihtoehtoisia lähtökohtia, (olemassaoleva luokittelu vs. sääntöpohjainen luokittelu) ennen lopullisen mallin valintaa
- Asiakaslähtöinen (so. kaupunki) jatkokouluttaminen olisi ajan salliessa onnistunut valitulla työkalulla hyvin.

# 6. Opit kokeilemisesta

- Selkeä tavoite ja helposti saatavat data-aineistot helpottivat työn aloitusta ja antoivat yhteistyökumppanille hyvät lähtökohdat.
- Avoimen rajapinnan päälle on helppo toteuttaa tekoälykokeilu.
- Kokeilun niputtaminen toisen GDPR-riippuvaisen kokeilun kanssa ja kaupungin Azure-kokeilualustan pystyttäminen lykkäsivät varsinaista aloitusta usealla kuukaudella → Kokeilulle jäi varsin vähän aikaa.
- Ehdotus: Luodaan pilivpohjainen ns. kokeiluhiekkalaatikko, jossa aloittaminen käy mahdollisimman kitkatta.
- Kokeilu auttoi validoimaan perusidean ja terävöittämään tavoitteita ja keinoja.
- Kokeilun aikana saatiin parannettua tavoitetta: a) yläkategorioiden hyvä muodostaa looginen kokonaisuus b) apuna voitaisiin käyttää myös Finton YSO-rajapintaa c) lisäkehitys avoimelta pohjalta.

# 7. Opit resursoinnista

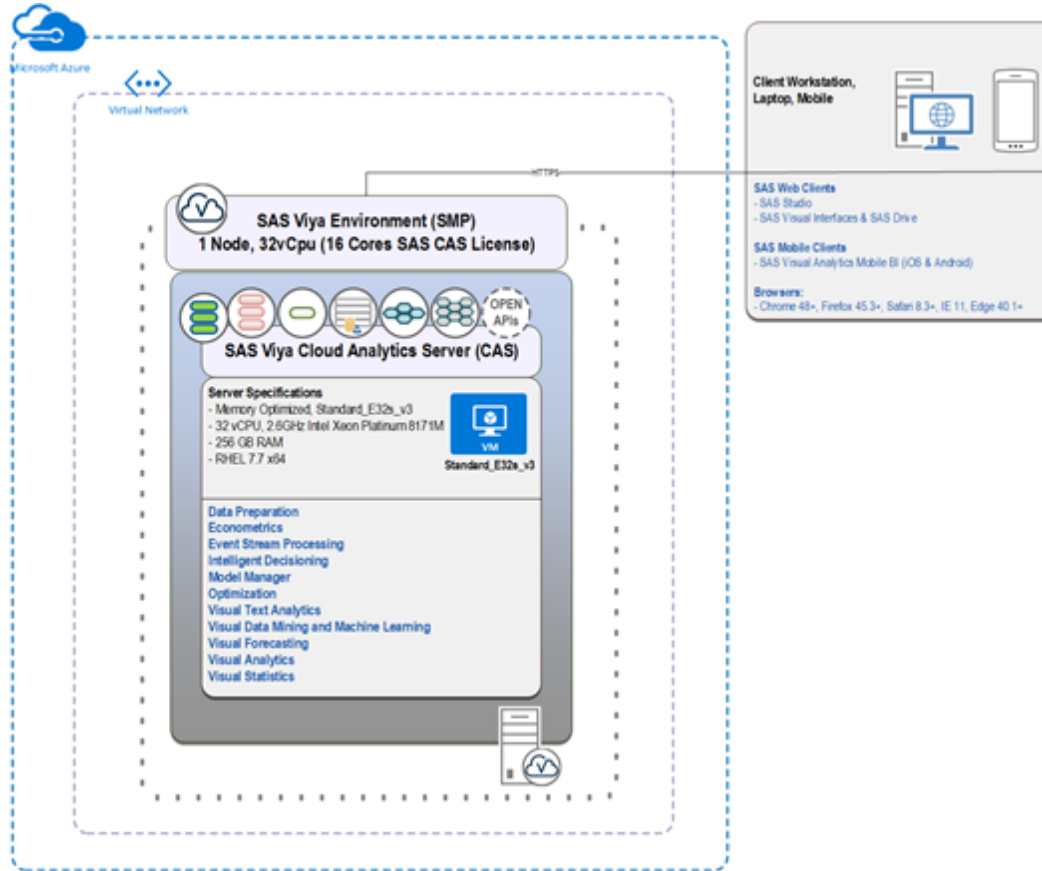
- Kokeilijoista riippumattomista syistä kokeiluaika jäi lyhyeksi siihen, että iteratiivista kehitystyötä olisi ehditty tehdä rauhassa.
- Tekninen toteutus oli onnistunut, mutta samasta syystä myös tuloksien havainnollistamiseen olisi tarvittu enemmän aikaa.
- Resursointi ja budjetti oli sopiva.
- Pidemmällä aikajaksolla olisi päästy lähemmäs periaatteessa tuotantokelpoista mallia mm. asiakaslähtöisen opettamisen kautta.

# 8. Kokeilun tekninen ympäristö

- Kokeilu toteutettiin kaupungin Azure-pilvessä SAS:n Viya-alustalla.
- SAS teki asennuksen ja Viya-alustan hallinnoinnin, Fujitsu hoiti palvelinpuolen.
- Data haettiin kokeilua varten dumpina Linked Events - tapahtumarajapinnasta.
- Koska SAS Viya tekee laskennan muistinvaraisesti, erillistä tietokantaa ei tarvittu.
- Ns. kokeiluhiekkalaatikko olisi tarpeellinen.

*Teknisen ympäristön kuvaus jatkuu seuraavalla sivulla →*

# 8. Kokeilun tekninen ympäristö

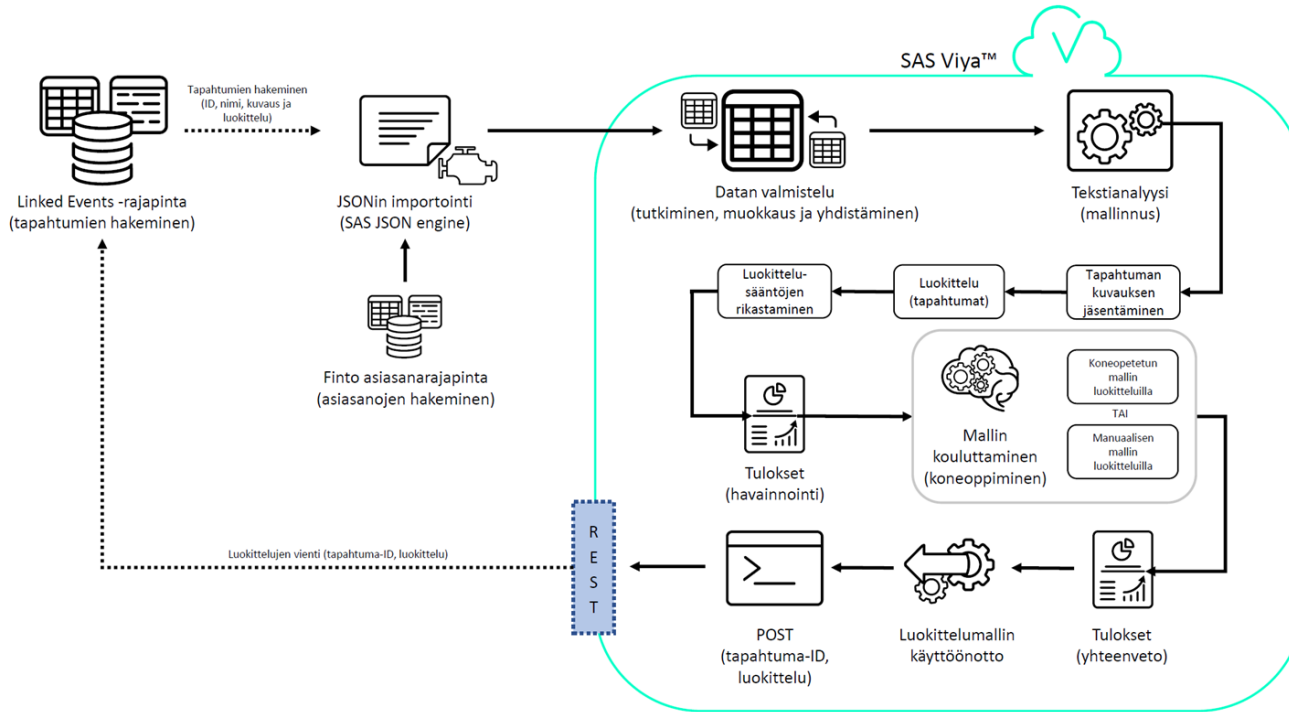


# 9. Kokeilun data

- Data saatiin avoimesta Linked Events -rajapinnasta vaivattomasti.
  - Luotiin datadumppi events.csv
    - Aineisto sisältää Linked Events -tapahtumarajapinnan tapahtumat ajanjaksolta 12/2018-12/2019
  - event\_keywords.csv
    - Aineisto sisältää tapahtumien yksilöivän ID:n ja olemassa olevien käyttäjien tekemien luokitteluiden koodit
      - Datojen yhdistäminen onnistui helposti tapahtuma ID:tä käyttäen
  - Näiden lisäksi käytössä oli YSO-koodit sisältävä listaus asiasanoista, joita kokeilussa käytettiin: 50kpl YSO-asiasanaa + myöhemmin kokeilussa lisätty 9kpl YSO-kohderyhmiä



# 10. Jatkopäätökset ja -ideat



# 10. Jatkopäätökset ja -ideat

Samalla teknologialla ei tällä hetkellä toteuteta tuotannollistamista.

**Perusteet:** a) SAS:n lisenssisopimus järkevä vain suuremmalla käyttäjätjoukolla (e. toimiala, kaupunki). b) Kokeilun kuluessa on tutustuttu varteenotettavaan avoimeen ratkaisuun (Annif), joka jatkokehittämisellä voisi sopia hyvin Linked Eventsille. Tällöin kaikki tuotokset jäisivät a) kaupungin b) muiden kaupunkien käyttöön.

Tuotannollistamiselle on saatu kokeilussa hyvät reunaehdot ja askelmerkit tarvittavista vaiheista. Seuraavalla kalvolla on hahmoteltu tuotantopalvelun toimintaprosessi ylätasoisesti.