



JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

Datan laadun ongelmat yksinkertaisesti esille Case: Tunkio

Tietoasiantuntija Jenni Kokko

Toiminnanohjaus, Jyväskylän yliopisto

IR-tekijäverkoston tapaaminen 24.4.2024



Rekisterien tiedon laatu

Lähtötilanne

- Datassa on virheitä
- Virheitä löydetään sattumalta
- Virheitä korjataan satunnaisesti



Tavoite

- Dataan ei synny virheitä
 - Virheiden syntyminen on estetty
- (Tunnetut) virheet löydetään systemaattisesti
- Kaikki virheet korjataan

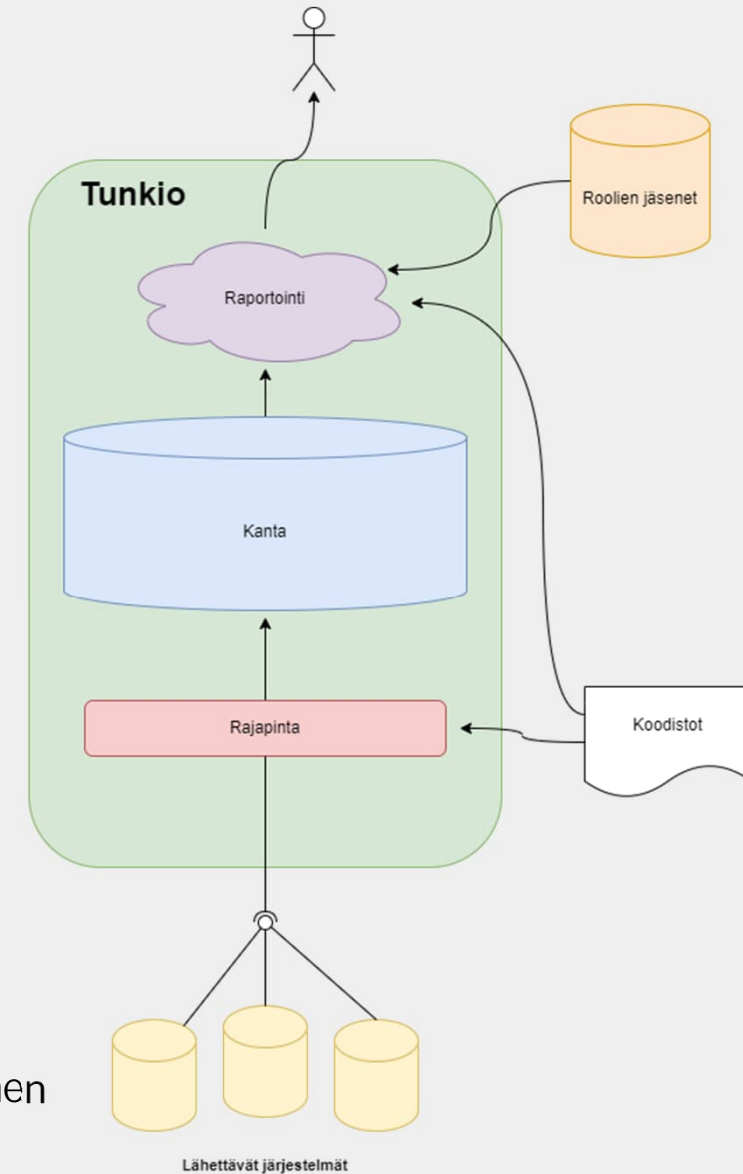
Lähdejärjestelmissä
kehitetään syötettävien
tietojen validointia





Ongelmatapausten selaamisen keskitetty palvelu Tunkio

- Itse Tunkio on
 - rajapinta (sisään)
 - tietokanta
 - raportointi
- Tunkio ei etsi ongelmia
- Vastuu ilmoitusten lähettämisestä on lähdejärjestelmän omistajalla
 - Lähdejärjestelmä tai erikseen lähettämistä varten rakennettu "palikka"
- Tunkio kokoaa ilmoitukset ja näyttää käyttäjälle hänelle kuuluvat ilmoitukset
- Suunnittelupäätös: Loppukäyttäjän näkökulma ei saa olla liian tekninen

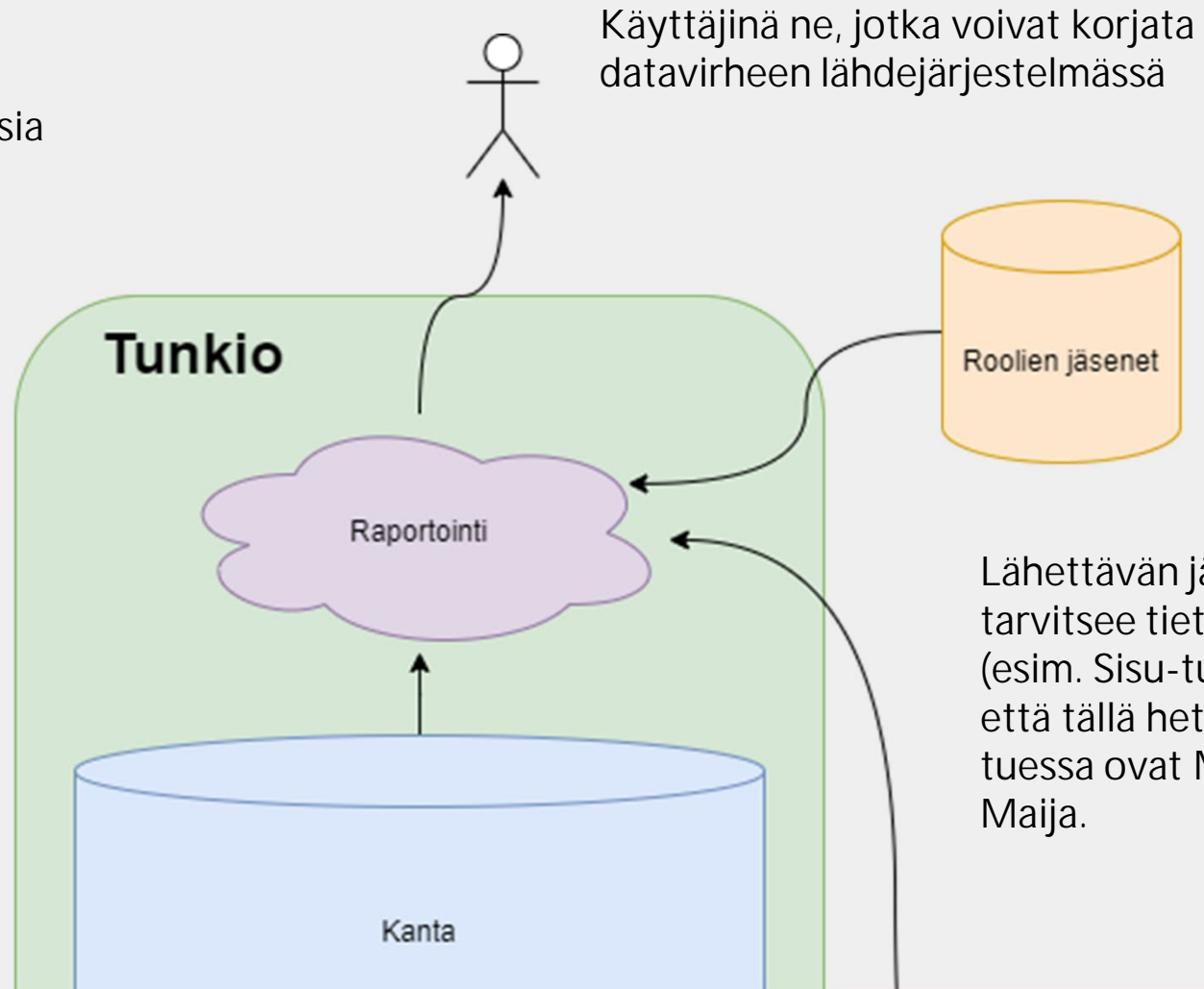




Tunkion raportointi

Tunkiossa ei kuitata ilmoituksia eikä korjata datavirheitä.


Raportointi = Power BI



Lähtevän järjestelmän tarvitsee tietää vain rooli (esim. Sisu-tuki), ei sitä, että tällä hetkellä Sisussa ovat Matti ja Maija.



Tunkion raportointi Power BI:ssä



Tunkio

TUNKIO

Tunkio

- Aikanäkymä
- Luokittelu
- Virheilmoitukset

Service: Kaikki

Tag: Kaikki

Short Message: Kaikki

Identifier & Short message: Kaikki

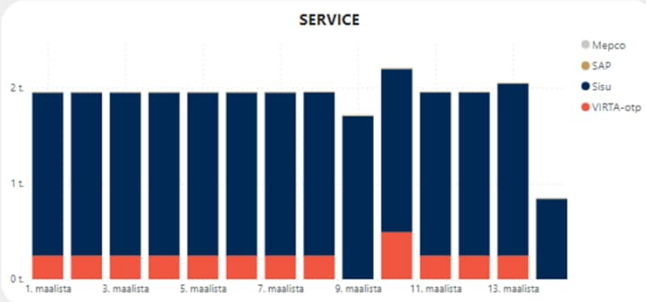
Date: Vime... 14 päivää
1.3.2023 - 14.3.2023

TAG

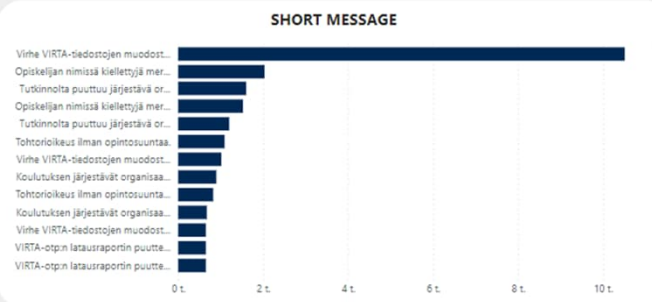
SHORT MESSAGE

SERVICE

SERVICE



SHORT MESSAGE



Notification Identifier	Notification Date	Host	Service	Tags	Target Roles	Target Persons	Short Message	Link	Full Message
Tunkio-testidataa hälytyksiä varten	14.3.2023	Jitterbit	Sisu	hälytys; powerbi; testaus; testilähetyskerta 2023-03-14 14:40:02	dw-tiimi		Tunkio-testidataa hälytyksiä varten	🔗	Tunkio-testidataa hälytyksiä varten
Tunkio-testidataa hälytyksiä varten	14.3.2023	Jitterbit	Sisu	hälytys; powerbi; testaus; testilähetyskerta 2023-03-14 14:40:02	dw-tiimi		Tunkio-testidataa hälytyksiä varten	🔗	Tunkio-testidataa hälytyksiä varten
Tunkio-testidataa hälytyksiä varten	14.3.2023	Jitterbit	Sisu	hälytys; powerbi; testaus; testilähetyskerta 2023-03-14 14:40:02	dw-tiimi		Tunkio-testidataa hälytyksiä varten	🔗	Tunkio-testidataa hälytyksiä varten
Tunkio-testidataa hälytyksiä varten	14.3.2023	Jitterbit	Sisu	hälytys; powerbi; testaus; testilähetyskerta 2023-03-14 14:40:02	dw-tiimi		Tunkio-testidataa hälytyksiä varten	🔗	Tunkio-testidataa hälytyksiä varten
Tunkio-testidataa hälytyksiä varten	14.3.2023	Jitterbit	Sisu	hälytys; powerbi; testaus; testilähetyskerta 2023-03-14 14:40:02	dw-tiimi		Tunkio-testidataa hälytyksiä varten	🔗	Tunkio-testidataa hälytyksiä varten



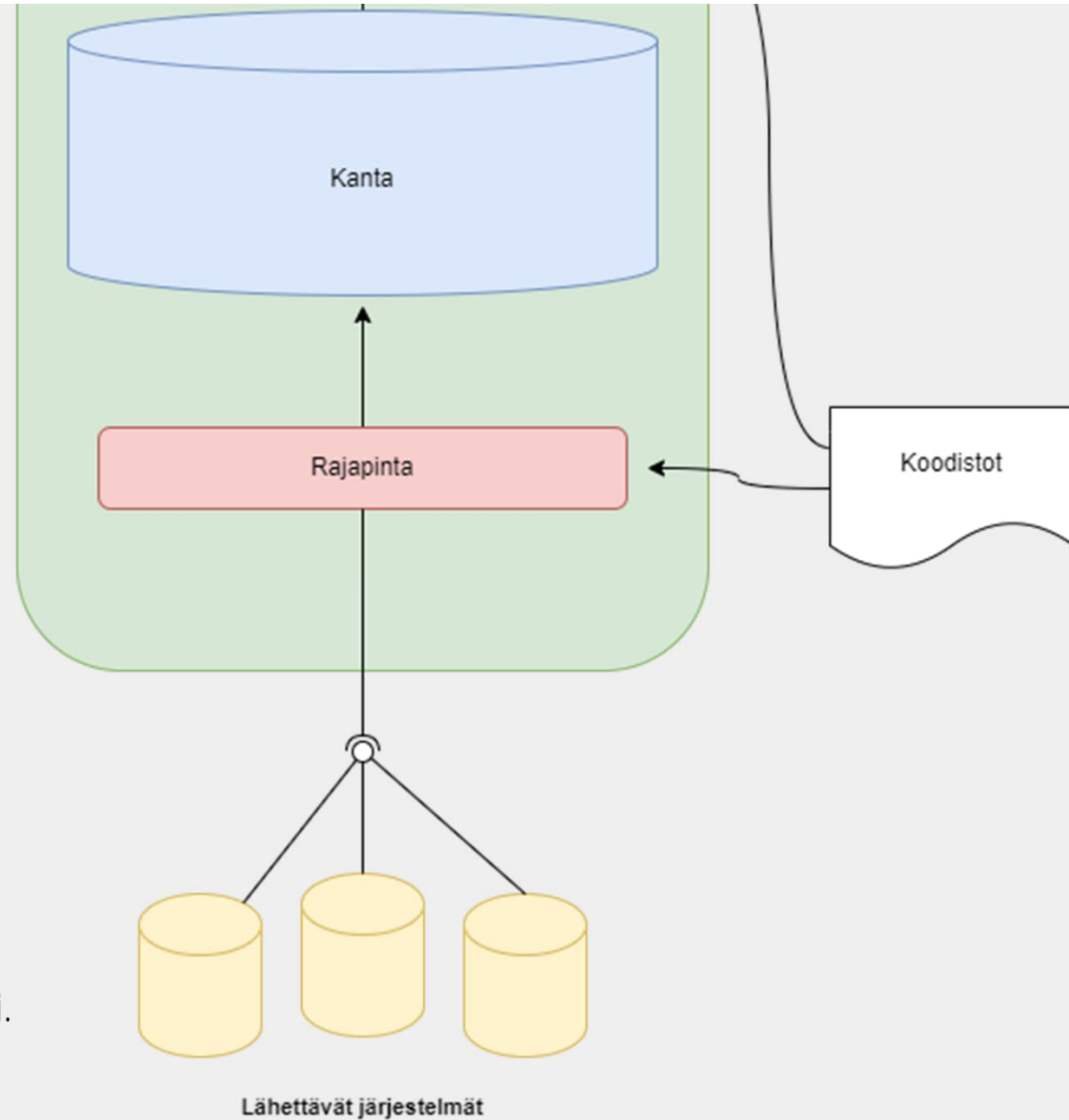
Tiedot Tunkioon

Tunkio ei ole arkisto.
Yli 2kk vanhat ilmoitukset poistetaan.

Rajapinnan käyttäjäksi pitää rekisteröityä.
Rajapinta varmistaa, että tulevat tiedot ovat vaatimusten mukaisia (mm. koodistot).

Tunkio ei itse etsi datavirheitä.
Tunkio vain vastaanottaa.

Lähettävä järjestelmä muuttaa omat tietonsa Tunkion tietomallin mukaisiksi.





Tunkio-ilmoituksen sisältö ja lähettäminen

- Kaikki tiedot eivät ole pakollisia
- Osa tiedoista luokitteluja
- Kenttinä esim.
 - Lähettäjä
 - Sisällön lähdejärjestelmä
 - Kohderoolit
 - Tunniste (esim. tutkinnon tunniste)
 - Otsikko
 - Linkki lähdejärjestelmään
- Ilmoitusten lähettäminen:
 - Https-pyyntö
 - Esim. curl

```
curl -X 'POST' 'https://xxxxx' -H 'api_key: 123456..' -H "Content-Type: application/json" -d '{"identifier": "9123456789", ..., "full_message": "Jotain vikaa identifierissä 9123456789. Korjaa lähdejärjestelmässä!" } ]'
```



Case: Miten Tietovarastossa keksitään uusia Tunkio-etsintäsääntöjä

"Miksi tiedekunnasta X on valmistunut tutkinto B?
Tutkinnot B valmistuvat vain tiedekunnasta Y."

Tiedontuotannossa
ihmetellään rekisterin tietoja
= löydetty virhe kirjauksessa

Raportoinnin suunnittelussa
valitaan tietty päättely



Laaditaan sääntö, jonka
perusteella lähetetään
Tunkioon ilmoituksia

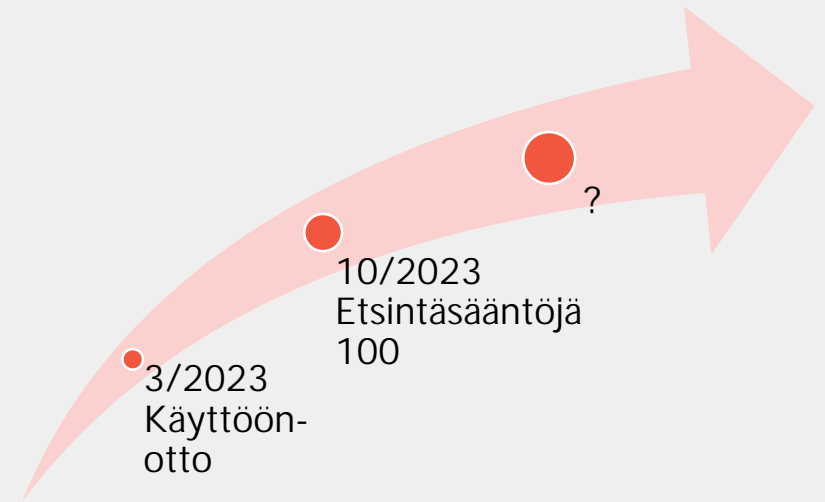
"Sanoivat, että tässä kentässä meillä tulee aina A tai B, ei koskaan C"

"Meillä tähän kirjataan aina vain yksi organisaatio, vaikka järjestelmä mahdollistaa useamman organisaation kirjaamisen."



Minkä kokoinen on Tunkio?

- Tunkiossa on n. 40 käyttäjää
 - Esim. Datatiimi, Sisu-pääkäyttäjät, Mobility Online – pääkäyttäjät, Talouspalvelut, tutkinto-opintojen palvelutiimi
- Tunkioon tulee joka päivä n. 3 800 ilmoitusta
- Tunkio otettiin käyttöön 3/2023
- Tietovarastossa nyt 100 tietojen sisällön laadun tarkastussääntöä, joiden tulokset lähetetään Tunkioon
 - Tekemistä vaille valmiina suunniteltu 33 uutta etsintäsääntöä tietovarastotiimille tehtäväksi





Kehittäjien näkökulma tulevaisuuteen

Kehitettävää

- Tunkio-ilmoituksessa ei riitä pelkkä tunniste, vaan tarvitaan myös esim. nimi
 - Kaikista lähdejärjestelmistä ei suoraan näe tunnistetta tai ei voi hakea sillä
- Tunkiossa pitää muistaa käydä
 - Ilmoitukset tulossa

Käyttö laajenee

- Tunkioon on pyydetty pääsyjä laajemmalle joukolle, jotta tietoja lähdejärjestelmässä käsittelevä saisi suoraan tiedon virheestä
 - Ei kierrätystä pääkäyttäjien kautta
 - Tunkioon tuodaan uusia rooleja
- Tunkio-käyttäjät ovat toivoneet lisää etsintäsääntöjä
- Tunkio-sääntöjen toteuttajat (kehittäjät) ovat itse lisänneet omaa työtään varten uusia etsintäsääntöjä



Käyttäjäpalaute

- "Yksi paikka mistä näkee kohtuu reaaliaikaisesti kaiken oleellisen samasta taulusta, saa kokonaiskuvaa ja voi tehdä toimenpiteitä sen mukaan"
- "Sinne saa kaiken, mitä vain osaa pyytää. Helpottaa huomattavasti virhetilanteiden tutkimista, kun ne kerätään yhteen paikkaan automaattisesti."
- "En muista käydä Tunkiossa: Pikemminkin en ole aina ehtinyt, mutta tähän varataan lisää aikaa ensi vuonna."
- "Kehitettävät asiat: Ei tule tässä vaiheessa mieleen. Varmasti asioita nousee esiin, kun ehtii käyttämään tarpeeksi."





Yhteenveto

Ongelmatapausten selaamisen keskitetty palvelu Tunkio

- Dataan syntyy virheitä
 - Jotain pitää tehdä
- "Tunkio on hyvä järjestelmä"
 - Käyttäjäpalaute
 - Käytön kasvu
 - Etsintäsääntöjen määrän kasvu
- Datan laadun ongelmat yksinkertaisesti esille

"Garbage in, mushrooms out"



Jenni Kokko
jenni.m.kokko@jyu.fi
Jyväskylän yliopisto



JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

