



SUOMEN
OLYMPIAKOMITEA

ESIMERKKI DATAN KÄYTÖSTÄ JA DATAEKOSYSTEEMISTÄ OLYMPIAKOMITEASSA

TOPIAS KOUKKULA / 2.10.2024



Tausta

- Miten hyödynnämme nykyaikaisia tietoteknisiä ratkaisuja raportoinnissa ja tiedolla johtamisessa.
- Miten teemme tietotulosten raportoinnista saavutettavaa ja käyttäjäystävällistä.
- Miten ratkaisemme ”kysyntä-tarjonta ongelman” (datatyön resurssit vs. verkoston laajuus) yhteisössämme ja maksimoimme tehokkuuden.
- Miten pääsemme toteuttamaan (demo)projektiluontoisesti urheilun datastrategiatyön mukaista toimintamallia.



Tausta

- Miten hyödynnämme nykyaikaisia tietoteknisiä ratkaisuja raportoinnissa ja tiedolla johtamisessa.
- Miten teemme tietotulosten raportoinnista saavutettavaa ja käyttäjäystävällistä.
- Miten ratkaisemme ”kysyntä-tarjonta ongelman” (datatyön resurssit vs. verkoston laajuus) yhteisössämme ja maksimoimme tehokkuuden.
- Miten pääsemme toteuttamaan (demo)projektiluontoisesti urheilun datastrategiatyön mukaista toimintamallia.

Tietolähteet

Tietokannat

Analytiikkaympäristöt

Ohjelmointikieliriippumattomuus

Raportointityökalut

**Tietoturva,
MyData, GDPR,
Tietojenluovutusso-
pimukset, sääntely**

Yksilöinti

Sportti-ID



Ekosysteemikuvaus

 = Toteutettu

 = Toteutumassa

 = Vielä selvittämättä



Tiedon lähteet



jne....



API-rajapinta haku



Suomisport
data

API-rajapinta haku



Connector lukeminen



Connector lukeminen



Analytiikka- / tiedon jalostamisympäristö



Raportointiympäristö

Databricks
notebook



Rajapintahaku



Rajapintahaku



Upottaminen



Reaaliaikaisuus
Tiedon yhdistäminen
Tekoäly/koneoppiminen



Tiedon lähteet



jne....



API-rajapinta haku



Suomisport
data

API-rajapinta haku



Connector lukeminen



Connector lukeminen



Analytiikka- / tiedon jalostamisympäristö

databricks



Raportointiympäristö

Databricks
notebook



Rajapintahaku



Rajapintahaku



Upottaminen



SUOMI
SPORT

Reaaliaikaisuus
Tiedon yhdistäminen
Tekoäly/koneoppiminen

Lopputuote: Lajin drop out –
näkömä lajiliittokäyttäjän Sporttitilillä
"reaaliaikaisena".



Esimerkkejä ekosysteemimallilla toteutetuista/toteutuvista urheilun datastrategiatyötä mukailevista projekteista

 = Toteutettu

 = Toteutumassa

 = Vielä selvittämättä



Liikunnan ja urheilun vaikuttavuusanalyysi

"Missä määrin liikunnan ja urheilun keskeiset muuttujat ovat aluetasolla tarkasteltuna yhteydessä muihin yhteiskunnallisiin sosioekonomisiin ja –demografisiin muuttujiin"

Reaaliaikaisuus
Tiedon yhdistäminen
Tekoäly/koneoppiminen

Tiedon lähteet



Sosioekonomiset ja –
demografiset tiedot



Lisenssitiedot

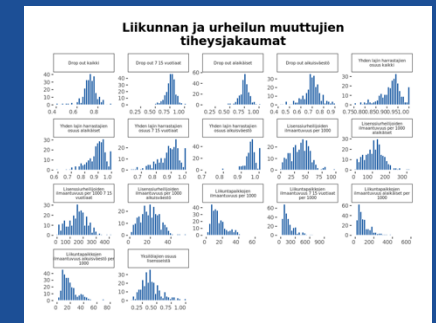
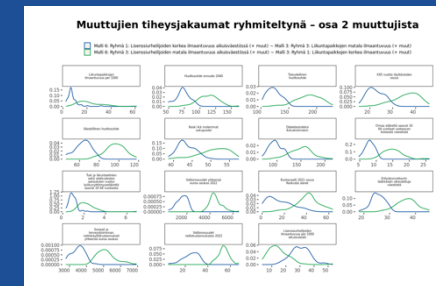


Suomisport
data

Analytiikka- / tiedon jalostamisympäristö



Tietotulokset



Yleisurheilun kilpailutulostanalyysi

"Yleisurheilun kilpailutulosten analysointi, yhdistettynä Suomisportin lisenssitietoihin"

Reaaliaikaisuus
Tiedon yhdistäminen
Tekoäly/koneoppiminen

Tiedon lähteet

TILASTOPAJA

Yleisurheilun
kilpailutulostiedot



Sportti-
ID

Suomisport
data

Lisenssitiedot



Analytiikka- / tiedon jalostamisympäristö

 databricks



Tietotulokset

"Kansainvälisen
yleisurheilun
vaatimustaso"

"Kansainvälisen
yleisurheilijan
kehityspolku (vs.
suomalaiset)"

"Drop outin /
monilajiharrastamisen ja
kilpailutulosten väliset
yhteydet nuorilla"



Yleisurheilun kilpailutulokset yhdistettynä harjoitustietoihin

Lisätään edellisen slaidin mukaiseen malliin harjoitustiedot (esim. 360training).

Tiedon lähteet

TILASTOPAJA



Suomisport
data

Yleisurheilun
kilpailutulostiedot

Harjoitustiedot

Lisenssitiedot

Sportti-
ID

Analytiikka- / tiedon jalostamisympäristö

 databricks



Reaaliaikaisuus
Tiedon yhdistäminen
Tekoäly/koneoppiminen

Tietotulokset

"Kansainvälisen
yleisurheilun
vaatimustaso"

"Kansainvälisen
yleisurheilijan kehityspolku
(vs. suomalaiset)"

"Drop outin /
monilajiharrastamisen ja
kilpailutulosten väliset
yhteydet nuorilla"

"Harjoittelun- ja
palautumisen vaikutus
tuloksentekoon /
tulosvarmuuteen"



Dataesimerkki: Case Yleisurheilun tulostietoanalyysi

- Tuodaan Tilastopajan tulostiedot raakadatana analytiikkajärjestelmään (Databricks-sovellutus)
 - ~ 7 miljoonaa datapistettä (v. 2022 eteenpäin)
 - analytiikkajärjestelmällä määritellään datan käsittelysäännöt ja tuotetaan tietotulokset raportoitavaan/siirrettävään muotoon
- Demoprojektina toteutettu 100 suomalaisurheilijan kilpailutulosten analyysi ja vertailu kansainvälisiin verrokkeihin
 - Tulostietotrendi, vertailut kv-verrokkeihin (“Pariisin rajat” sekä Pariisin menestyjät)
- Tietotulosten tekninen siirtäminen mahdollinen muihin järjestelmiin (esim. 360training-palvelun analytiikkanäkymään)
- Urheilijakohtaiset tiedot kohdennetaan Sportti-ID:llä muihin tietoihin

