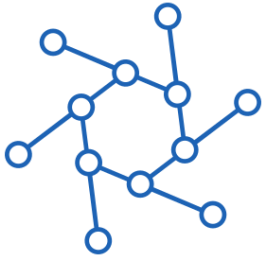


AURIA CLINICAL
INFORMATICS

Kohti kansallista sote-tietojen saatavuuspalvelua

Kansallisen saatavuustietokannan **tiekartta**
-ehdotus arkkitehtuurista
-keinot löytää ja rikastaa nyt hajanaisten
aineistojen sisältämä tieto.



Kaupallisia saatavuustietokantoja:

- testasimme Belgialaisen Custodix-yrityksen ylläpitämä InSite Platformia → kehitetty EU-projektissa IMI-projektina +16 milj. € budjetilla 10 lääkeyrityksen ja 25 muun akateemisen/yksityisen kumppanin kanssa tietokantarungosta(12b2).
- Espoossa vuonna 1997 perustettu BC Platforms tarjoaa omaa, erityisesti geenitietoon erikoistunutta, mutta kaiken klinisen tiedon kattavaa globaalia saatavuuspalveluaan. Asiakastestaus päästiin aloittamaan vasta 08/2018



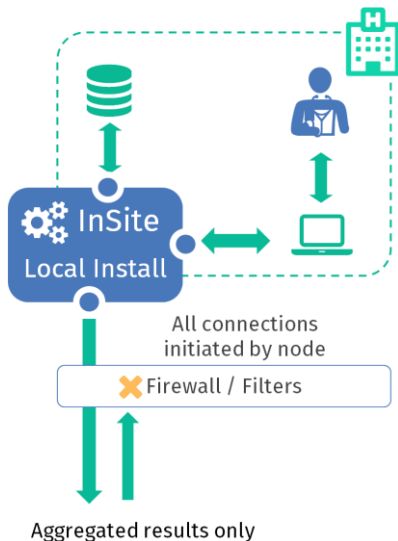
Insite Link



Insite Platform



Laadun validointi
tuotannossa:
Fibstroke(stroke/TIA)
Q1-Q2/18



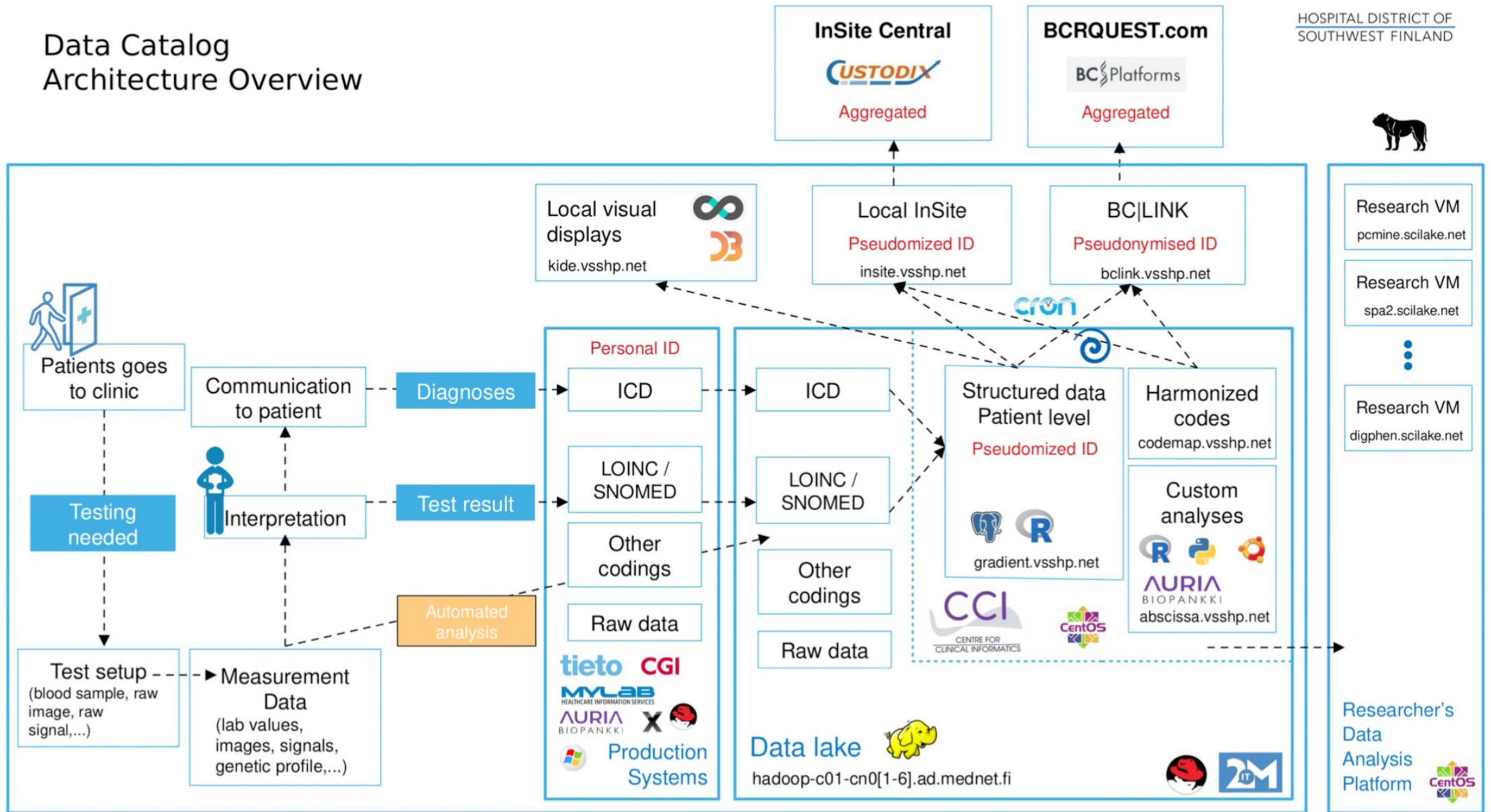
BC|Link



BC|REQUEST Platform
Q3/18

Data Catalog Architecture Overview

HOSPITAL DISTRICT OF
SOUTHWEST FINLAND



Kuva 1. Tieto-, järjestelmä- ja teknologia-arkkitehtuurin kuvaus saatavuuspalvelun käyttäjän näkökulmasta (VSSH).

Käyttökelpoisuus

-----TESTICASE KOHORTTI-----

Aikuisia (>18 vuotiaita) 1.1.2003 diagnosoihetkeltä alkaen

1. Potilaalla Diagnosoitu eteisvärinä → Atrial Fibrillation (AF) or atrial flutter I48
 2. Yksi tai useampi seuraavista aivoverenkierron häiriöistä → Disturbances in cerebral blood flow during **01.01.2003-31.12.2012**:
 - a. Aivoinfarkti → Ischemic stroke, I60.0-I60.9, I61.0 - I61.9, I62.0 - I62.9, I63.0 - I63.9, I65.0-I65.9, I66.0-I66.9, I69.0 - I69.9 or
 - b. Ohimenevä aivoverenkiertohäiriö → TIA, G45.0 - G45.9, G46.0 - G46.9 or
 - c. kallonsis. Verenvuoto → Intracranial bleeding, S06.0- S06.9
- Kuinka monella potilaalla oli **eteisvärinä AF diagnoosi ENNEN ensimmäistä diagnoosia a, b tai c.** ("ENNEN" määritelty, ainakin 1 päivä ennen diagnoosia a, b tai c).



-Insite palvelee hyvin kohortin muodostusta.



-perusidea selvittää **tietyt kriteerit omaavat potilasmäärät** per sairaala (aikamääre, diagnoosi, toimenpide, lääkitys, sukupuoli, ikä diagnosoihetkellä jne)
Tähän tarkoitukseen se soveltuukin hyvin ja sen käyttö/implementointi osaksi sairaalan järjestelmää ei maksa mitään ja mahdollistaa näyteikkunan klinisen tutkimuksen maailmanmarkkinoille. Maksajana toimii tutkimuksen tekijä.



Käyttöliittymä näyttää selkeästi potilasmäärän vähenemisen sitä mukaa kun kriteereitä lisätään. **Insite listaa potilasmääriä.** Fibstroke-tutkimuksessa oli listattu myös mm. **aivotapahtumien** lukumääriä per potilas, minkälainen haku on Insitellä mahdotonta tai ainakin hyvin hankalaa.





Insite ajolla kohdepopulaation määrä oli kaksinkertainen Fibstroke-tutkimuksen tuloksiin verrattuna.

→Ei yhtä kohortin suuruuseroja selittävää tekijää. 

→Tällaisiä syitä voisi olla mm. se, että väitöskirjassa on alkuperäisen kohortin siivoamiseen/data curatointiin käytetty tapoja joita ei ole selitetty niin että tutkimus voitaisiin toistaa mielekkäästi.

→Poiminta tuolloisesta SAS-järjestelmästä on voitu toteuttaa hieman eri tavoin→ ei selitetty osana väitöskirjaa.

→Toki voi olla, että Insite on tarkempi.

→Mainittakoon vielä, että tiedustelimme useampaan kertaan Insitelta asiaa ja tunnistimme, että insite ottaa hakuihinsa mukaan myös potilaita, joilla exclude kriteeri päällä, eli potilaita jotka eivät pitäisi tulla mukaan mikäli potilaspoiminta suoritettaisiin SQL-skriptin nykyosaamisemme valossa.

6230



2920



RESULTS

	Insite search	Fibstroke manual search
Number of patients who suffered from stroke, TIA or Intracranial bleeding during 2003-2013 and who had ever been diagnosed of AF (see Fig 1)	6230	2920
Number of events (stroke, TIA or/and intracranial bleeding) in total during 2003-2012 (as some had more than one event).	9114	3762
AF diagnosed before the event. Please see explanation in the end of this document why the last values (4493 and 2177) are not comparable with each other.	4493	2177

ical
kbench

InSite Studies

Cohort Analysis

Recruiting Studies

NiinaN

Notification

Cohort Studies

autovalidate

Cohort builder

Cohort Builder

Save Cohort & Explore Patients

Save Cohort

Recalculate results

Show matches

Choose cohort:

1 016 137
IN COHORT

Age 18 or more on 1.1.2003 (1 016 137 patients)

Reference date:

31 DEC 2012

FILTERS

Filters + Add filter

1. first diagnosis of Atrial fibrillation and flutter

23 865
IN COHORT

2. the following events:

first diagnosis of Nontraumatic subarachnoid hemorrhage

or first diagnosis of Nontraumatic intracerebral hemorrhage

or first diagnosis of Other and unspecified nontraumatic intracranial hemorrhage

or first diagnosis of Cerebral infarction

or first diagnosis of Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction

or first diagnosis of Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction

or first diagnosis of Sequelae of cerebrovascular disease

or first diagnosis of Vascular syndromes of brain in cerebrovascular diseases

or first diagnosis of Intracranial injury

or first diagnosis of Transient cerebral ischemic attacks and related syndromes

happened at most 10 Years before now

6 230
IN COHORT

BC|REQUEST.COM SEARCH ANALYZE RESULTS COLLECTIONS COLLECTION SEARCH Minna Blomster

Q c50

Demographics Findings & Observations Diagnoses Lab Values

Malignant neoplasm of axill...
Malignant neoplasm of bre...
Malignant neoplasm of cen...
Malignant neoplasm of low...
Malignant neoplasm of low...
Malignant neoplasm of nip...
Malignant neoplasm of ovrl...
Malignant neoplasm of uns...
Malignant neoplasm of upp...
Malignant neoplasm of upp...

Group 1

Drag in here to search

Select gene
Select Collection List
Query Name
ID1528444713864

Run

1. Miten kaikki C50 Group 1 saadaan laatikkoon? Voiko useampia valita esim jollain näppäin yhdistelmällä?

2. Olisiko mahdollista että tällaisessa haussa vastaukset skaalautuisi niin, että näki specifikaatiot?

🔍 Type search criteria (for example: gender, age, ...)

Demographics

Gender

☐ Male ☐ Female
☐ Unknown

Tobacco use

☐ Never ☐ Former
☐ Current

Age in years

Findings & Observations

Body height measure (m) [5...

Body mass index (BMI) 25-...

Body mass index (BMI) 30+ ...

Body mass index (BMI) (kg/...

Body mass index less than ...

Body weight (kg) [27113001]

Hip circumference (m) [284...

Waist circumference (m) [2...

Waist/hip ratio (m) [248367...

Diagnoses

Malignant neoplasm of axill...

Malignant neoplasm of axillary tail
of unspecified male breast

1
2

Malignant neoplasm of bre...

Malignant neoplasm of breast

1

Malignant neoplasm of cen...

Malignant neoplasm of central
portion of unspecified male breast

1
2

46, XX true hermaphrodite [...]

Abdominal actinomycosis [...]

Abdominal and pelvic pain [...]

Abdominal aortic aneurysm...

Abdominal aortic aneurysm...

Lab Values

A-3520 (LOIN

A-3521 (LOIN

A-3526 (LOIN

Alanine amin

Albumin [Ma

Albumin [Ma

Alkaline phos

Alpha 1 antit

Ammonium i

Amylase [En

Angiotensin

Aspartate am

Bicarbonate

Bicarbonate

Bilirubin.dire

Group 1 ✕

Include

Malignant neoplasm ... ✕

Malignant neoplasm of axillary tail
of unspecified male breast

and

Include

Malignant neoplasm ... ✕

Malignant neoplasm of breast

and

Include

Malignant neoplasm ... ✕

Malignant neoplasm of central
portion of unspecified male breast

Group 2 ✕

Include

Malignant neoplasm o... ✕

Malignant neoplasm of axillary tail
of unspecified male breast

and

Include

Malignant neoplasm o... ✕

Malignant neoplasm of central
portion of unspecified male breast

+

3. En kyennyt lisäämään kaikkia C50 koodeja

samaan ryhmään group 1

–Onko se mahdollista jos niin miten?

4. Auttaisi jos diagnoosi ICD koodin 3-ekaa

nr:ta ainakin näkyisi tässä –Onko
mahdollista lisätä? Peruste helpottaa
kyselyn luomista.

BC-Platforms

R|QUEST

-----TESTICASE KOHORTTI-----

Aikuisia (>18 vuotiaita) 1.1.2003 diagnosoitua alkaen

1. Potilaalla Diagnosoitu eteisvärinä → Atrial Fibrillation (AF) or atrial flutter I48
2. Yksi tai useampi seuraavista aivoverenkierron häiriöistä → Disturbances in cerebral blood flow during 01.01.2003-31.12.2012:
 - a. Aivoinfarkti → Ischemic stroke, I60.0-I60.9, I61.0 - I61.9, I62.0 - I62.9, I63.0 - I63.9, I65.0-I65.9, I66.0-I66.9, I69.0 - I69.9 or
 - b. Ohimenevä aivoverenkiertohäiriö → TIA, G45.0 - G45.9, G46.0 - G46.9 or
 - c. kallon sis. Verenvuoto → Intracranial bleeding, S06.0- S06.9

- Kuinka monella potilaalla oli eteisvärinä AF diagnoosi ENNEN ensimmäistä diagnoosia a, b tai c. ("ENNEN" määritelty, ainakin 1 päivä ennen diagnoosia a, b tai c).

Turku University Hospital Collection

BC-R|QUEST löytää 10934 potilasta kohorttiin

Huom. pvm ehtoja ei kyetty asettamaan.

Ei myöskään ehtoja mikä tapahtuma tapahtuu ennen toista tapahtumaa.

Result :

COLLECTION

Turku University Hospital Collection

COUNT

10934

Query Parameters

GROUP	FIELD	VALUE
Group1	Atrial fibrillation and flutter [I48]	I48 >= 18 [year]
Group2	Nontraumatic subarachnoid hemorrhage [I60]	I60
Group2	Nontraumatic intracerebral hemorrhage [I61]	I61
Group2	Other and unspecified nontraumatic intracranial hemorrhage [I62]	I62
Group2	Cerebral infarction [I63]	I63
Group2	Occls and stenosis of precereb art, not reslt in cereb infrc [I65]	I65
Group2	Occls and stenosis of cereb art, not reslt in cerebral infrc [I66]	I66
Group2	Sequelae of cerebrovascular disease [I69]	I69
Group2	Transient cerebral ischemic attacks and related syndromes [G45]	G45
Group2	Vascular syndromes of brain in cerebrovascular diseases [G46]	G46
Group2	Intracranial injury [S06]	S06
Groups Rule		OR

saatavuuspalvelu	Kohortin koko	Huomio
Insite Platform, Custodix	6230 Potilasta	Insight Platformilla kohortin muodostus pystytään tekemään. Haku verifioitiin myös Custodixin asiantuntijoiden toimesta sillä käyttöliittymä ei ole näin monimutkaisen (Fibstroke aineiston Autovalidate study) kohortin luomiselle ihan yksiselitteinen.
BC-R QUEST, BC Platforms	10934 potilasta	Kohortin rakentaminen ei onnistu autovalidate tutkimuksen edellyttämällä tavalla. -Pvm ehtoja ei kyetty asettamaan. -Ehtoja mikä tapahtuma tapahtuu ennen toista tapahtumaa ei kyetä asettamaan.
VERTAILUKOHDE: AUTOVALIDATE STUDY (SAS:illa poimittu FIBSTROKE aineisto)	2920 Potilasta	tutkimusaineistoon/väitöskirjaan ei ollut kirjattu kriteereitä tai tietoja potilaista, joita ei oltu kelpuutettu mukaan manuaalisen haun perusteella.



Validointitutkimuksen kokemuksen perusteella **Insite soveltuu parhaiten uusien tutkimusten suunnittelutyökaluksi** sopivien potilaspopulaatioiden ja heitä hoitavien sairaaloiden löytämiseksi. Antaa selkeän kuvan siitä kuinka potilasmäärät kohdesairaalassa pienenee kun kriteerit lisäntyy.



Insitellä ei niinkään pystytä havaitsemaan diagnoosien ja lääkehoitojen onnistumisen trendejä tai analysoimaan dataa potilastapahtumien aikajanaa ajatellen. Kliinisen tutkimuksen näkökulmasta Insitea voidaan käyttää alustavaan arvioon, siitä kuinka monta potilasta kohorttiin saataisiin kyseisestä sairaalasta.

Insiten käyttökokemuksista: oli vaikea hahmottaa Insite termejä “at most”(maksimissaan) ja “at least”(ainakin).

Inklusio- ja eksklusio-kriteerien erotelu ja tarkastelu kaipaa selkeytystä käyttöliittymässä.

Monessa kohdassa inhimillisen käyttäjätulkinnan virheen riski on suuri, sillä käyttöliittymä ei ole vielä tarpeeksi selkeä.

Liiketoiminnan malli

Saatavuuspalvelut joita testasimme	Ansaintamalli – liiketoiminnan malli
Insite Platform, Custodix 	Maksajina toimivat tutkimusorganisaatio - Champpion ohjelman aikana sairaala ei maksanut verkostoon liittymisestä. (CRO/Pharma/Akateemikot) jotka saavat kohortin Insiten kautta.
BC R QUEST, BC Platform 	Ansaintamalli on vielä kehitteillä. BC Platformsin edustajan mukaan on todennäköistä, että ansaintamalli on pitkälle sama kuin Eurooppalaisten markkinojen muilla saatavuuspalvelua tarjoavilla toimittajilla.



TrinetX-Ratkaisun testaus ei ollut osana projektia. Osallistuimme 18.10.2018 videotapaamisessamme demoon.

TrinetX tarjoaa aggregoidun datan pohjalta kohortin muodostuksen lisäksi kohorteille suoraan mm. myös statistiikan: esim tunnusluvut, riskisuhteet p-arvoineen ja Kaplan-Meierit yms.

Kokonaisuutena alusta oli vaikuttava verratuna projektissa testaamiimme alustoihin. Sairaallalle järjestelmän implementointi, tekninen ylläpito ja käyttö ei maksa mitään ja sitoutumispakkoa pitkälle ajalle ei ole ja erinlaisia hallintamalleja on helppo toteuttaa osana tätä ratkaisua.

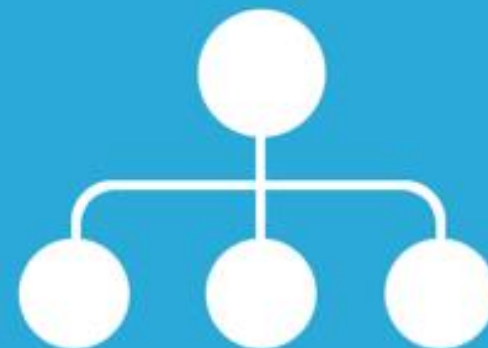
TrinetX on parhaillaan laajentamassa toimintaansa Eurooppaan ja moni toimija joka on ollut mukana investoimassa EHR4CR IMI projektin kautta Custodixin Insite Platform ratkaisuun ovat ottaneet ratkaisut rinnakkain. TrinetX liiketoiminnan mallissa maksajana toimii CRO, Pharma ja tutkimuksen teettäjä.

Voi ostaa oikeuden nähdä kaikki verkoston sairaalat kaikkine terapia-alueinnen tai sitten rajoitetusti ennalta valitut terapia-alueet pienempää kertakorvausta vastaan.

Trinex on investoinut pian 100milj dollaria alustan kehitykseen ja se näkyy käyttöliittymän intuitiivisena suunnitteluna ja lisäpalveluina, jotka auttavat käyttäjää statistisin infograafein muodostamaan potilaskohorin tarjoamista tutkimusmahdollisuuksista selkeämmän kokonaiskuvan.

Mukana on jo 200kpl sairaaloita verkostossa, pääosin USA:n markkinoilta ja 32 pharma yhtiötä.

(Vertailuna EHR4CR IMI projektin budjetti josta Custodixin ratkaisu syntyi oli 16milj €, ja mukana Pharma-firmoja joilla mm. AstraZenecalla ja Sanofilla nyt molemmat järjestelmät käytössään.)





Luokitukset - Koodistot

Lääketieteellisten luokitusten yhdenmukaisella käytöllä mahdollistetaan potilastietojen käsittely algoritmien avulla sekä mm. saatavuuspalvelu, tutkimus- ja tilastointitoiminnan, sekä kansainvälinen vertailu.

Rakenteisempia tietoja syntyy:

- uusien yhä kattavampien **käsiteluokitusten**
- käyttöliittymien luokitukset huomioivien **päivitysten myötä** (mikäli sillataan osa kertomusmuotoisesta potilastiedosta rakenteisiin)
- **tekoälyn & tekstinlouhinnan menetelmien kehittyessä.**

Ydintietojen laadun varmistamiseksi tarvitaan paitsi käyttöliittymä, joka tukee yhdenmukaista kirjaamista, myös hyvät vakiintuneet käytännöt, joiden jalkautuksen merkitys vaikuttaa usein unohtuvan.

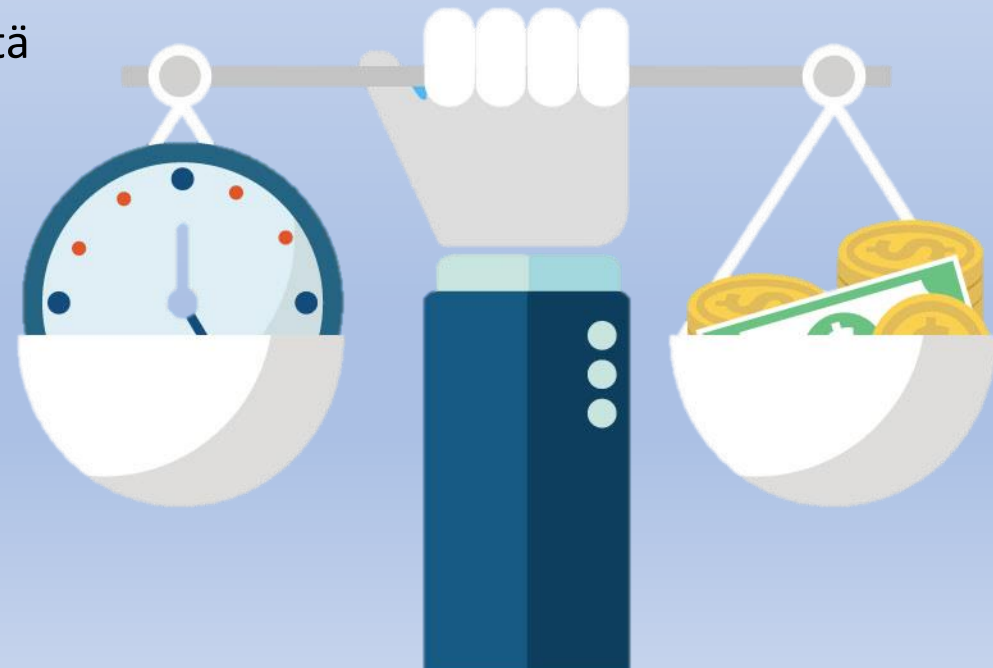


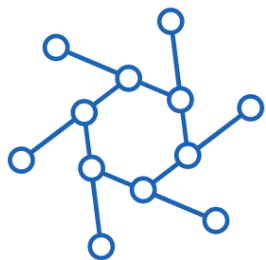
EHR4CR-saatavuuspalvelu ratkaisussa suosittelemme käyttämään lähdejärjestelmien käyttämiä alkuperäiskoodistoja niin pitkälle kuin mahdollista, koska mitä vähemmän tarvitsee silata, sitä parempi laatu (vähemmän virheitä syntyy hakumoottorin tuottamissa kyselyissä).

Toki kansainvälisesti hyödynnettävissä ratkaisuissa aika ja riittävä siltaustarkkuus joka nopeasti hyödynnettävissä useassa paikassa ratkaisee.

Toisin sanoen mitä lähemmällä alkuperäistä rakenteista dataa ollaan, sitä pienempi virhemarginaali tuloksissa ja validointiresurssien allokoinnin tarve.

Tulevaisuudessa implementointi lienee helpottuu, kun koodistojen käyttöä jatkuvasti yhteismitallistetaan paremmin kansainvälisesti yhdenmukaiseksi.





EHR4CR koodistot

SNOMED CT:stä ennustetaan EHR-standardia, mutta tällä (Q3/2018) hankkeessa mukana olevat datakatalogien kehittäjät, BC Platforms ja Custodix, eivät siltaa toimenpiteitä ja laboratorio koodistoja SNOMED CT laadullisista syistä.

Custodix Insite EHR4CR-implementaatio on näistä kahdesta testaamastamme saatavuuspalvelusta pidemmälle kehitetty ja heidän koodistosiltauksensa ovat Aurian tietopalvelussa parhaillaan validoitavina.

TYKS-Lab laboratorio koodistojen mittauksista 82,3% on sillattu Custodixin Insitessa LOINC-koodistoon (Logical Observation Identifiers Names and Codes (**LOINC**))

Toimenpiteitä NCSP:tä ICD10PCS-koodistoon on sillattu 40% kattavuudella. → Muita siltauksia ei ole katsottu tarpeelliseksi toteuttaa tässä Custodix vetoisessa pilotissa huolimatta koodistojen kansainvälisistä käyttöeroista.

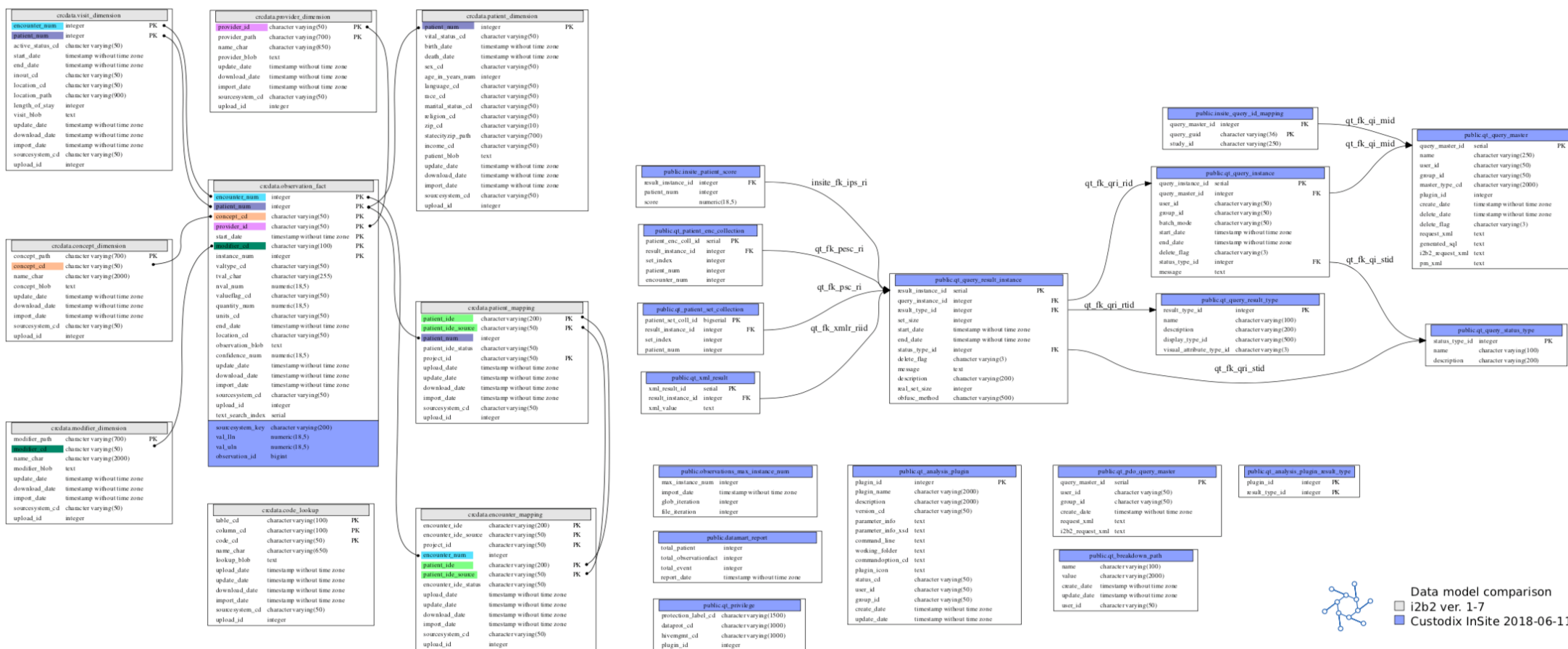
BC|REQUEST saatiin ulkopuoliseen testaus käyttöön vasta syyskuun alussa, jonka takia sen testaus on jäänyt hyvin vähäiseksi. **Insite datakatalogi ei tätä nykyä siltaa toimenpiteitä SNOMED CT:n syystä, että se ei tuo lisähyötyjä koska harva lähdejärjestelmä ilmaisee SNOMED CT:nä muotoisena vielä mitään. (EPIC on poikkeus.)**

YDINTIEDOT

- NÄYTETYYPPI
- TUTKIMUSNIMIKE
- ELINDIAGNOOSIPARIT
([SNOMED](#), KANSALLISESTI ERI MURTEITA JA KÄYTÄNTÖJÄ) - AIKALEIMA
- POTILAAN TUNNISTE
- [ICD-10](#) DIAGNOOSI
- TOIMENPIDEKOODIT ([NCSP](#)→[ICD10 PCS](#))
- LABORATORIO TULOKSET/LOYDOKSET ([TYKS_LAB](#)→[LOINC](#))
- LÄÄKEKOODIT ([ATC-KOODI](#)) + GENEERINEN NIMI/SELITE

I2B2-tietomalli vs. Custodix implementaatio

I2B2 suunniteltu nimenomaan "**kohortin tunnistukseen**" ympäristöön jossa on useita erinlaisia lähdekirjastoja. Toimii myös pseudonymisoidussa ja aggregoidussa tietoympäristössä. Custodixin tietomalli pohjaa I2B2-tietomalliin ja on sen kanssa yhteensopiva.



Palveluoperaattori

Palveluoperaattori on kansallisen saatavuuspalvelu toteutusedellytys.

Saatavuustiedon yhdistämiseen on useita teknisiä tapoja, joita voidaan käyttää myös yhtä aikaa.

Yhdistetty rivitason data

Projektin kokemusten nojalla tapa (3.4 Taulukko 1 kohta 3)

- vaikuttaa toteutuskelpoisimmalta vaihtoehdolta.
- Palveluoperaattorin järjestelmien tietoturvasta huolehditaan erityisen tarkasti.
- Tarjoaa tehokkaan tavan laskea tavanomaisimmat saatavuustilastot tarkasti, sekä mahdollisuuden kehittää tilastoja edelleen.



Datan kerääminen ja yhteismitallistaminen useista lähteistä on vaikeaa ja työlästä, mutta tämä voidaan tehdä vaiheittain, mikä helpottaa työn organisointia.

	Saatavuustilastojen toteutustapa	Toteutus	Hyödyt	Haitat
1	Hajautettu tilastointi	Jokainen rekisterinpitäjä laskee määrämuotoisen tilaston aineistoistaan	Siirrettävä data on täysin anonyymiä	Tilastot voivat vääristyä pahasti, koska rekisterinpitäjiä käsitellään erillisinä ja yksittäisen henkilön tiedot eivät yhdisty
2	Yhdistetty tilastointi	Jokainen rekisterinpitäjä tuottaa tilastointiin riittävän rivitason datan, minkä palveluoperaattori yhdistää	Tilastointi on tarkkaa, koska tieto yhdistyy rekisterinpitäjien yli	Palveluoperaattorin aineisto on rajattu
3	Rivitason datan yhdistäminen	Palveluoperaattori säilyttää kattavaa ja laajaa kopiota rekisterinpitäjien rakenteisesta datasta	Ennalta määrättyjen tilastojen lisäksi voidaan tehdä datan muutakin analyysiä, kuten yksittäisen asiakkaan toivomia erityistilastoja	Palveluoperaattorin keskitetyn järjestelmän pitää olla erittäin huolellisesti suojattu
4	Laskenta pyydettyäessä	Kuten yhdistetty rivitason data, mutta laskenta tehdään lennosta rajapintakutsujen avulla aina uusimmalla datalla	Tarkka ajanmukainen tilasto	Monimutkainen tekninen toteutus ja laskennan hitaus. Soveltuu todennäköisesti vain eräajoihin.

Palveluoperaattori

Ongelmat kiteytyvät seuraavasti:

- (1) datan saantivaikeuksiin eri lähteistä – resurssi/osaaminen
- (2) tietojen yhteismitattomuuteen
- (3) tietoturvavaatimusten ja suuren tallennustilan aiheuttamiin – toisinaan hyvin perusteltuihin – hidasteisiin
- (4) kaupallisen mallin mahdollisiin ongelmiin ja
- (5) loppukäyttäjän kokemaan palvelun toimimattomuuteen
- (6) erityyppisiin lakitulkintoihin käytännön tasolla

Real World Data at your fingertips

Example—>supports the assessments of the safety, efficacy and appropriate use of new medicines.

Supporting Citizen's overall well-being and health

