

## 2.2.2 Koodistojen siltaus kansainvälisiin koodeihin

### Sisällysluettelo

1 SAATAVUUSPALVELUN KOODISTOT.....	2
1.1 Johdanto.....	2
1.2 Saatavuuspalvelun implementoidut koodistot ja niiden siltaukset.....	3
1.3 SNOMED CT -koodiston nykytilanne.....	5
2 SAATAVUUSPALVELUN YDINTIETOMALLI.....	7
3 LYHENTEET.....	9

# 1 SAATAVUUSPALVELUN KOODISTOT

## 1.1 Johdanto

Lääketieteellisten luokitusten yhdenmukaisella käytöllä mahdollistetaan potilastietojen yhdenmukainen tutkimus, tilastointi ja kansainvälinen vertailu. Rakenteisia tietoja syntyy uusien yhä kattavampien käsiteluokitusten, käyttöliittymien, tekoälyn ja tekstinlouhinnan menetelmien kehittyessä. Ydintietojen laadun varmistamiseksi tarvitaan paitsi käyttöliittymä, joka tukee yhdenmukaista kirjaamista, myös hyvät vakiintuneet käytännöt, joiden jalkautuksen merkitys usein unohtuu.

Luokitellun tiedon avulla voidaan tulevaisuudessa yhä paremmin rakenteistaa osa kertomusmuotoisesta potilastiedosta. Suomessa nykyisin käytössä olevista luokituksista mahdollisesti tunnetuin on Maailman terveysjärjestön (WHO) määrittelemä ICD-tautiluokitus (International Classification of Diseases and Related Health Problems), josta tällä hetkellä on käytössä kymmenes versio (ICD-10). ICD-10 tautiluokitus on ollut Suomessa käytössä 22 vuoden ajan saavuttaen vakiintuneen aseman. Luokituksen alkuperäinen käyttötarkoitus sairastavuuden ja kuolemansyiden tilastoinnissa on laajentunut käsittämään myös kliiniset, potilaan hoitoon liittyvät käyttötarkoitukset. Kansalliset hoitosuositukset sekä alueelliset ja paikalliset hoitoketjut on suurelta osin määritelty ja indeksoitu ICD-10 diagnoosien mukaisesti. Luokituksen mukaisia diagnooseja hyödynnetään myös klinisen päätöksentuen järjestelmissä. Suomalainen ICD-10 luokitusversio noudattaa suurimmalta osin WHO:n luokitusta. WHO on aloittanut ICD-11 tautiluokituksen laatimisen. THL seuraa luokitustyötä ja on varautunut osallistumaan siihen.

Uudempi lääketieteellinen käsitejärjestelmä on graafirakenteeseen pohjautuva SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms), jota ylläpitää kansainvälinen järjestö IHTSDO (International Health Terminology Standards Development Organisation). SNOMED CT:tä ei tosin ole vielä käytössä yhdessäkään suomalaisessa terveydenhuollon lähdejärjestelmissä (EPIC aikoo implementoida sen Q3/2018 aikana).

Edellä mainitut koodijärjestelmät on kehitetty diagnoosien dokumentointiin, mutta kehitystyön näkökulmat eroavat toisistaan ja tämän seurauksena koodien käytettävyys saatavuuspalvelun runkona eroaa toisistaan. SNOMED CT pyrkii yhdistämään käsitteistöönsä myös toimenpideluokituksen ja laboratoriotutkimusnimikkeistöt, jotka ovat nykyiseltään Suomen sairaanhoitopiirin käyttämissä lähdejärjestelmissä erilliset koodistot. SNOMED CT -koodiston graafisen ontologian perustellaan mahdollistavan myös vaikeammin toteutettavien hakujen tehokkaamman toteutuksen, sillä konteksti on paremmin läsnä rakenteessa.

Terveydenhuollon organisaatiot kuitenkin myös tietoisesti suunnittelevat ja toteuttavat räätälöityä sisältöä lääketieteen käsitejärjestelmiin eri sairaanhoitopiireissä, jolloin koodistot eivät ole täysin yhteneviä edes kansallisesti, saati kansainvälisesti. Tästä syntyy laadullisesti epäluotettavia päällekkäisiä tietoja, jotka eivät ole verrattavissa. Yksi esimerkki eri murteisiin jakautuneesta koodistosta on Suomessakin käytetty patologian SNOMED 2 -koodisto.

Suomessa potilastietojärjestelmiä varten kulloinkin voimassa oleva luokitus on saatavilla Terveyden ja hyvinvoinnin laitoksen (THL) kansalliselta koodistopalvelimelta, jonne sitä myös päivitetään tarvittaessa. Luokituksen viimeisin, virallinen voimassa oleva versio vietään koodistopalveluun THL:n nimeämän toimitusneuvoston hyväksymisen jälkeen. Toimitusneuvosto koostuu erikoislääkäriyhdistysten edustajista ja lääkäriasiantuntijoista. Suomalaiset luokitukset kuitenkin poikkeavat kansainvälisistä luokituksista jossain määrin ja tietyissä koodistoissa on eroja luokitusten perusteella muodostuneissa ydintiedoissa myös kansallisesti. Eroihin johtaneiden käytäntöjen taustalla on varmasti useita eri syitä, mutta mainittakoon, että koodistojen ja ydintietojen hallinnan merkitystä ei ole laajasti sisäistetty eikä ydintietojen hallintaa ole keskitetty. Siksi koodistojen oikaisu ja validointi kunkin saatavuuspalveluratkaisun yhteydessä on suositeltavaa automatisoitujen (osa-aineisto- eli) kohorttihakujen laadun parantamiseksi. Kliinisen tutkimuksen automatisoituja kohorttihakuja tietyin kriteerein toteutettavissa saatavuuspalveluratkaisuissa voidaan toki tinkiä laadusta, kunhan tutkimusta tekevät organisaatiota kokevat automaattisten saatavuuskyselyiden nopeuttavan prosessia, jossa tutkimusten asettelua vasta mietitään. Saatavuustietokanta on parhaimmillaan reaaliaikainen potilastiedon näyteikkuna kliinisen tutkimuksen tekijöille, kuten kansainvälisille lääkeyrityksille, tutkimusta koordinoivalle yritykselle (Contract Research Organisation, **CRO**) tai akateemisille tutkijalle.

## 1.2 Saatavuuspalvelun implementoidut koodistot ja niiden siltaukset

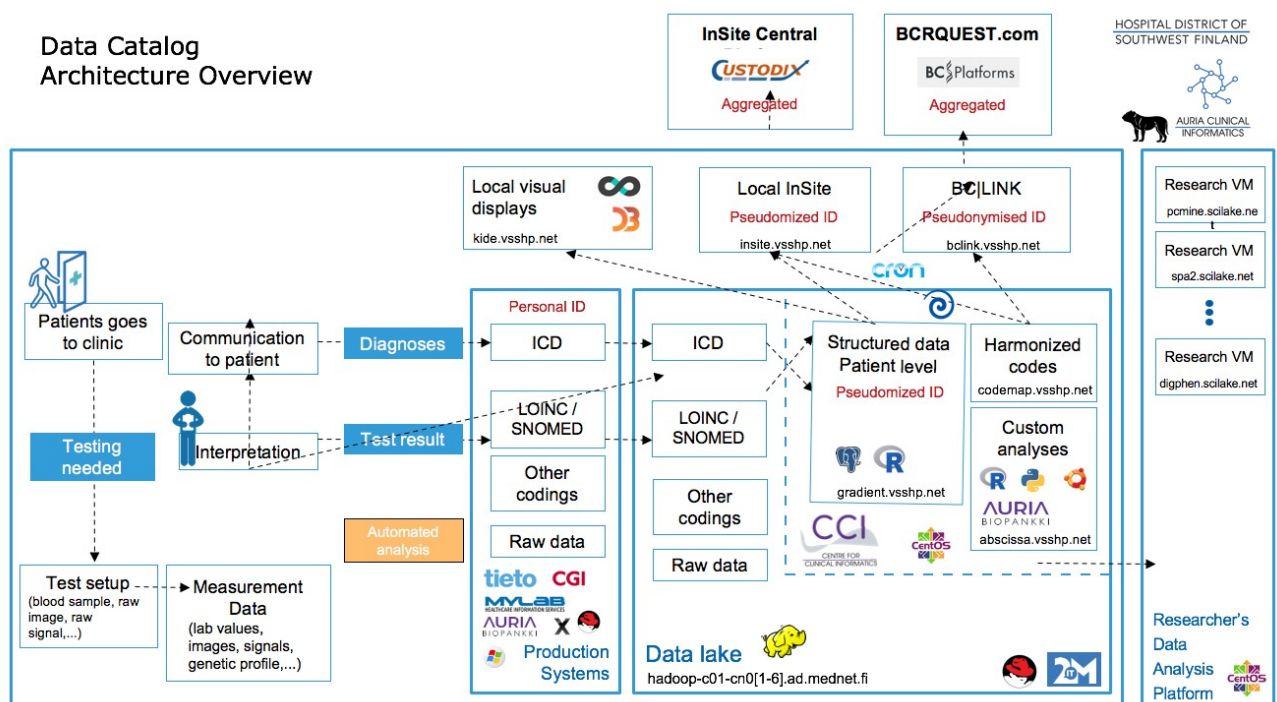
Saatavuuspalvelua toteutettaessa suosittelemme käyttämään lähdejärjestelmien käyttämiä alkuperäiskoodistoja niin pitkälle kuin mahdollista. Mitä vähemmän koodeja tarvitsee sillata, sitä vähemmän virheitä syntyy hakumoottorin tuottamissa kyselyissä. Toisin sanoen, mitä lähemmällä alkuperäistä rakenteista dataa ollaan, sitä pienempi virhemarginaali tuloksissa.

Tulevaisuudessa tilanne helpottunee, kun koodistojen käyttöä jatkuvasti yhteismitallistetaan paremmin kansainvälisesti yhdenmukaiseksi. Luokitusten ja rakenteellisen tiedon harmonisointia eteenpäin vievänä voimana järjestelmäkehityksen lisäksi ja katalyyttinä toimii mm. tekoälyn hyödyntämisen tarve hoidossa.

Esimerkkinä voimme käyttää Custodix Insite EHR4CR-ratkaisussa hyödynnettyjä koodistosiltauksia, jotka ovat Aurian tietopalvelussa parhaillaan validoitavina. Insitessa siltaukset on

tehty ainoastaan TYKS-Lab laboratorio koodistojen mittauksista (82,3% on sillattu Custodixin Insitessa) LOINC-koodistoon (Logical Observation Identifiers Names and Codes) ja toimenpiteitä on sillattu Suomen toimenpideluokituksesta, joka pohjautuu Pohjoismaiseen toimenpideluokitukseen NCPS:ään. NCSP:tä on siis sillattu ICD10PCS-koodistoon on 40% kattavuudella suhteessa erilaisiin toimenpiteisiin. Muita siltauksia ei ole katsottu tarpeelliseksi toteuttaa tässä Custodix-vetoisessa pilotissa huolimatta koodistojen kansainvälisistä käyttöeroista. BC R|QUEST -tuote saatiin ulkopuoliseen testauskäyttöön vasta syyskuun alussa, jonka takia sen testaus kyseisen saatavuuspalvelun päällä on jäänyt vähäiseksi. Insite datakatalogi ei tätä nykyä siltaa toimenpiteitä SNOMED CT koodistoon, koska se ei tuo lisähyötyjä: Harva lähdejärjestelmä ilmaisee SNOMED CT:nä muotoisena vielä mitään. (EPIC on poikkeus tästä.)

Kansainvälisiin referenssiterminologioista, joita EHR4CR ratkaisussa hyödynnetään on syytä mainita keskeisimmät: ICD-10, SNOMED-CT, ICD-O(onkologia) ,NCSP(toimenpide), ATC(lääkkeet), ja LOINC (laboratorio).



**KUVA 1.** VSSH:n saatavuuspalvelutestauksen mahdollistava arkkitehtuuri. Harmonisoidut koodistot on sovittu jaettavan yhteisen tietokannan (codemap.vssh.net) kautta.

### 1.3 SNOMED CT -koodiston nykytilanne

SNOMED CT: ssa termien ja käsitteiden väliset suhteet eli noodit ovat paremmin rakenteistettuja ja soveltuvat hyvin koneellisesti tulkittavaksi. SNOMED CT:n hyödyntäminen voi jäädä puutteelliseksi, mikäli SNOMED CT:n tarjoamia mahdollisuuksia ja sen topologiaa ei kunnolla ymmärretä. SNOMED CT:n kliiniset termit on suunniteltu käytettäväksi ohjelmistosovelluksissa, kuten sähköisessä potilastietokannassa, päätöksenteon tukijärjestelmissä ja tukemassa tietojen sähköistä tiedonsiirtoa eri klinisten sovellusten välillä. On siis siirretty termistöön, missä luetellaan lääketieteellisiä käsitteitä ja niiden suhteita. Sen suunnittelijoiden tavoite on, että SNOMED CT:stä tulee terveydenhuollon hyväksytty kansainvälinen terminologinen resurssi, joka tukee yhteisten käsitteiden monikielisiä terminologisia esityksiä. SNOMED CT:n terminologiaa voidaan laajentaa, mutta sen käyttäminen ilman tietokonetta ei ole mielekäästä. Samoin rakenne ei enää seuraa automaattisesti.

THL teki 26.05.2018 myönteisen päätöksen Suomen valtiokohtaisen lisenssin hankkimisesta. Tämä on myönteinen uutinen, sillä julkisen terveydenhuollon patologian laboratorioilta

on puuttunut diagnoosi- ja elin-rekisterien käyttämä ajantasainen yhteinen koodisto. SNO-MED koodiston voi toki ladata käyttöönsä tuotekehitystarpeita ajatellen ilman lisenssiäkin. SNOMED CT on tarkoitettu ensisijaisesti sisällöllisen yhteen toimivuuden varmistamiseen määrittelemällä yhteiset käsitteet ja termit (käytetään järjestelmien sisäisesti ja järjestelmien välisessä integraatiossa). Maalisenssin maksamalla saa käyttöönsä myös vastaavuusjoukot ("map reference sets"), esim SNOMED CT , ICD-10, ICD-O3, ICPC-2, SNOMED 2 → SNOMED CT. Kirjoitushetkellä emme ole vielä saaneet kattavia siltauksia käyttööme, sillä maalisenssiä ei ole vielä maksettu.

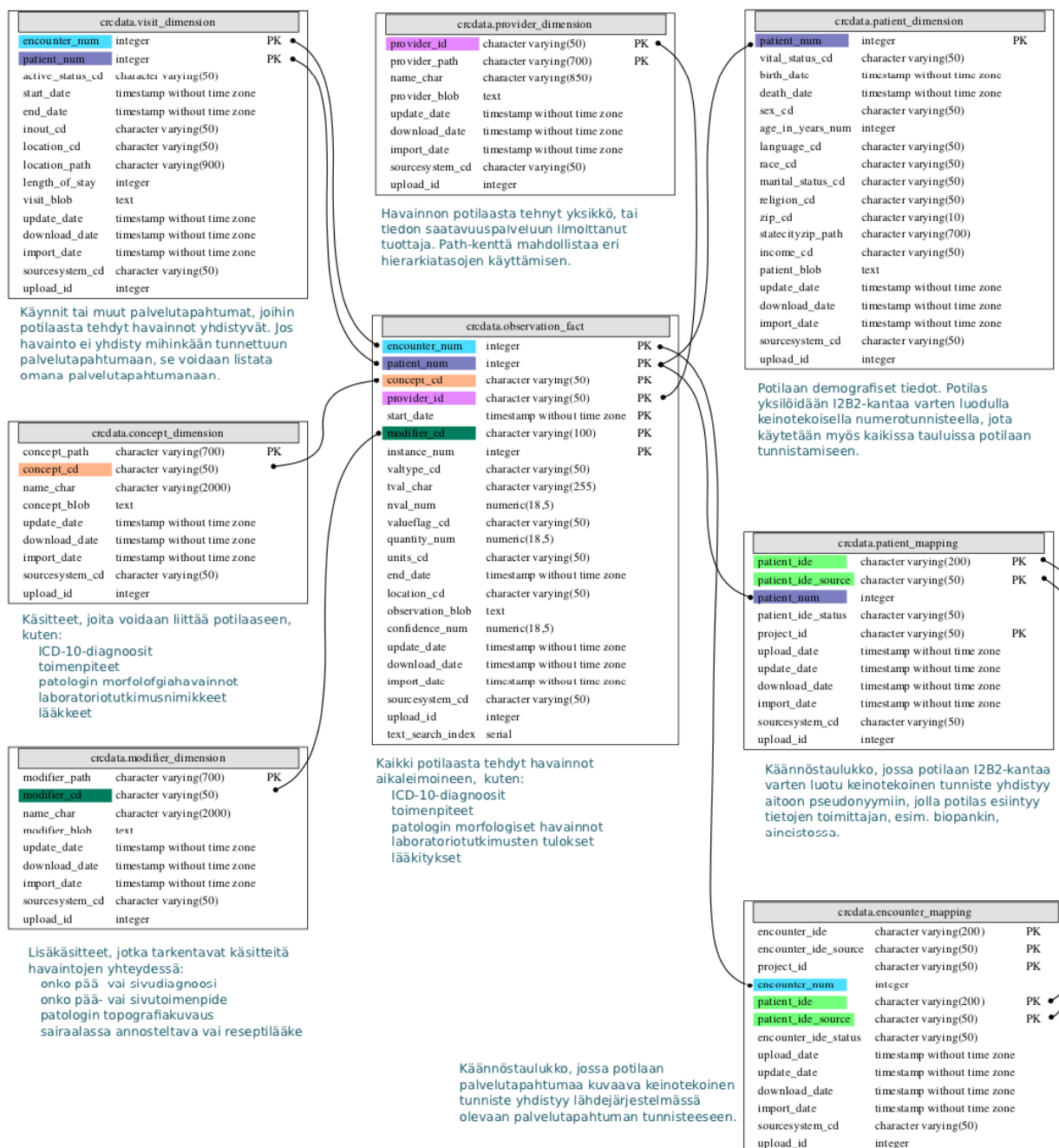
SNOMED CT:stä ennustetaan EHR-standardia, mutta tällä (Q3/2018) hankkeessa mukana olevat datakatalogien kehittäjät, BC Platforms ja Custodix, eivät siltaa toimenpiteitä ja laboratorio koodistoja SNOMED CT -koodeihin laadullisista syistä. Custodix Insite EHR4CR-implemентаatio on näistä kahdesta testaamastamme saatavuuspalvelusta pidemmälle kehitetty ja heidän koodistosiltauksensa ovat Aurian tietopalvelussa parhaillaan validoitavina. TYKS-Lab laboratoriokoodistojen mittauksista 82,3% on sillattu Custodixin Insitessa LOINC-koodistoon (Logical Observation Identifiers Names and Codes) ja 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 R|QUEST 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 saatavuuspalvelukäytössä.

Varsinais-Suomen Sairaanhoidopiirin Auria tietopalveluissa (<https://auria.fi>) kehitetään tällä hetkellä suomalaisten SNOMED-murteiden siltaamistyökalua. Tarkoitus on kehittää syksyn 2018 aikana työkalu, jolla saadaan kunkin paikkakunnan patologiedustajat yhdistämään kunkin sairaanhoidopiirin patologian laboratorion käyttämät nyky SNOMED-murteet SNOMED CT:n (<http://browser.ihtsdotools.org>) siltaukseen ja validoimaan työn tulokset kansallisesti. SNOMED CT:n sisältämää hierarkiaa hyödyntämällä saamme esiin tarkimman yhteisen koodiston laboratorioiden kesken. Siltaamistyökalu tuotetaan Apache 2-lisenssillä ja se jaetaan Aurian tietopalvelut GitHub-organisaatiotilin (<https://github.com/auriadih>) kautta hyödynnettäväksi myös muihin koodistosiltaustöihin.

## 2 SAATAVUUSPALVELUN YDINTIETOMALLI

### YDITIEDOT SAATAVUUSPALVELUSSA

- NÄYTETYYPPI
- TUTKIMUSNIMIKE
- ELINDIAGNOOSIPARIT (SNOMED, KANSALLISESTI ERI MURTEITA JA KÄYTÄNTÖJÄ)
- AIKALEIMA
- POTILAAN TUNNISTE
- ICD-10 DIAGNOOSI
- TOIMENPIDEKOODIT (NCSP ICD10 PCS)
- LABORATORIO TULOKSET/LÖYDÖKSET (TYKS\_LAB LOINC)
- LÄÄKEKOODIT (ATC-KOODI) + GENEERINEN NIMI/SELITE



**KUVA 2 I2B2-TIETOMALLIN KUVAUS. TIETOMALLIN AVULLA YDINTIETOJA KUSTAKIN LÄHDEJÄRJESTELMÄSTÄ (ERI LUOKITUKSIA NOUDATTAEN) VOIDAAN DATAKATALOGIRATKAISUSSA ESITTÄÄ TEHOKKAASTI. TÄMÄ TIETOMALLI MAHDOLLISTAA MYÖS KANSALLISTEN HARMONISOITUIJEN OHJEIDEN LUONNIN DATAN TUOTTAJILLE SAATAVUUSPALVELUA VARTEN.**



### **3 LYHENTEET**

**ATC** (LÄÄKETIEDOT), ANATOMICAL THERAPEUTIC CHEMICAL

**EHR4CR**, ELECTRONIC HEALTH RECORDS FOR CLINICAL RESEARCH

**ICD-10** TÄMÄN HETKINEN TAUTILUOKITUS

**ICD**, INTERNATIONAL CLASSIFICATION OF DISEASES AND RELATED HEALTH PROBLEMS

**NCSP**, POHJOISMAINEN LEIKKAUSLUOKITUS

**NCSP**, NOMESCO CLASSIFICATION OF SURGICAL PROCEDURES

**SNOMED**, SYSTEMATIZED NOMENCLATURE OF MEDICINE

**SNOMED CT**, SYSTEMATIZED NOMENCLATURE OF MEDICINE CLINICAL TERMS