



**Anne Pitkäranta**

professori, tutkimusjohtaja  
HUS



**Mikko Rotonen**

IT-kehitysjohtaja  
HUS



**Visa Honkanen**

dosentti, kehittämisjohtaja  
HUS

HUS-tietoallas on Big Data-järjestelmä, joka kykenee käsittelemään suuria tietomääriä ja mahdollistaa suurten tietomäärien analytiikan. Tietoaltaalla on monia ulottuvuuksia. Se on tarkoitettu toimimaan osana potilastietojärjestelmiä, tukemaan tiedolla johtamista sekä erilaisia lääke- ja terveystieteen tutkimushankkeita tarjoamalla tiedon tallennukseen, hallintaan ja analysointiin sopivia työkaluja.

# HUS-tietoallas tiedolla johtamisen ja tutkimuksen tukena

**T**ietoallas tarjoaa tehokkaan alustan sovellus- ja raportointikehitykseen ja se tukee operatiivisen tuotannon tarpeita tekoälypalveluiden taustajärjestelmien avulla. Tietoaltaan dataan saatava näkymä antaa työkaluja potilaskohtaisen, potilasryhmäkohtaisen ja väestötasoisennalta ehkäisevän hoidon suunnitteluun.

## HUS-tietoallas on Big Data-järjestelmä

Terveystieteen suurissa organisaatioissa on jo vuodesta 1990 ollut käytössä tietovarastoja, joihin on siirretty eri potilastietojärjestelmistä ensisijaisesti erilaista hallinnollista tietoa. HUS Tietohallinto käynnisti vuonna 2015 kehittämishankkeen, jonka tavoitteena oli kehittää laajojen tietoaaineistojen käsittelyyn kykenevä tietovaranto, joka olisi tietoallas-tyyppinen ratkaisu. Kun Sitra käynnisti Isaacus-hankkeen, niin HUS esitti Sitralle, että tietoallas toteutettaisiin yhteistyössä. Perusperiaatteena on tietoaltaan tietotekniikan avoin lähdekoodi ja tietoaltaan lisensoiminen avoimen lähdekoodin periaatteella pilvipohjaisena ratkaisuna.

HUS-tietoallas on tietoteknisesti Big Data-järjestelmä, joka kykenee käsittelemään suuria tietomääriä ja mahdollistaa suurten tietomäärien analytiikan.

Tietoallas on CE-merkitty ja osa HUSin potilastietojärjestelmäkokonaisuutta. HUS Tietoallas täyttää samat tietoturva- ja GDPR-vaatimukset kuin muutkin HUSin käyttämät potilas- ja hallinnon tietojärjestelmät, ja sille on teetetty GDPR:n mukainen riskien arviointi.

Tietoaltaan rekisterit ovat käytävissä samoin periaattein kuin alkuperäisten järjestelmien rekisterit. HUS-tietoaltaassa ei yhdistetä eri rekistereitä, vaan niiden aineistot ovat lähtökohtaisesti teknisesti erillään toisistaan. Tietoaltaassa olevat aineistot on säilytetty pseudonymisoituina, eli henkilötietoja ei voida käsittelyn jälkeen enää yhdistää tiettyyn henkilöön ilman lisätietoja. Aineistossa oleva sosiaaliturvatunnus on salattu (kryptattu) yhtenevästi ja kaikki muut henkilötiedot, kuten nimet ja yhteystiedot, on poistettu. Salattun ”pseudohetun” avulla kaikki HUS-tietoaltaan aineistot ovat yhdistettävissä keskenään.

HUS-tietoallas on HUS Tietohallinnon palvelu HUS-yhtymähallinnolle, joka toimii potilastietojärjestelmien rekisterinpitäjänä ja sairaanhoitopiirin alueella toimivien julkisen sektorin terveydenhuollon yhteisrekisterin rekisterinpitäjänä. HUS-tietoaltaan avulla eri rekistereiden tiedot ovat myös tutkijoiden hyödynnettävissä. Hallinnon

Taulukko 1.

## HUS Tietoaltaan sisältämät aineistot

- Potilastietorekisterit
- Laboratoriotutkimukset
- Leikkaukset
- Patologiset tutkimukset
- Teho-osastot ja anestesiatiiedot
  - Teho-osastoilla ja vastaavissa yksiköissä/erikoisaloilla käytössä oleviin järjestelmiin kirjatut ja tallennetut tiedot, kuten tapahtumat, toimenpiteet, monitoriarvot jne.
  - Myös lasten ja vastasyntyneiden teho-osastojen aineistot löytyvät näistä tietokannoista.
- BCB-laaturekisterit
  - yli 80 valtakunnallisesti käytössä olevan BCB-laaturekisterin aineistot HUSin potilaiden osalta.
- Kuvantaminen
- HUS Tietovaraston raportointiaineistot
- Ensivasteen aineistot
- Synnytystietojärjestelmän tiedot

Taulukko 2.

## HUS-tietoallas sisältää valtavan määrän dataa.



3,5 miljoonaa henkilöä



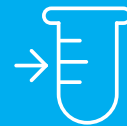
2,7 miljoonaa hoitokertomusta



37 miljoonaa potilaskertomusta



80 miljoonaa potilaskertomustekstiä



790 miljoonaa laboratoriotulosta

ja tutkijoiden lisäksi HUS-tietoaltaan asiakkaita voivat olla myös HUS-alueen kunnat, kuntayhtymät ja ERVA-alueen sairaanhoitopiirit, kukin omana rekisteripitäjänään.

### Mitä aineistoja HUS-tietoaltaasta löytyy?

Tähän mennessä HUS Tietoaltaaseen on integroitu HUSin kliinisessä käytössä olevat potilastietojärjestelmät (taulukko 1, kuva 1). Näiden järjestelmien integraatioihin sisältyvät jatkuvat päivitykset saapuvat HUS Tietoaltaaseen vähintään kerran vuorokaudessa. Lisäksi vanhojen käytöstä poistuneiden potilastietojärjestelmien aineistoja tuodaan HUS Tietoaltaaseen kaiken aikaa. HUS-tietoaltaassa on valtavasti dataa (taulukko 2), jonka määrä kasvaa koko

ajan. Tietoaltaaseen viedään myös potilasmonitorien tuottamaa elintoimintojen mittaustietoa.

### Tietoallas-kokonaisuus

Tietoaltaan lisäksi kokonaisuuteen kuuluu HUS-tietoallaspalvelu, joka sisältää tietoaltaan, tuen, käyttöpalvelut ja tietoallas-asiantuntija- ja sovelluskehityspalvelut. Tutkijat voivat hankkia korvausta vastaan HUS Tietoallas -ympäristöstä oman analytiikkatyötilan, jossa tutkimusluvan mukainen aineisto on siirretty valmiiksi omaan erilliseen sivutietoaltaaseen. Analytiikkatyötiloja on saatavilla erilaisilla laskentaresursseilla ja analytiikkatyökaluilla (Python, R tai Spark), joihin on pääsy vain kyseisen tutkimuksen tekijöillä. Sekä HUSin työntekijät että ulkoiset käyttäjät voi-

vat hakea analytiikkatyötilaan käyttöoikeudet. Käyttövaltuuksien myöntämisen jälkeen tutkijat pääsevät omiin analytiikkatyötiloihinsa. Tutkimuksen päättyessä analytiikkatyötila suljetaan.

### Tietoallas tiedolla johtamisen tukena

HUS käyttää tietoaltaan tietoja hallinnollisessa raportoinnissa. Tietoaltaan tiedoista muodostetaan tietovarasto (infokanta), jossa on sovitulla tarkkuustasolla erilaisia tietoaineistoja käytössä olevia eri raportointivälineitä varten (Hus Total ja PowerBI). Tiedot siirtyvät tärkeimpiin johtamisessa käytettäviin raportointijärjestelmiin vähintään päivätasoisesti. Osa tietoaltaan tiedoista raportoidaan 15 minuutin välein tilannekuvaraportointia varten.

## Tietoallas tutkijan tukena

Tutkimusryhmille luovutetaan aineistoja tutkimuskäyttöön HUSin tutkimusluvan sallimassa laajuudessa aivan samoin periaattein kuin muussakin tutkimuksessa. HUS Tietoaltaan kailta tutkijoilta edellytetään aina tutkimuslupa. Rekisteritutkimuksiin riittää tutkimuslupa, mutta jos aineistoa tarvitaan muuhun kuin rekisteritutkimukseen, on tutkimukselle haettava myös eettisen toimikunnan puoltava lausunto.

» **Tietoaltaiden avulla voidaan potilaiden hoidosta kertyvää valtavaa tietomassaa hyödyntää kliinisessä päätöksenteossa ennennäkemättömällä tavalla.**

Tietopyynnöissä luovutettava henkilötunniste ei kuitenkaan ole HUS-tietoaltaan pseudohetu, vaan kutakin aineistoluovutusta varten muodostetaan uusi tutkimuskohtainen pseudo-

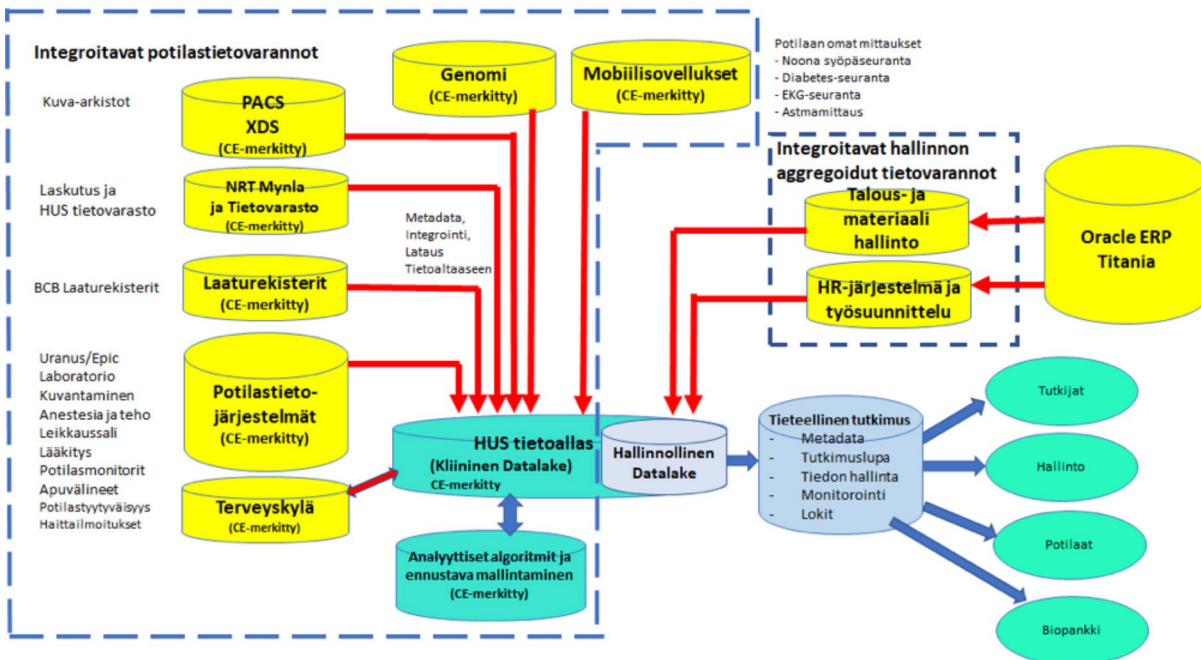
hetu, jolla voidaan yhdistää ainoastaan kyseisen tutkimusluvun alaisia aineistoja. Aineistoluovutuksien yhteydessä on myös mahdollista poikkeustapauksessa saada mukaan potilaan oikea henkilötunniste. Tällöin asia tulee kuitenkin olla selvästi mainittu tutkimusluvassa ja tutkimusluvun myöntäjä harkitsee erityisen huolellisesti tällaisen luvan myöntämistä. Tarvittaessa poimintakohortti voidaan muodostaa myös henkilötunnuslistauksen perusteella.

Tutkimuslupahakemuksessa pitää määritellä mahdollisimman tarkkaan, miten potilaskohortti muodostetaan ja mitä aineistoja hakemus koskee. Tästä syystä tutkijan on pyydettävä esiselvi-

tys. Esiselvityksessä varmistetaan siitä, että toivottu aineisto on saatavilla, ja että tutkija saa määriteltyä tietopyynnön riittävällä tarkkuudella tutkimuslupahakemusta varten.

## Lopuksi

Tietoaltaiden avulla voidaan potilaiden hoidosta kertyvää valtavaa tietomassaa hyödyntää kliinisessä päätöksenteossa ennennäkemättömällä tavalla. Tiedon hyödyntäminen on vielä alussa, mutta tekniikan, menetelmien, osaamisen ja tietoturvan kehittymisen myötä tulevaisuuden näköalat ovat huikeat.



Kuva 1.

Tietoaltaalla on integroituna eri CE-sertifioiduista potilastietojärjestelmistä, joka on tuotu kunkin potilastietojärjestelmän käyttämissä tietomallissa ja tietoformaateissa.