



Parempia päätöksiä modernilla, automatisoidulla tietoaalustalla

1. D# (aiemmin SmartDW)

D# (lausutaan "dee sharp") sisältää **menetelmän** ja sitä tukevat **työkalut** ketterään tietovarastokehitykseen. Menetelmä pohjautuu toiminnan käsitteiden mallintamiseen tarvelähtöisesti. Käsittemalli on samalla yhteinen kieli kaikille toimijoille, tärkeä määrittelydokumentti ja lähtökohta tietovaraston tekniselle ratkaisulle.

D#:n tärkeänä työkaluna on **SmartEngine**, jonka avulla automatisoidaan merkittävä osa tietovaraston kehitystyöstä. Käytännössä SmartEngine generoi käsitemallin perusteella tietovaraston rakenteiden, latausmekanismeihin sekä rajapintoihin tarvittavan standardi SQL-koodin täysin automaattisesti.

2. Tietoaalustan yleinen arkkitehtuuri

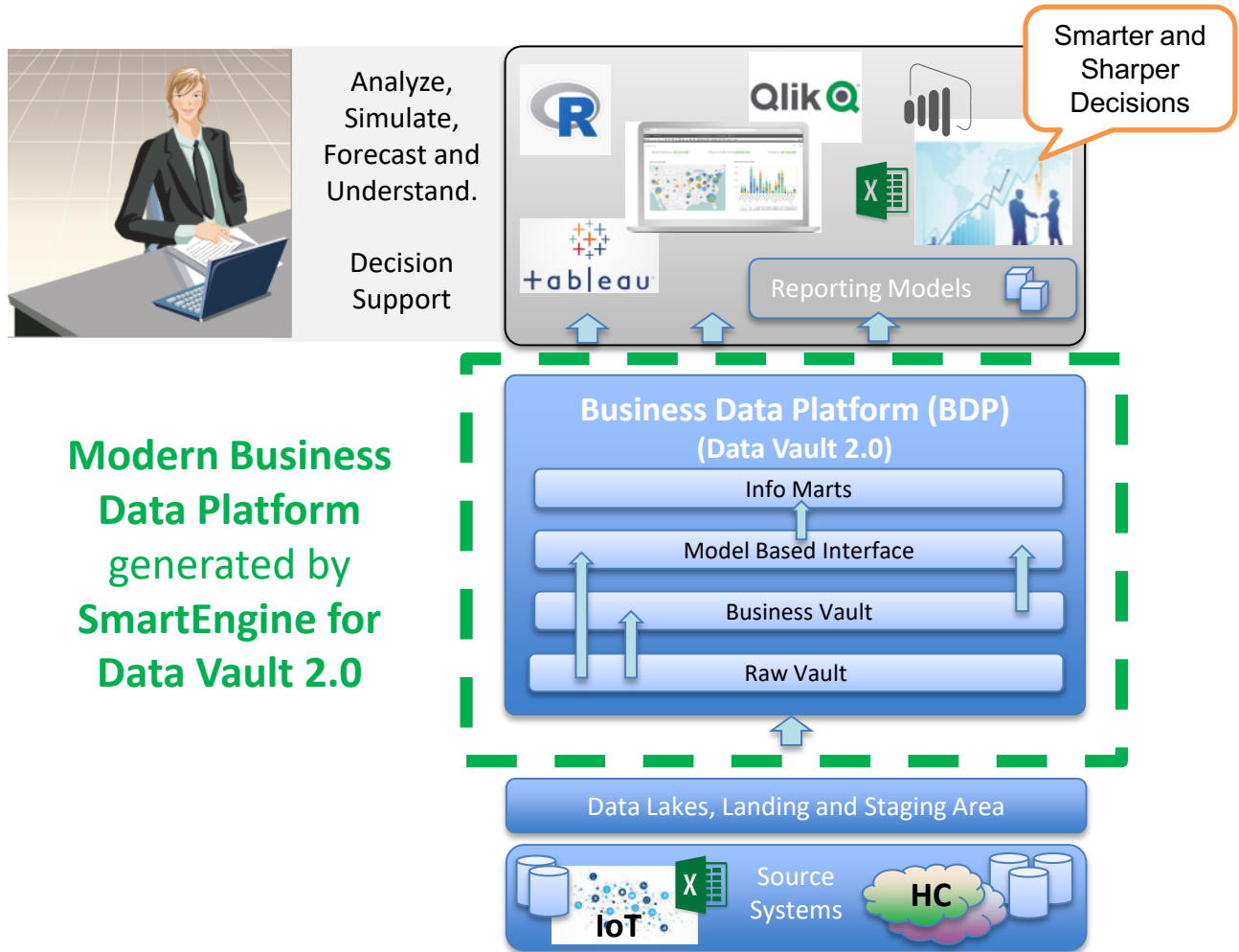
Modernien tietovarastojen de facto standardiksi on muodostunut Data Vault 2.0 -malli. Data Vault 2.0 on erityisesti tietovarastojen ketterään kehittämiseen luotu menetelmä, joka sisältää sekä kehitysproessin että tietokannan rakenteen mallinnustavan. Lopputuloksena syntyvä varsinainen tietokantarakenne on optimoitu tiedon tallentamiseen alusta asti ratkaisun laajentamisen näkökulmasta. Raportoinnin näkökulmasta menetelmä tukee myös standardeiksi muodostuneiden tähti- ja lumihuutalemalleihin perustuvia raportointikokonaisuuksia, eikä rajaa millään tavoin, miten tietovarastoon vietyä tietoa voi hyödyntää.

Data Vault -kokonaisuuden olennaiset tasot ovat Raw Vault ("Tiedon tallennus"), Business Vault ("Liiketoimintalogiikka") sekä Info Mart ("Julkaisurajapinta"):

- *Raw Vault* muodostaa varsinaisen tiedon tallennuskerroksen, ja sen muodostamisperiaatteet on tarkasti määritelty Data Vault 2.0-standardissa.
- *Business Vaultin* tarkoitus on tarvelähtöisesti puhdistaa ja jalostaa Raw Vault -tasolle tallennettu data käyttökelpoiseksi, soveltaen asiakkaan johdon, liiketoiminnan avainhenkilöiden ja data-ammattilaisten kanssa määriteltyä liiketoimintalogiikkaa. Tähän tasoon ei tallenneta mitään pysyvää, vaan se sisältää pelkästään raakadatasta johdettua tietoa, joka voidaan aina generoida uudestaan. Business Vault -taso elää tietovaraston elinkaaren aikana, tarjoten tallennetusta tiedosta aina sen halutun ja ajantasaisen tulkinnan.
- *Info Mart* on tietovaraston julkinen tiedon julkaisurajapinta. Tähän toteutetaan kaikki täsmäraportointiratkaisut sekä esim. tietovarastosta julkaistavien muuhun kuin raportointitarkoituksiin tarkoitettujen tietojen rajapinnat.



Yksinkertaistettu sovellusarkkitehtuuri on esitetty alla olevassa kuvassa.



**Modern Business
Data Platform
generated by
SmartEngine for
Data Vault 2.0**

3. Tietoalustan automatisointi SmartEngine-työkalulla

Data Vault 2.0 -standardin mukaisen tietovaraston toteutukseen tarvittava SQL-koodi generoidaan käsitelmällin pohjalta täysin automaattisesti. SQL-koodin generointiin käytetään Cerion SmartEngine automaatiovälinettä. Tästä seuraa merkittäviä laadullisia ja taloudellisia hyötyjä, koska normaalisti virhealtis manuaalinen työ on korvattu automaatiolla. Lisäksi työkalu abstraktoi pois monimutkaisen Data Vault 2.0 -rakenteen ja julkaisee tietovaraston sisällön kehittäjille tehokkaassa ja ymmärrettävässä muodossa, mikä käytännössä sekä yksinkertaistaa että nopeuttaa kaikkea kehitystyötä. Automaation muita hyötyjä ovat henkilösidonaisuusien väheneminen ja lopputuloksen johdonmukaisuus ja ennustettavuus.



D#-menetelmä pohjautuu UML-luokkakaavioilla tai ER-kaavioilla mallinnettuun käsitemalliin. Mallissa käytetään ei-tekniisiä päivittäistoiminnassa käytettäviä termejä siten, että liiketoiminnan asiantuntija tunnistaa arkipäivänsä kaavioista. Käsitemallia voidaan käyttää kommunikaatiovälineenä sekä tiedon toimittajan suuntaan (kehitysvaiheessa ”tätä tietoa tarvitsemme”) sekä oman organisaation asiantuntijoiden/tiedon tarvitsijoiden suuntaan (”tämä tieto meillä on tarjota teille, miten haluatte sitä käyttää”).

Käsitemallit voidaan luoda esim. Visual Paradigm -työvälineellä. Visual Paradigm on edullinen ja helppokäyttöinen mallinnusväline, joka tukee tietovaraston automaattista generointia SmartEngine-työvälineellä. D#-menetelmä tukee myös Ellie-työvälinettä, eli haluttaessa käsitemallinnus voidaan tehdä myös Elliellä. Käsitemallin perusteella SmartEngine generoi kaiken tietovarastorakenteisiin, latausproseduureihin ja latausproseduurien orkestrointiin tarvittavan SQL-koodin automaattisesti. Käytännössä käsitemalli viedään mallinnusvälineestä SmartEnginelle, joka tämän perusteella suorittaa edellä mainitun generoinnin. Kaikki SmartEnginen generoima SQL-koodi (taulut, näkymät, funktiot, proseduurit) asennetaan esim. Azure SQL Database - palveluun.

D#:n erityispiirre on, että se luo automaattisesti käsitemallin mukaisen kehitysrajapinnan Business Vault ja Info Mart -tasojen väliin, jolloin infomarttien kehittämisessä hyödynnetään pelkästään tätä Kehitysrajapintaa. Tämä nopeuttaa kehitystä huomattavasti verrattuna siihen, että tähtimallien kehittämisvaiheessa kyselyt kohdistettaisiin hajautettuun Data Vault 2.0 -rakenteisiin (sekä Raw että Business Vault -rakenteisiin). Kehittäminen ei näin ollen edellytä asiantuntijoilta syvällistä Data Vault -osaamista. Kehitysrajapinta mahdollistaa myös esim. pääkäyttäjille koko tietovaraston datan hyödyntämisen (ilman infomartissa tehtyjä rajoituksia ja logiikkaa) ilman Data Vault -osaamista.

SmartEngine-työkalun perusominaisuuksiin kuuluu käsitemallin laajennettavuus ja laajennetun mallin pohjalta tietokantarakenteen päivittäminen vastaamaan muutettua mallia. Koska D#-menetelmässä kehittäminen pohjautuu mallintamiseen, pysyy itse käsitemalli ja siitä johdettava muu dokumentaatio aina ajan tasalla.

Kumppanimme Cerion Solutions Oy:n kaikki tietovarastotoimitukset on tehty D#-menetelmän mukaisesti vuodesta 2009 lähtien. Viimeinen suuri kehitysaskel otettiin vuonna 2018, kun välineeseen lisättiin tuki Data Vault 2.0 -mallin mukaisten tietovarastojen generointiin.