Hajautetut sovellukset -seminaari Syksy 2002
Tietojenkäsittelytieteen laitos
Helsingin yliopisto
Jenni Löytynoja 27.9.2002

Datan integrointi Common Warehouse Metmodelin (CWM) avulla

Suuri osa erilaisten organisaatioiden tiedoista on tallennettu sähköisessä muodossa tietokantoihin. Tällaisten suurien tietomäärien integrointi, yhtenäiseksi kokonaisuudeksi yhdistäminen, auttaa tehostamaan organisaation toimintoja. Usein tallennettu data on kuitenkin hallitsematonta, minkä seurauksena sitä on vaikea paikantaa ja käyttää. Tietovarastoinnin avulla tallennettu data voidaan integroida hyödylliseksi ja luotettavaksi tietolähteeksi.

Tietovarastojen rakentamiseen, ylläpitoon ja käyttöön tarvitaan metadataa. Metadata kuvaa tietovarastoon tallennettavaa tietoa, kuten sen merkitystä, muotoa ja rakennetta. Tietovarastoinnissa ongelmaksi muodostuu se, että organisaation eri järjestelmät käyttävät usein erilaista metadataa eikä tietoja näin ollen voida siirtää suoraan järjestelmästä toiseen.

Eri muotoa olevat metadatat käyttävät erilaisia metadatamalleja, joiden avulla määritellään metadatan muoto. Datan integroimiseksi tarvitaankin eri työkalujen välillä sopimus metadatan esitysmuodosta. Common Warehouse Metamodel eli CWM määrittelee kehykset metadatan esitysmuodolle ja syntaksin tietovarastoinnin operaatioille. Näin CWM helpottaa metadatan siirtoa tietovarastoinnin eri työkalujen, alustojen sekä metadatavarastojen välillä ja mahdollistaa organisaation datan integroinnin.

CWM on OMG:n (Object Management Group) standardi, joka perustuu pääpiirteissään kolmeen muuhun OMG:n standardiin. Meta Object Facility (MOF) määrittelee kehykset metamallien kuvaamiseen ja tietovarastojen käyttöön, Unified Modeling Language (UML) on monipuolinen olioperustainen mallinnuskieli, jonka avulla kuvataan CWM:n metamalli ja XML Metadata Interchange (XMI) mahdollistaa metadatan siirron XML -muodossa. Seminaarityössä tullaan perehtymään CWM -malliin ja sen suhteeseen datan integrointiin.

Lähteet

OMG. *Common Warehouse Metamodel (CWM) Specification*, helmikuu 2001.

John D. Poole. *Model-Driven Architecture: Vision, Standards and Emerging Technologies*, huhtikuu 2001.

Doug Tolbert. *CWM: A Model-based Architecture For Data Warehouse Integration*, toukokuu 2000.

Daniel T. Chang. *CWM Enablement Showcase: Warehouse Metadata Interchange Made Easy Using CWM*, maaliskuu 2001.

Sridhar Iyengar. *CWM Audio Briefing: The Key to Integrationg Business Intelligence*, ?, 2000.