

XML

- XML voidaan nähdä rakennetason mallina puolirakenteiselle (semistructured) tiedolle
- Puolirakenteisuus: tietoelementit eivät ole yhtä tiukasti tyypitettyjä kuin esim. SQL:ssä,
 - muoto on vaihtelevampi kuin tiukka tauluesitys
- XML:ssä tiedot esitetään dokumenttina, joka jakautuu hierarkkisesti elementteihin

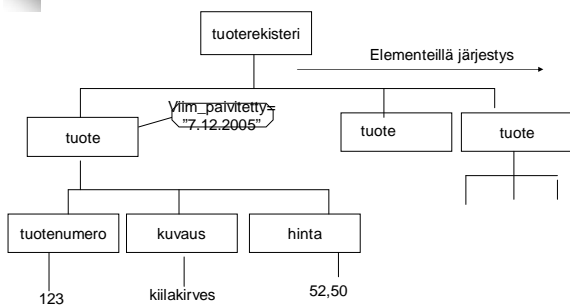
1

XML:n käyttötilanteita

- Tiedonsiirto ja järjestelmien välinen kommunikointi - järjestelmäraja-alue
- Yhtenäinen näkymä heterogeenisiin (erimuotoisiin) tietoihin
 - välittäjä rakenne
- Asetustiedostojen yleinen muoto
- Tiedon tallennustapa (grafiikka (SVG, DrawingML), tekstidokumentit (OOXML, ODF, ...),
- Originaali tai generoitu esitysmuoto
 - Esimerkiksi relaatiotietokannasta voidaan generoida XML-muotoista tietoa

2

Xml-dokumenttipuuna



3

XML-elementit

- Kullakin elementillä on nimi ja sisältö
- Sisältö voi olla
 - tyhjä
 - tekstidataa,
 - toisia elementtejä tai
 - sekä tekstidataa että elementtejä
- Elementtiin voi myös liittyä attribuutteja (määrittäjiä)
 - Attribuutilla on nimi ja arvo
- Data voidaan siten esittää joko elementin sisältönä tai elementin attribuutin avulla.
- Tietojen välisiä yhteyksiä esitetään järjestyksen ja sisältyvyyden avulla. Näiden lisäksi voidaan tietoja kytkeä toisiinsa elementtien sisällön avulla.

4

XML-tekstiesitys

- XML-tekstiesitys = miten dokumentti esitetään tekstinä
- Elementin aloittaa alkunimiö (start tag) ja päättää loppunimiö (end tag)
 - <nimi></nimi>
- Alkunimiön ja loppunimiön välissä sisältö, joka voi olla tekstidataa tai elementtejä
- Alkunimiöön voi sisältyä attribuuttilausekkeita
 - Attribuutti="arvo"

5

XML-tekstinä (esimerkki)

```
<tuoterekisteri>
  <tuote viimeksi_paivitetty="7.12.2005">
    <tuotenumero>123</tuotenumero>
    <kuvaus>kiilakirves</kuvaus>
    <hinta>52,50</hinta>
  </tuote>
  <tuote ...>...</tuote>
  <tuote ...>...</tuote>
</tuoterekisteri>
```

- Käytettävät nimet, attribuutit ja elementtien järjestys kiinnitetään XML-kielen määrittelyssä

6

XML-kielen määrittely

- n XML-dokumentin rakenne (=XML-kieli) määritellään joko DTD-tyyppimäärittelyn (document type definition) tai XML-kaavion (XML-schema) avulla.
- n Jälkimmäisellä tavalla pystytään määrittelemään attribuuttien arvot täsmällisemmin.
- n Tyyppimäärittely tai kaavio määrittelee XML-rakenteisen kielen

7

dtd-määrittelyä

```
<!ELEMENT tuoterekisteri (tuote*)>  
<!ELEMENT tuote (tuotenumero, kuvaus, hinta)>  
<!ELEMENT tuotenumero (#PCDATA)>  
<!ELEMENT kuvaus (#PCDATA)>  
<!ELEMENT hinta (#PCDATA)>  
<!ATTLIST tuote  
viimeksi_päivitetty NMTOKEN #REQUIRED>
```

Lue: Tuoterekisteri-elementti muodostuu joukosta tuote-elementtejä.
Tuote-elementti muodostuu tuotenumero-, kuvaus- ja hinta-elementeistä.
Näistä kukin sisältää tekstidataa. Tuoteella on pakollinen attribuutti.
Huom: jokaisella yrityksellä voi olla oma XML-kielensä tuoterekisteriä varten

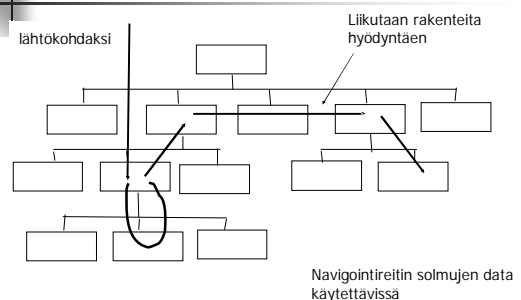
8

XML:n käsittely

- n XML-rakenteen käsittely eroaa relaatiomallin tavasta hakea tietoa relaatioalgebran operaatioiden avulla.
- n Tiedonhaun periaatteena navigointi dokumenttipuussa
 - n Valitaan dokumenttipuun solmu, josta lähdetään liikkeelle
 - n Valinta voi perustua elementin nimeen, sen asemaan suhteessa muihin elementteihin tai sen sisältöön
 - n Valitusta lähtöelementistä voidaan jatkaa etenemistä dokumenttipuun tarjoamia yhteyksiä hyväksikäyttäen
 - n Solmun lapsisolmuun (alas), velisolmuun (samalla tasolla), isäsolmuun (ylös), mahdollisesti tasoja tai solmuja ohittaen

9

Navigointi



10

Navigointi

- n Navigointiin voi liittyä silmukoita ja ehtoja
- n Tulos muodostuu 'proseduraalisesti' (vaiheittain)
- n Hierarkiat ja verkot sql:ää kätevämmiin hoidettavissa.
- n Isot tietomäärät, rinnakkaisuus ja muuttuvat tiedot ongelmallisia

11

XML:n käsittely

- n Määriteltyjä kyselykieliä: XPath, XQuery,..
- n Määritelty muunnosmekanismi datan esittämiseksi: XSL
- n Muutamissa relaatiotietokannoissa mahdollisuus XML-tiedon tallennukseen ja XML-näkymät eli miten relaatiot näkyvät XML:nä

12