

TIETOKANNAT

JOHDANTO

© JOUNI HUOTARI & ARI HOVI 2000 - 2017



jamk.fi

TAUSTAA

- Yritykselle tiedot ovat tärkeä resurssi
 - päätöksenteon tukena (JIT)
 - varastointi ja käyttö vaativat investointeja
 - vrt. energia (lähde, tuotanto, hallinta, oikeudet...)
- Tiedonhallinta: kaikkea tiedon määrittelyyn, suunnitteluun, toteuttamiseen, ylläpitoon ja käyttämiseen liittyvää toimintaa



SISÄLTÖ

- Tekniset talletustavat

 - Tietokanta

 - Tietokannan hallintajärjestelmä

 - Tietokantajärjestelmä

- Tietokannan käyttö

- Tietokantajärjestelmän edut



Tehtävä

- Määrittele tietokanta
- Mitä tietokantoja olet käyttänyt?
- Minkälainen on mielestäsi hyvä tietokantaratkaisu (mainitse esimerkkejä)?
- Millä välineillä tietokanta voidaan toteuttaa?
- Ketä tarvitaan, että tietokanta saadaan käyttäjiä palvelevaksi ratkaisuksi?



Tietokanta on (opiskelijoiden määritelmiä)...

- ”Iso kasa tietystä järjestyksessä olevaa tietoa”
- ”Tiedon (tekstiä, kuvia ym.) tallennuspaikka”,
”tietovarasto”
- ”Kokoelma yhteen liittyviä tietoja”
- ”Tiedot ovat helposti saatavilla”
- ”Ei tarpeetonta, ylimääräistä tietoa”
- ”Määrämuotoinen rakenne (taulukkomainen)”,
”koostuu tauluista” ...



Tietokanta

- Data Base (database), lyh. DB
- "Tietokoneistettu kortistojärjestelmä"
- "Kokoelma toisiinsa liittyviä tietoja, jotka on talletettu ilman haitallista ylimäärää ovat usean sovelluksen käytettävissä hallitaan yhtenäisen ohjelmiston avulla"
"... ja jonka rakenteen kuvaus on erotettu ohjelmista" (Nachmens)
- "Reaalimaailman tietomallin toteutus"



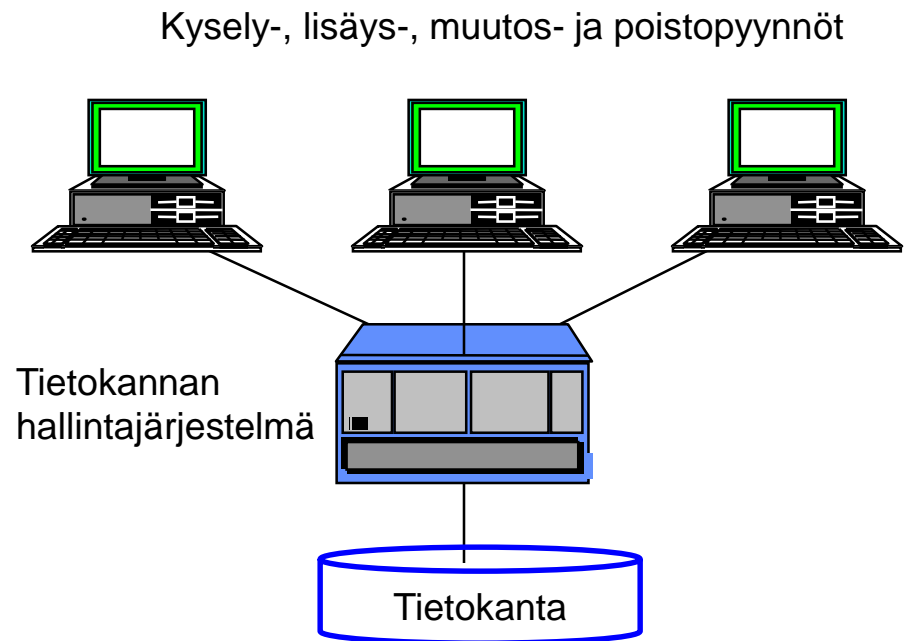
Tehtävä

- Tutustu JaNet-kirjastotietokantaan:
<http://janet.amkit.fi/>
= eräs esimerkki tietokannasta
=> hae tietoa esim. aiheesta "Tietokannat"
- Lisätehtävä: tutustu myös muihin kirjaston sivulle koottuihin tietokantoihin



Tietokannan hallintajärjestelmä

- DataBase Management System eli DBMS (suom. TKHJ)
- "Ohjelmisto, joka hoitaa kaikki tietokantaoperaatiot" (Date 1975)

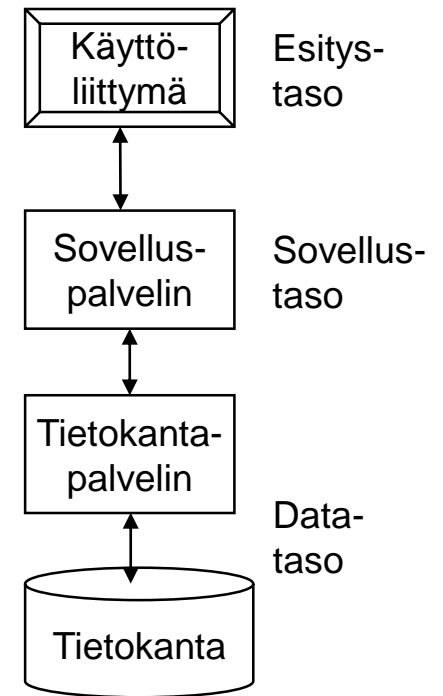
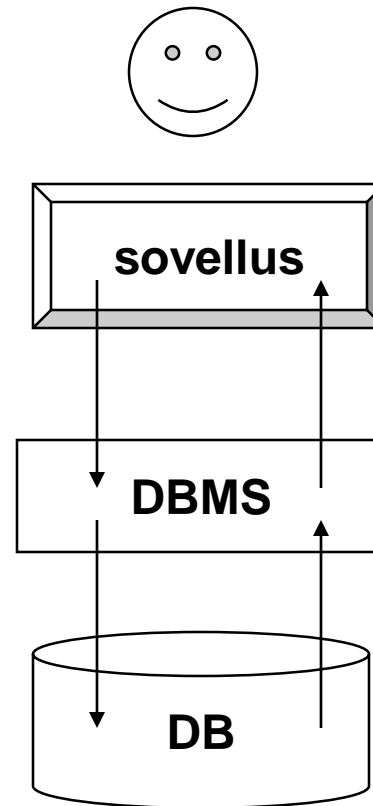


Huom.: tietokantaohjelmisto \neq TKHJ. Tietokantaohjelmistolla luodaan tietokanta ja siihen liittyvät määrytykset (kuten käyttöoikeudet).



Tietokantajärjestelmä

- Tietokantaa käyttävät ohjelmat +
- Tietokannan hallintajärjestelmä +
- Tietokanta



Sovellusluston komponentit



Pasi Mäkinen 2008

<http://seminarit.codezone.fi/video/20070821>

/1/

Microsoft Access

- Sisältää nk. tietokantamoottorin, joka vastaa TKHJ:ä
- Se on sovelluskehitin, koska sillä voidaan luoda
 - tietokanta (taulut ja niiden yhteydet) ***
 - kyselyt
 - käyttöliittymä (nk. lomakkeet)
 - raportit
 - sovelluslogiikka (sisältää mm. VBA-ohjelmointikielen)
 - tietokannan käyttöoikeudet
- Se on siis sekä sovelluskehitin että se voi toimia myös tietokantajärjestelmänä

*** Tietokannan taulut voidaan liittää ODBC:n avulla toisesta TKHJ:stä Microsoft Accessiin



Tehtävä

- Käynnistä koneellasi oleva tietokantaohjelmisto, esim. Microsoft Access
- Avaa esimerkkietokanta, esim. Northwind
- Tutustu tauluihin: avaa ainakin yksi taulu, selaa taulun sisältöä ja tarkastele myös sen rakennetta
- Tutustu myös muihin esimerkkietokannan ”komponentteihin” (kyselyihin jne.)



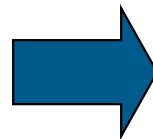
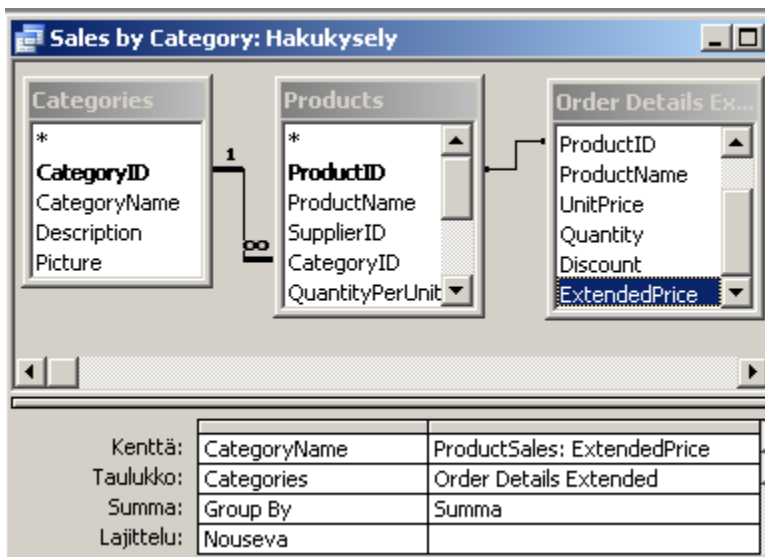
Tietokannan käyttö

- Kyselyt eli haut (Queries)
- Tietojen ylläpito:
 - lisäys (insert, append)
 - muuttaminen, päivittäminen (edit, update)
 - poistaminen (delete)
- Tietokannan koko: sadoista tietueista / kymmenistä kilotavuista miljooniin tietueisiin tai tuhansiin Gigatavuihin (Teratavuihin)
- Tietokanta ja taulukkolaskentaohjelma ovat yhdessä mainio pari

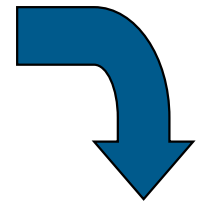


Esimerkki

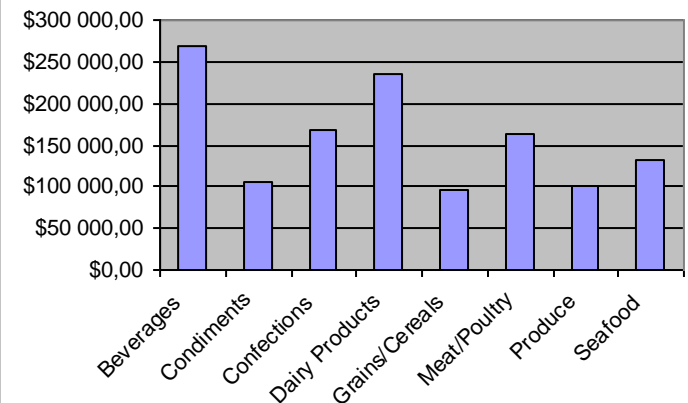
- Haku eli kysely Northwind-tietokannasta



Category Name	Product Sales
Beverages	\$267 868,16
Condiments	\$106 047,05
Confections	\$167 357,16
Dairy Products	\$234 507,25
Grains/Cereals	\$95 744,59
Meat/Poultry	\$163 022,37
Produce	\$99 984,57
Seafood	\$131 261,73



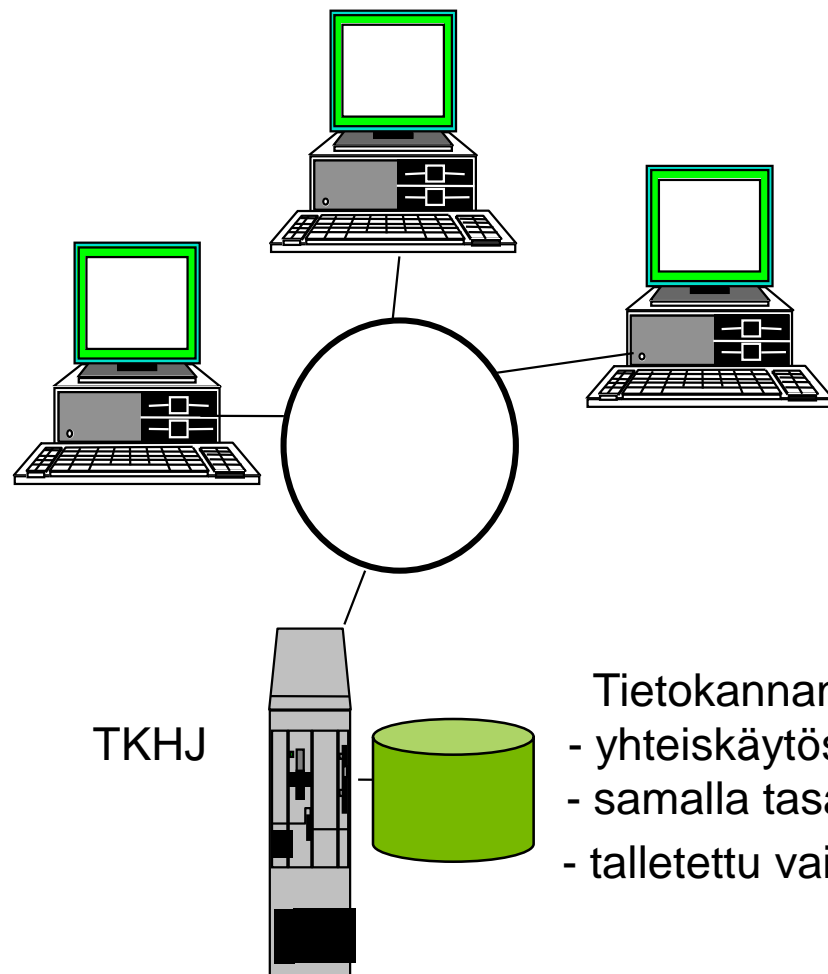
Product Sales



```
SELECT Categories.CategoryName, Sum([Order Details Extended].ExtendedPrice) AS ProductSales
FROM Categories INNER JOIN (Products INNER JOIN [Order Details Extended] ON Products.ProductID = [Order Details Extended].ProductID) ON Categories.CategoryID = Products.CategoryID
GROUP BY Categories.CategoryName
ORDER BY Categories.CategoryName;
```

Tietokantojen ominaisuuksia

- Hakujen nopeus
- Tiedot saadaan järjestyksessä
- Tietojen ajan tasalla pitäminen
- Tieto tallennetaan vain kerran
- Yhteiskäyttö
- Omien näkymien (view) luonti
- Tulokset näytölle, paperille tai tiedostoon
- Tietojen suojaus

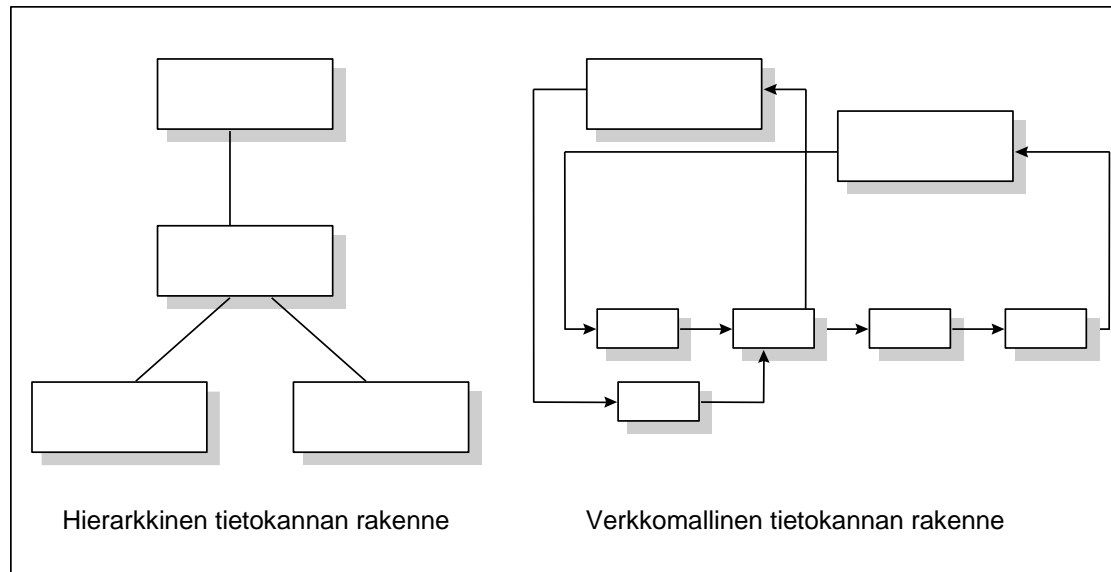


Tietokannan tiedot

- yhteiskäytössä
- samalla tasalla
- talletettu vain kerran

Tietokantatyypit

- Historia: kortistoista tiedostoihin ja sen jälkeen hierarkkisiin tai verkkomallisiin tietokantoihin



- Relaatiotietokannat yleisimpiä
- NoSQL-kannat yleistymässä

Tauko ?



RELAATIOTIETOKANNAT

- Relaatiomalli: E.F. Codd v. 1970
- Perustuu matematiikkaan, joukko-oppiin (relaatioalgebraan: yhdiste, erotus, leikkaus, ristitulo, valinta, projektio, liitos, jako) ja predikaattilogiikkaan
- Käsittely SQL:llä (Structured Query Language); SQL voidaan upottaa sovellukseen
- Esimerkki esitetty [aikaisemmin](#)



Esim. kahden relaation leikkaus

- Leikkaus sisältää ne monikot, jotka kuuluvat molempiin lähtörelaatioihin

Relaatio X:

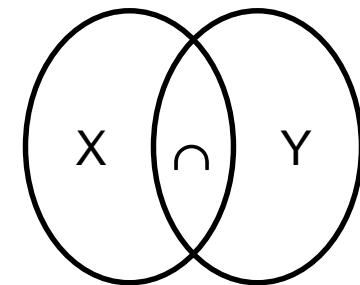
A	B	C
01	a	x
02	b	x
03	c	y

Relaatio Y:

A	B	C
02	b	x
04	d	z

Leikkaus $X \cap Y$

A	B	C
02	b	x



Taulu

Sarake

Sarake-
nimet

YRITYS

<u>yrtun</u>	<u>nimi</u>	<u>postinro</u>	<u>postitmpka</u>	<u>puhelin</u>
2245	Ali Hewi Oy	00200	HELSINKI	09-348298
3124	Eilan Kahvila	02150	ESPOO	09-876938
2143	Kokotieto Tmi	40640	JYVÄSKYLÄ	014-453623
3211	Matsmickelson	90650	OULU	08-6371180
2214	Tapionmäki Ltd	99101	LUUPLANKKA	040-342569

Rivi

"isätaulu"

Perusavain



Viiteavain



HENKILO

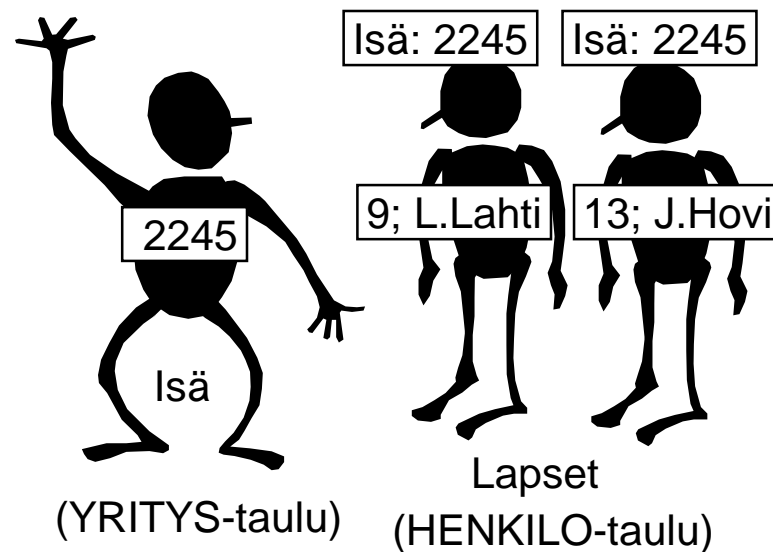
<u>hnro</u>	<u>etunimi</u>	<u>sukunimi</u>	<u>syntvuosi</u>	<u>pvm</u>	<u>yrtun</u>
9	Laura	Lahti	1961	15.3.1991	2245
13	Julia	Hovi	1984	10.4.2003	2245
2	Lydia	Huotari	1988	15.2.2003	2143
1	Jukka	Jokinen	1959	13.3.1992	3211
5	Kerkko	Jokinen	1959	13.3.1992	3211
7	Lea	Ranta	1971	28.9.1994	3211
15	Lyly	Wrede	1956	1.12.2002	3124

"lapsitaulu"

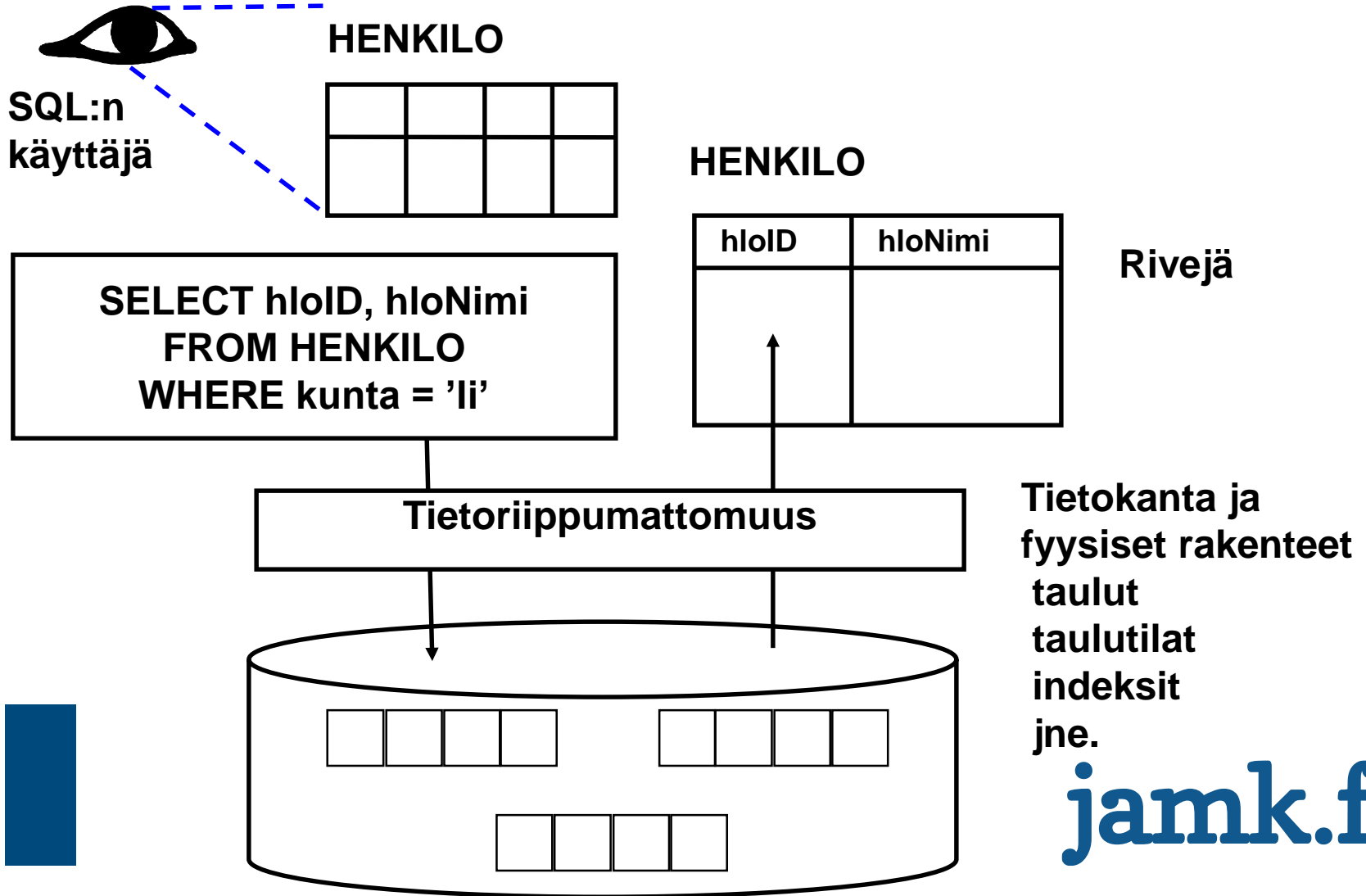


Yhteys taulujen välillä

- YRITYS- eli FIRMA-taulun isällä voi olla monta lasta HENKILO-taulussa (Asiakastunnus on perusavain)
- Lapsesta näkee, kuka on isä (Asiakastunnus on viiteavain)

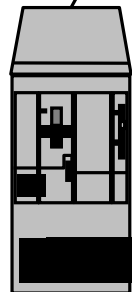
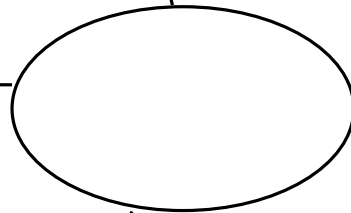
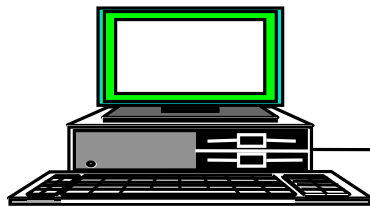
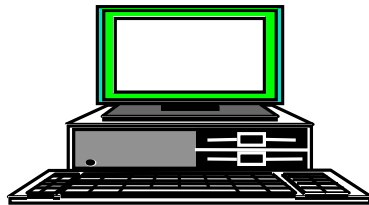


Tietoriippumattomuus



Tietokantapalvelimet

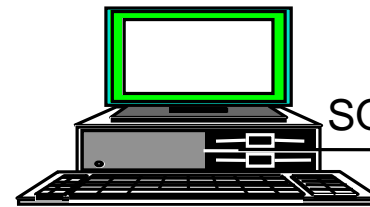
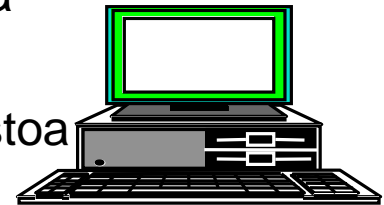
Sovellus työssämassa, samoin tietokantaohjelmisto.



Taulut tiedostossa.
Ei tietokantaohjelmistoa.

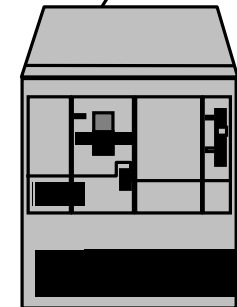
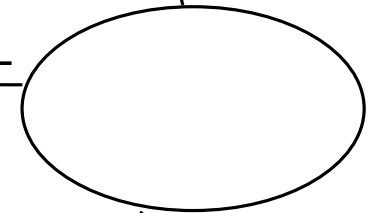
Tiedostopalvelin

Sovellus työssämassa lähettää SQL-kielisiä käskyjä SQL-palvelimelle; ei tietokantaohjelmistoa



SQL

SQL



SQL-kanta;
SQL-tietokantaohjelmisto,
sisältää taulut

Tietokantapalvelin

TIETOKANTOJEN KÄYTTÖTAPOJA

- Operatiivinen tietojenkäsittely

tietokantaa päivitetään runsaasti (transaktioita eli tapahtumia useita / sekunti)

Esim. pankkitilit ja taloushallinto

- Tietovarastot (Data Warehouse, DW)

tietoja kopioidaan (eräajolla) operatiivisista järjestelmistä tietovarastoon säännöllisin väliajoin

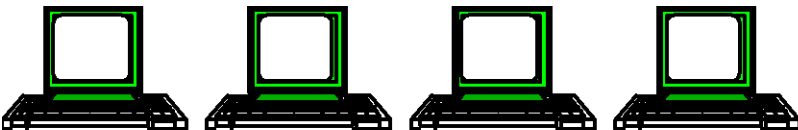
kyselyjä, raportointia ja tietojen analysointia varten

business intelligence (BI)

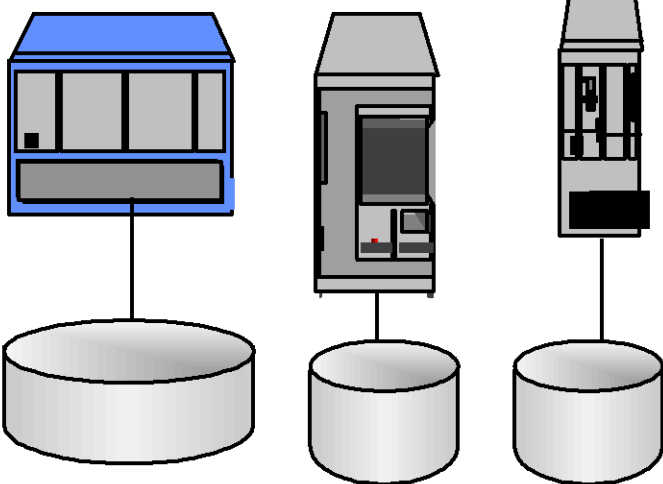
Big Data

- Internet-/mobiili- ja muut tietokannat

Tietokanta pilvipalveluna eli Database-as-a-Service, DaaS tai DBaaS



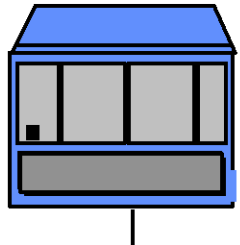
Tapahtumankäsittely,
lisäyksiä, päivitystä, poistoja, kyselyjä



**Perusjärjestelmien
operatiiviset tietokannat**



Raportit, kyselyt, analyysit,
tietoja vain luetaan



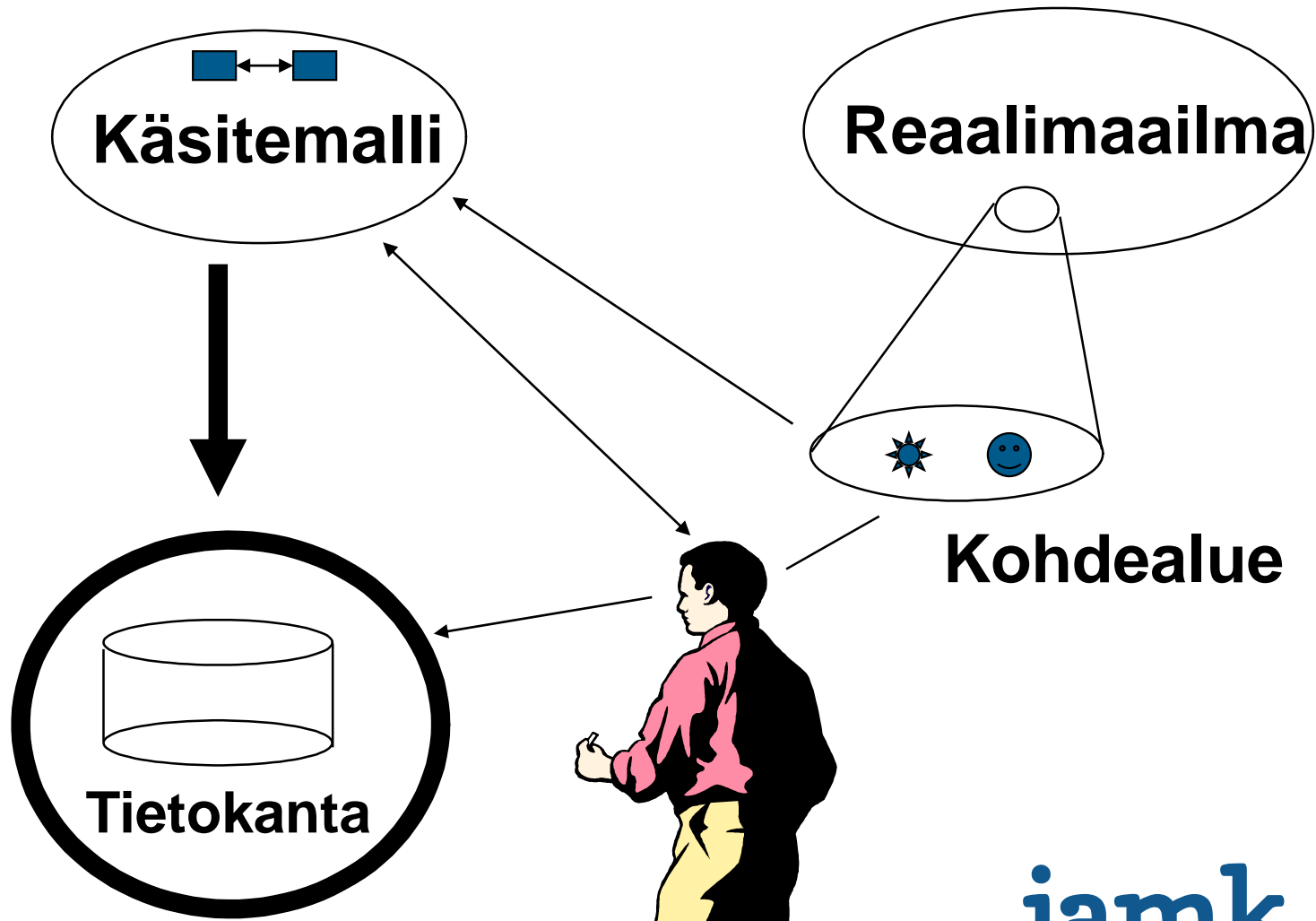
Tietovarasto

TIETOKANTOJEN SUUNNITTELU

- Käsiteanalyysi
- Käsittemalli: tietokannan piirustukset
- Haetaan käsitteitä (kohde, objekti, eng. entity), esim. Asiakas, Tilaus, Tilausrivi, Osasto, Tuote, Tuoteryhmä, Henkilö, Tili, Tilivienti
- Katso tarkemmin kalvot Käsiteanalyysi (suunnitteluputki) ja/tai kirjasta luvut 2 - 4



Tietokannan toteutus



Tietokantojen käyttäjäryhmät

- Tietokantojen määrittely, suunnittelu ja toteutus voi jakautua seuraaviin ryhmiin:
 - järjestelmäsuunnittelija tekee vaatimusmäärittelyt
 - tietokantasuunnittelija tekee kuvauksen tietokannan rakenteesta ja tiedoista (raja-arvot, tietotyypit jne.)
 - järjestelmänvalvoja (administrator) määrittää tietokannan kaikki oikeudet
 - sovellusohjelmoijat toteuttavat toimintalogiikan (mm. SQL-lauseiden upottaminen koodiin)

• Käyttäjät hakevat ja päivittävät (ylläpitävät) tietoja



Turvallisuusnäkökohtia

- Varmistukset

vaihtuvat: päivittäin, viikoittain
pysyvät: kuukausittain tai vuosittain



- Eheys

tiedot oikein (esim. tarkistussäännöt)
viite-eheydestä huolehtiminen

- Tietosuoja

lain huomioiminen (esim. EU:n kyberturvalaki)
käyttöoikeuksien valvonta (salasanat, tietokantalokien seuranta)

- Laittomien SQL-kyselyjen ja –injektioiden esto



Tehtävä

- Tutustu seuraavilla sivulla oleviin linkkeihin
- Minkä relaatiotietokantatuotteen valitsisit oman harjoitustyösi tekemiseen ja miksi?



Linkkejä 1

- Sanasto:

<http://www.cs.helsinki.fi/~laine/relaationsanasto/>

- SQL:

<http://www.firstsql.com/tutor.htm>

<http://www.managedtime.com/freesqlbook.php3>

<http://www.contrib.andrew.cmu.edu/~shadow/sql.html>

www.sqlcourse.com ja www.sqlcourse2.com

<http://www.guru99.com/sql.html>



Linkkejä 2: tuotteita

- http://en.wikipedia.org/wiki/Comparison_of_relational_database_management_systems
- Oracle: <http://www.oracle.com/>, <http://otn.oracle.com>
- DB2: <http://www-4.ibm.com/software/data/db2/>
- SQL-Server: <http://www.microsoft.com/sql/>
- MySQL: <http://dev.mysql.com/>
- PostgreSQL: <http://www.postgresql.org/>
- Ingres: <http://opensource.ca.com/projects/ingres>
- Solid Server:
<https://teambblue.unicomsi.com/products/soliddb/>

