

# Nem konvencionális adatbázisok

Dr. Balázs Péter

Szegedi Tudományegyetem

Képfeldolgozás és Számítógépes Grafika Tanszék

# Tematika

- Relációs adatmodell, EK-diagram, normalizálás, SQL
- Objektum-relációs modell, SQL3
- Térbeli adatbázisok
- Temporális adatbázisok
- XML alapú adatbázis kezelés, XPath, Xquery
- NoSQL adatbáziskezelés
- Képi- illetve multimédia adatbázisok
- Képi tartalom alapú keresés
- Konkrét speciális adatbázisok (műszaki, orvosi, stb.)

# Segédanyag, irodalom

- Előadás, gyakorlat anyaga: [www.inf.u-szeged.hu/~pbalazs](http://www.inf.u-szeged.hu/~pbalazs)
- Katona Endre: Térképi adatbázisok, Typotex, 2011, online elérhető
- Garcia-Molina H., Ulmann J.D., Widom J.: Adatbázisrendszerek megvalósítása, Panem, 2001
- Shashank T.: Professional NoSQL, John Wiley & Sons, 2011
- Richard T. Snoggrass: Developing Time-Oriented Database Applications in SQL, Morgan Kaufmann, 1997
- Oracle referenciák
- ...

# Mitől nem konvencionális?

- *Hatalmas méretű:* a tárolt adatmennyiséget az elérhető hagyományos adatbázisrendszerek és adatfeldolgozó eljárások képtelenek kezelni
- *Temporális:* Nem csak az adatok aktuális állapota, hanem azok időbeni változása is nyomon követhető/követendő
- *Térbeli, multimédiás:* speciális objektumok (térkép, úthálózat, videó, hanganyag) tárolása, keresés
- *Félig strukturált:* jelölések segítségével az adatok szemantikus elemei elkülöníthetők, de táblák/relációk nincsenek
- *Strukturálatlan:* az adatok még jelölésekkel sincsenek ellátva, speciális eljárásokkal kell az értelmezhető információt kinyerni
- + a fentiek kombinációja

# Multimédiás adatbázis

- Kép, audio, video tárolása részben vagy egészben az adatbázisban
- Adatok + (metaadatok)
- Félig strukturált vagy strukturálatlan
- Minimális követelmények:
  - tárolási és visszakeresési lehetőségek
  - ismertebb formátumok támogatása
  - könnyű elérhetőség a WEB-ről
- Visszakeresés:
  - a metaadatokon egyszerű
  - a strukturálatlan adatokon komplex → tartalom alapú keresés
- Oracle interMedia, Oracle Visual Information Retrieval

# Temporális adatbázisok

- Fontos az adatbázis változásának nyomon követése is:
  - banki tranzakciók
  - betegek kórtörténete
  - kölcsönzések során a kölcsönzött tárgy állapotának változása
- Megvalósítás:
  - RDBMS: beépített dátum adattípusokkal
  - ODBMS, ORDBMS: absztrakt adattípusokkal
  - új megoldások
- Terradata Database 13.10, Oracle, IBM DB2, PostgreSQL, TimeDB

# NoSQL adatbázisok I.

- a hagyományos relációs modellel szakító irányzatok
- Kulcs/érték tárolók (key/value stores):
  - string kulcsokkal és string értékekkel dolgoznak (hash-tábla)
  - keresés a kulcs alapján
  - Amazon Dynamo (zárt) → Riak, Apache Cassandra (nyílt forráskódúak)
- Dokumentum tárolók (document stores):
  - hasonló az előzőhöz, de a szerkezetileg hasonló dokumentumok („értékek”) kollektciókat alkotnak
  - MongoDB, CouchDB (nyílt forráskódúak)

# NoSQL adatbázisok II.

- Oszloporientált tárolók (column-oriented stores):
  - középút a relációs adatbázisok és az objektum tárolók között
  - táblába szervezett adatok (sorok halmaza), de oszlopok helyett oszlopcsaládokat definiálunk
  - Google Bigtable (zárt) → Hypertable, Apache HBase (nyílt forráskódúak)
- Gráf adatbázisok (graph databases)
  - csúcsok, élek, mindkettőnek lehetnek tulajdonságai
  - szociális hálózatok
  - Neo4j (JAVA nyelven implementált, nyílt forráskódú), FlockDB, HyperGraphDB



# XML alapú adatbázis kezelés

- Ismeretlen adatstruktúrák fogadása (teljesen strukturált adatokat csak a séma ismeretében tudjuk beolvasni)
- Különféle forrásból eredő adatok integrálása
- Rugalmasabb adatstruktúrák alkalmazása
- Címkék fejezik ki az adatok struktúráját és típusait
- Az adatokat képes fa- illetve gráf struktúrában leírni
- XPath, XQuery lekérdezőnyelvek