
A SINTAGMA szemantikus integrációs rendszer – áttekintés

Lukácsy Gergely

lukacsy@cs.bme.hu

Budapesti Műszaki és Gazdaságtudományi Egyetem
Számítástudományi és Információelméleti Tanszék

-
- ① Motiváció, információintegráció
 - ② A rendszer architektúrája
 - ③ A modelltárház, az integráció folyamata
 - ④ Szemantikus komponensek

AZ ALAPTECHNOLÓGIA

Bevezetés

- Tanszékünk 2005 januárjától részt vesz az NKFP által támogatott SINTAGMA projektben
 - fejlesztői oldalról rajtunk kívül a KFKI és a SZTAKI vesz részt
 - alkalmazói oldalról az MTI, az OSZK és egy közép vállalkozás képviselteti magát
- A SINTAGMA rendszer már létező technológiákra, a SILK és a LOBO projekt eredményeire épít

Szemantikus integráció

- A projekt fő célja egy olyan szemantikus integrációs technológia kialakítása, amely lehetővé teszi a legkülönbözőbb típusú információk egységes és együttes kezelését
- Szemantikus integráció alatt olyan megoldást értünk, amely tudásreprezentációt és következtetést használ a különféle integrációs feladatok megoldása során

BEVEZETÉS

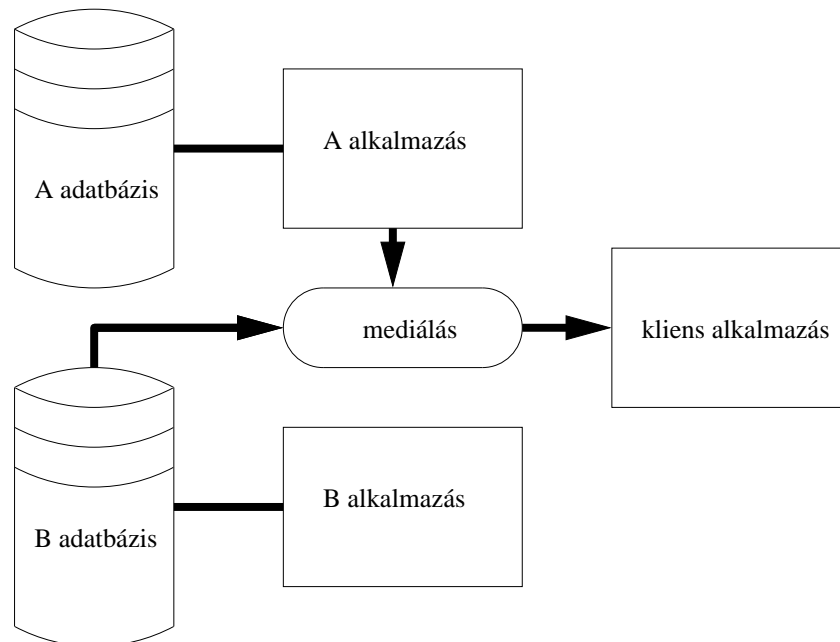
Motiváció

- egy szervezetnél az aktuális igényeknek megfelelően különböző adatforrások és alkalmazások vannak
- előbb-utóbb előáll egy olyan helyzet, ahol bizonyos adatokat több rendszer is tárol
 - felmerül a kérdés, hogy hol is keressünk
 - megoldás lehet a redundáns adatok konzisztenciájának biztosítása, amihez egy köztes réteget kell beiktatni
- az is elképzelhető, hogy olyan tudásra van szükségünk, amely csak szétosztva van jelen
- az integráció során ezekre a problémákra keresünk olyan megoldást, amely újrahasznosítja a már meglévő adatokat és alkalmazásokat

INTEGRÁCIÓS MEGKÖZELÍTÉSEK

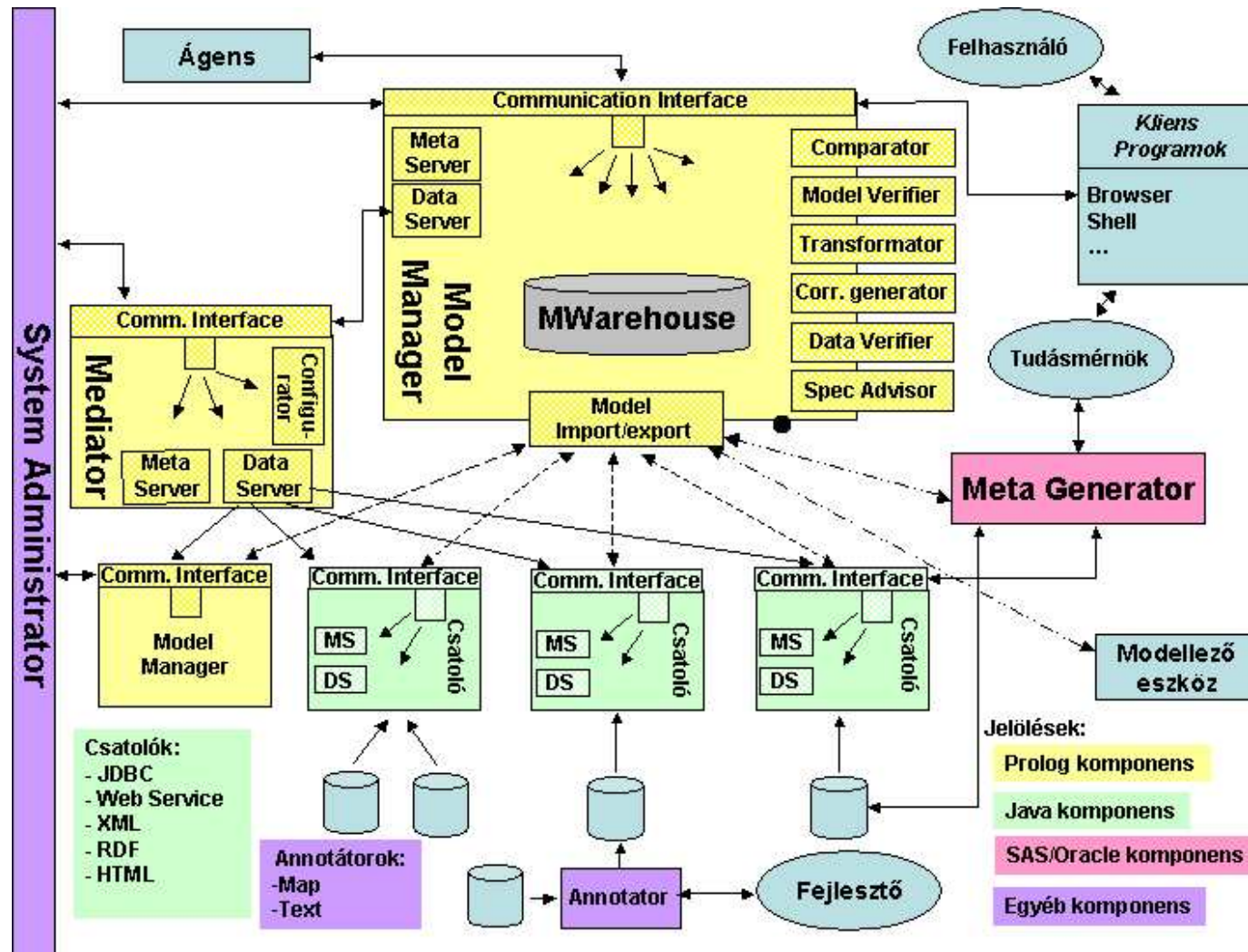
Működési minták

- Ad-hoc integráció
- Alkalmazásintegráció
- Adatintegráció
- **Információintegráció**



- egyirányú kommunikáció
- a mediátor adatforráshoz és alkalmazáshoz is csatlakozhat
- nincsen adatduplikátum
- minden kommunikáció **on-line**
- a kliens trendanalízis helyett friss információkat kap
- a közös séma benépesítése és lekérdezése egy lépésben történik meg
- a mediátornak kérdésről kérdésre más lekérdezési tervet kell készítenie
- a kliens oldaláról a mediátor egy virtuális adatbázist mutat fel

A SINTAGMA RENDSZER FELÉPÍTÉSE



A MODELLTÁRHÁZ

Modelltárház tartalma - hibrid modellezés

- Alapvetően UML alapú objektum-orientált modelleket tartalmaz
 - struktúrális és egyéb jellemzőket ír le velük
 - OCL korlátokat használ például az invariánsok megadására
 - néhány helyen eltér az UML specifikációtól
- A felhasználók tudását leíró logikákkal (OWL-DL) is megadhatjuk
- Kérdések és lekérések is itt tárolódnak

A modellek osztályozása

- A modelleket két dimenzió szerint osztályozzuk
- Absztrakciós szint szerint megkülönböztetünk alkalmazási- és fogalmi szintű modelleket
- A lefedettség szerint megkülönböztetünk lokális- (kapcsolati-) és egyesített modelleket

A MODELLTÁRHÁZ NYELVE

A SILan nyelv

- Az modelltárházban lévő modelleket szöveges alakban írja le
- Az alábbiakban egy autó- és kerékgyár modelljét mutatjuk be

```
model Factory {  
  class Product {  
    attribute String serial;  
    primary key serial;  
  };  
  
  class Car: Product {  
    attribute String make;  
    attribute Integer price;  
    constraint self.wheel.size > 2;  
  };  
  
  class Wheel: Product {  
    attribute Integer size;  
  };  
  
  association 'Car-Wheel' {  
    connection Car;  
    connection Wheel composite;  
  };  
};
```

LEKÉRDEZÉS MAGAS SZINTEN

Absztrakciók

- Ahhoz, hogy magas szinten tehessünk fel kérdéseket, első lépésként kapcsolatot kell teremteni a magasabb és alacsonyabb szintű modellek között
- Egy ilyen kapcsolatot hívunk absztrakciónak
- Az alábbiakban azt láthatjuk, hogy a Peugeot járművekkel foglalkozó forrást hogyan kötjük össze az előzőekben megadott magas szintű modellel

```
abstraction (s: Peugeot::Vehicle -> c: Factory::Car) {  
  constraint s.type = "Car" implies s.serial_number = c.serial  
    and "Peugeot" = c.make  
    and s.price*1000 = c.price;  
};
```


LEKÉRDEZÉS MAGAS SZINTEN

Lekérdezések megadása

- Lekérdezéseket meg lehet fogalmazni bármilyen szintű modellen
 - lekérdezhetjük akár a kapcsolati modelleket is
 - feltehetünk kérdéseket a legmagasabb szintű modellen is
- Az alábbi lekérdezés az olyan összetartozó kocsik és kerekek azonosítóját adja vissza, ahol a kocsi olcsó és a kerék kellően nagy (a mértékegységek már egységesek)

```
query CheapBigWheelCar
select w.serial, w.car.serial
from w: Factory::Wheel
where w.size > 4 and w.car.price < 5000;
```

- Amennyiben a megfelelő absztrakciók léteznek és az adatforrások elérhetők, a lekérdezés végrehajtható

SZEMANTIKUS KOMPONENSEK

Csatolók

- XML csatoló: modell DTD-ből, illetve XML sémából
- WS/SOAP csatoló: modell WSDL-ből
- HTML csatoló: modell egyedi építés alapján
- RDF/RDFS/OWL csatoló: modell a séma alapján
- természetes nyelvű csatoló: modell egyedi építés alapján

Dugaszok

- modell ellenőrző: leíró logikai, korlátalapú
- modell összehasonlító: strukturális és lexikai
- modell transzformátor

Lekérdezés - mediátor