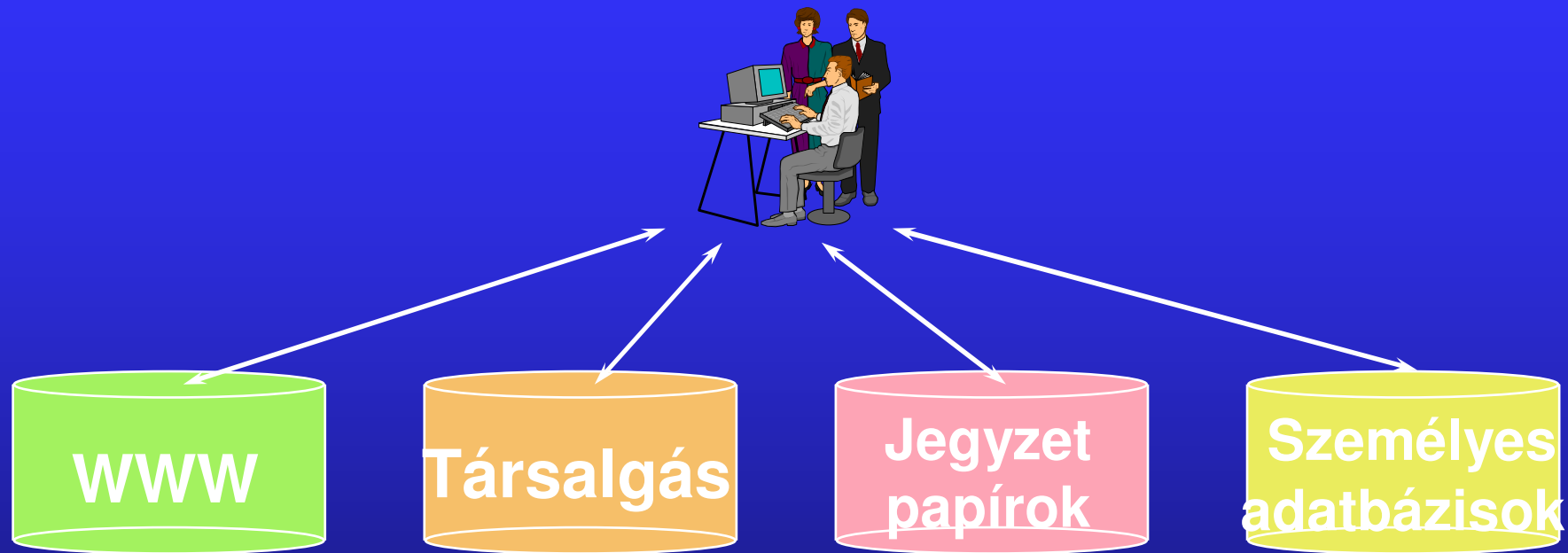


TSIMMIS

egy lekérdezés centrikus megközelítés

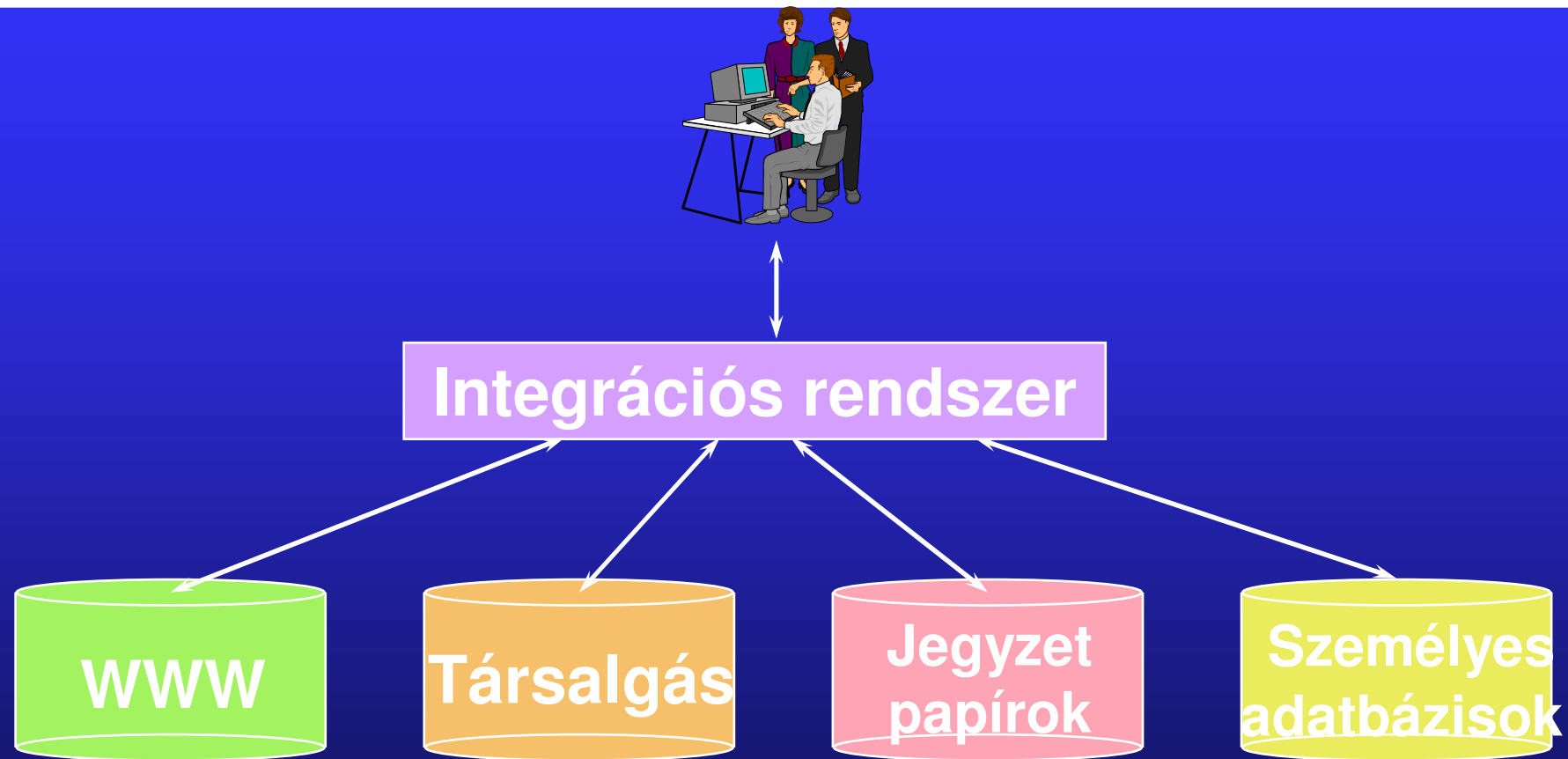
- TSIMMIS célok, technikák, megoldások
- TSIMMIS korlátai
- További lehetőségek

Információk heterogén információs forrásokban érhetőek el



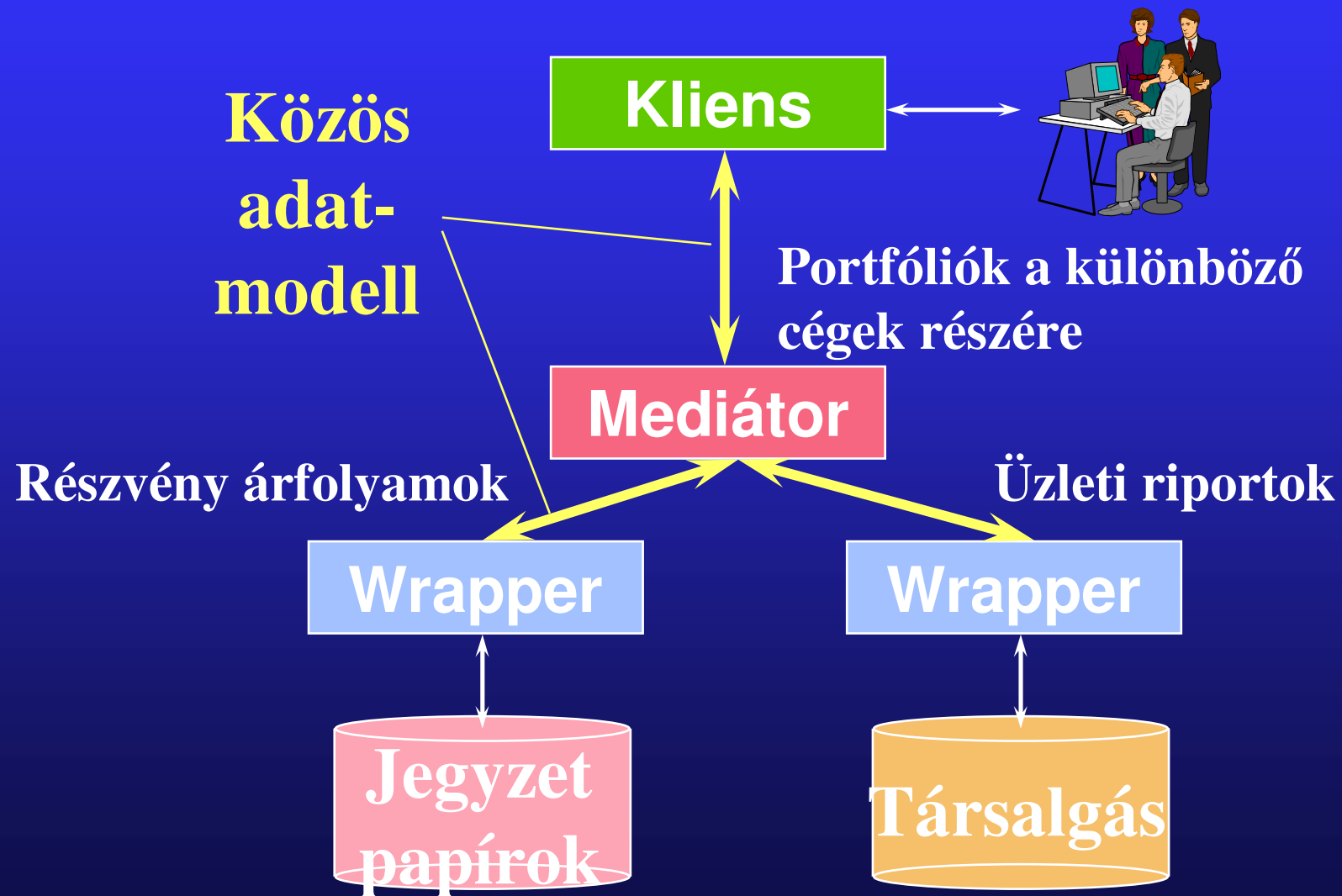
- **Eltérő interfészek**
- **Eltérő adatrepresentációk**
- **Redundáns és ellentmondásos információk**

Cél: Egységes nézetet és elérést biztosítani az adatoknak

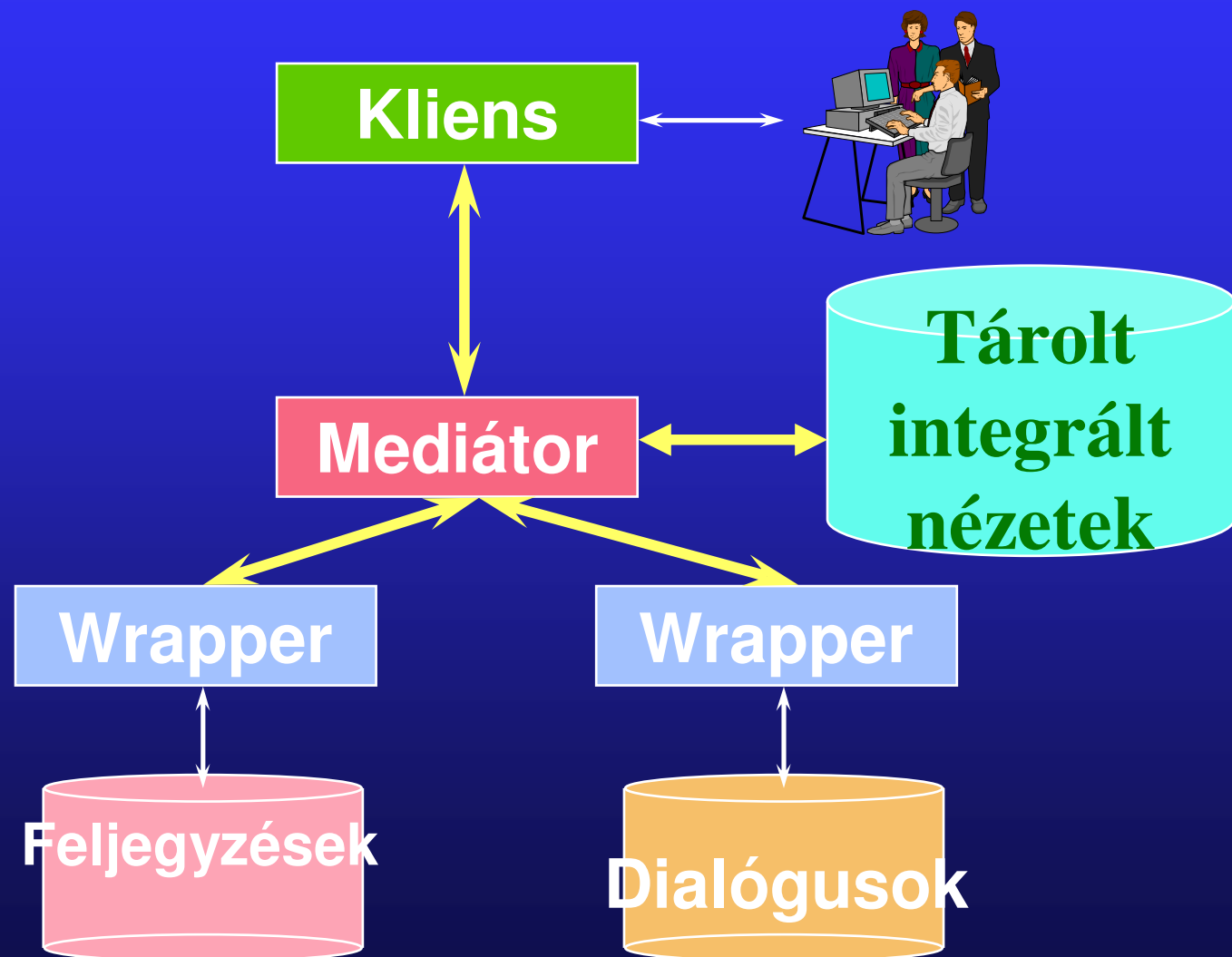


- Összegyűjt és összekombinál információkat
- Integrált felhasználói nézetet és egységes felhasználói felületet biztosít

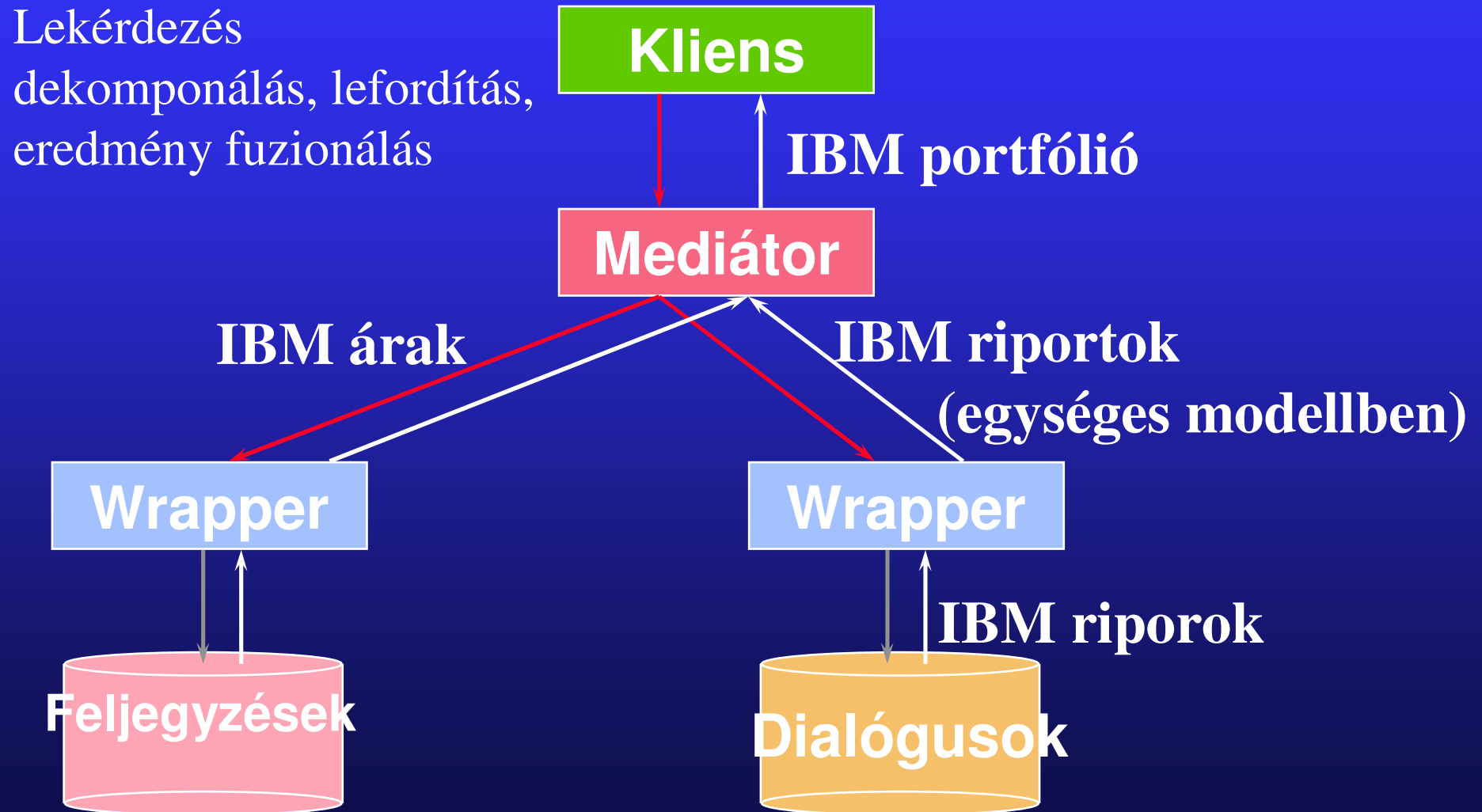
A mediátor megközelítés



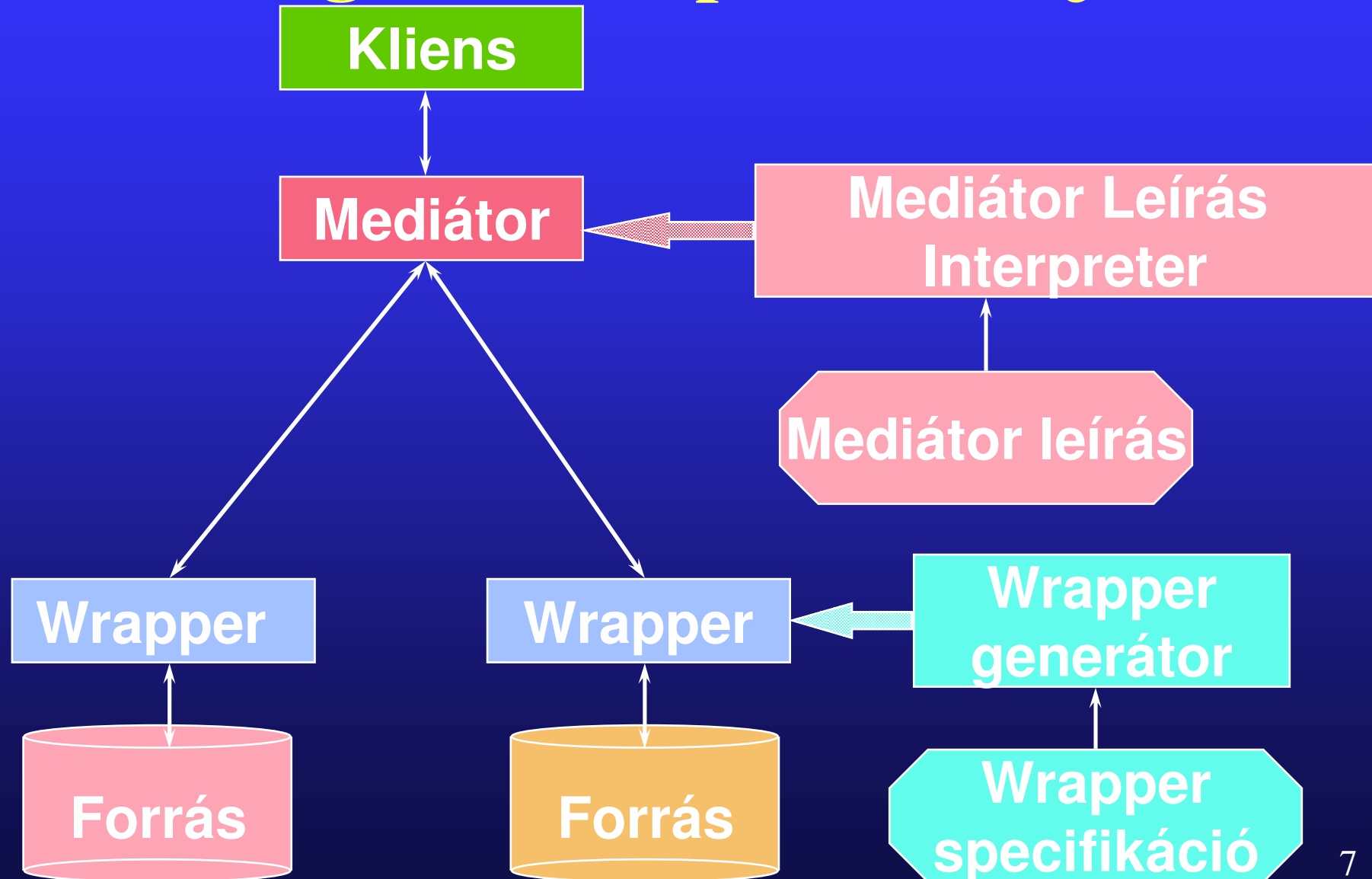
Adattárház megközelítés



Virtuális integráció megközelítés



Wrapper-ek és mediátorok magasszintű specifikációja



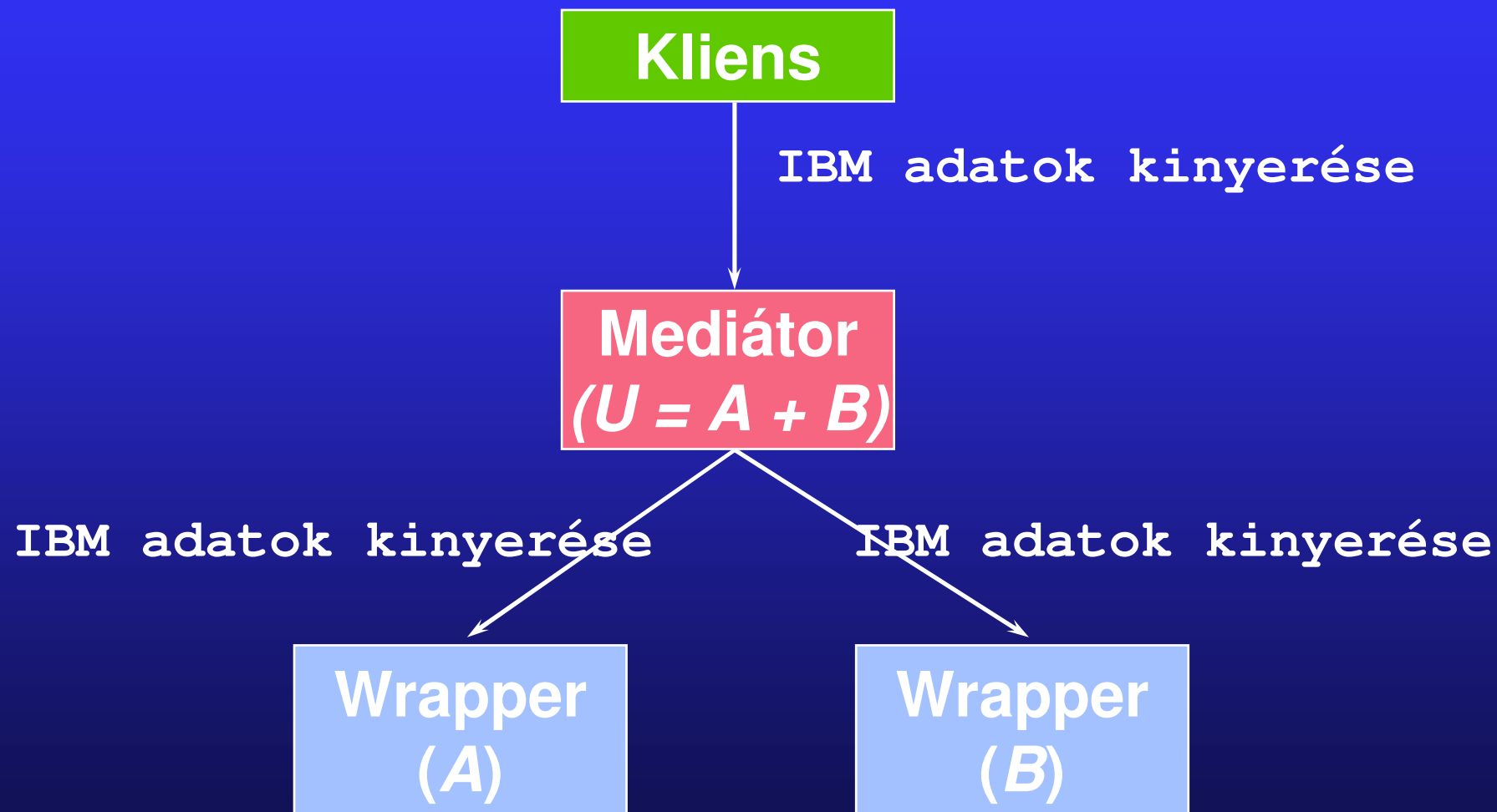
Kihívás: Jól strukturált forma nélküli források elemzése

- Félig-strukturált
 - irreguláris
 - Mélyen beágyazott
 - kereszthivatkozások
- Hiányos séma információ
 - autonóm
 - dinamikus

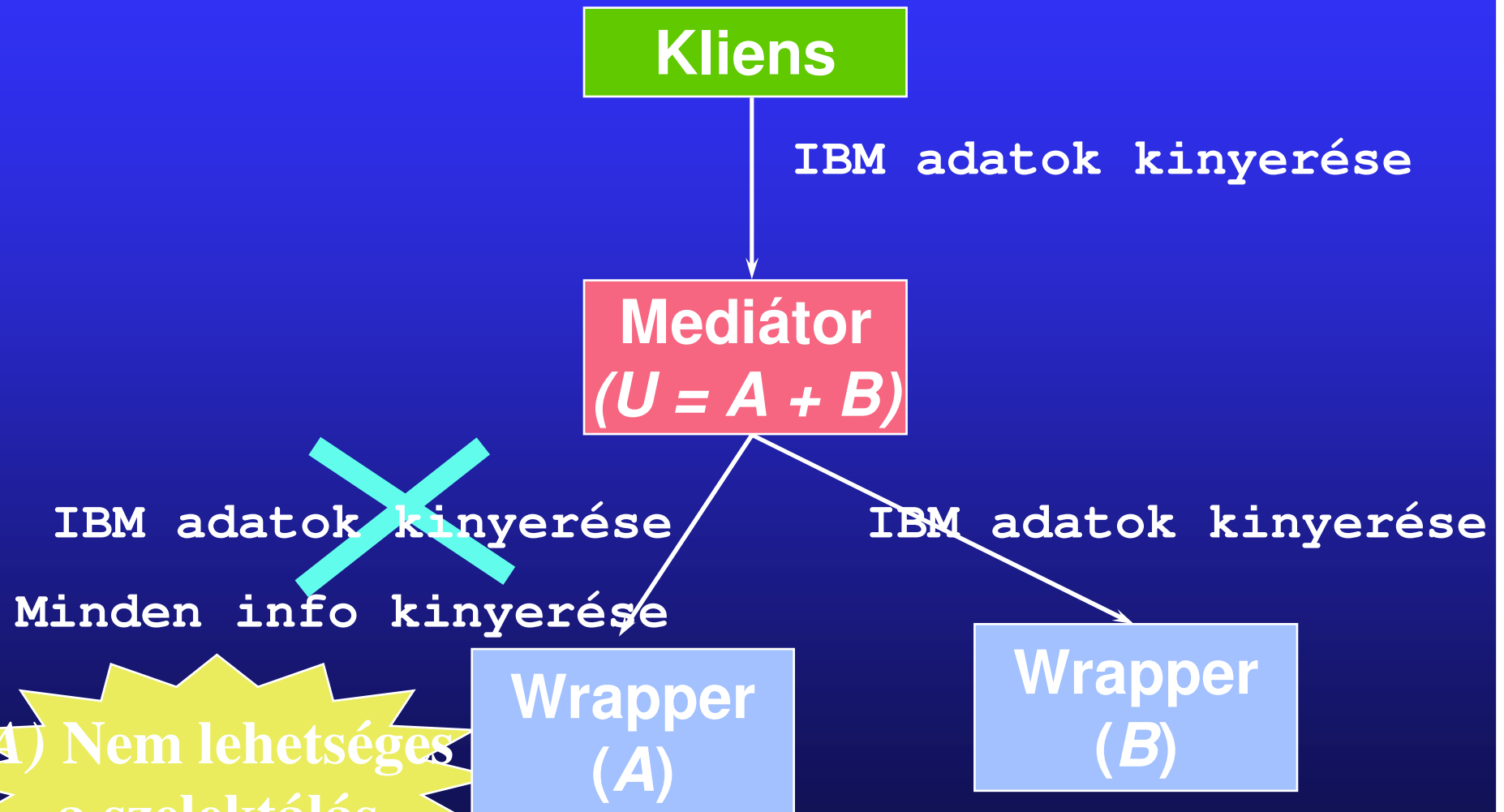
Példák

- HTML lapok
- SGML dokumentumok
- Gén adatbázisok
- Kémiai elemek
- Könyvtári információk
- Integrációk eredményei

Kihívás: Különböző forrás képeségek



Medátornak illeszkednie kell a források szolgáltatásaihoz



Adatreprezentáció

- ➡ • **Félig-strukturált adat reprezentáció**
 - Mediátor generálás
 - Wrapper generálás
 - Képességeket figyelembe vevő átírás

Félig-strukturált információk reprezentálása OEM modellel



OEM – Gráf reprezentáció

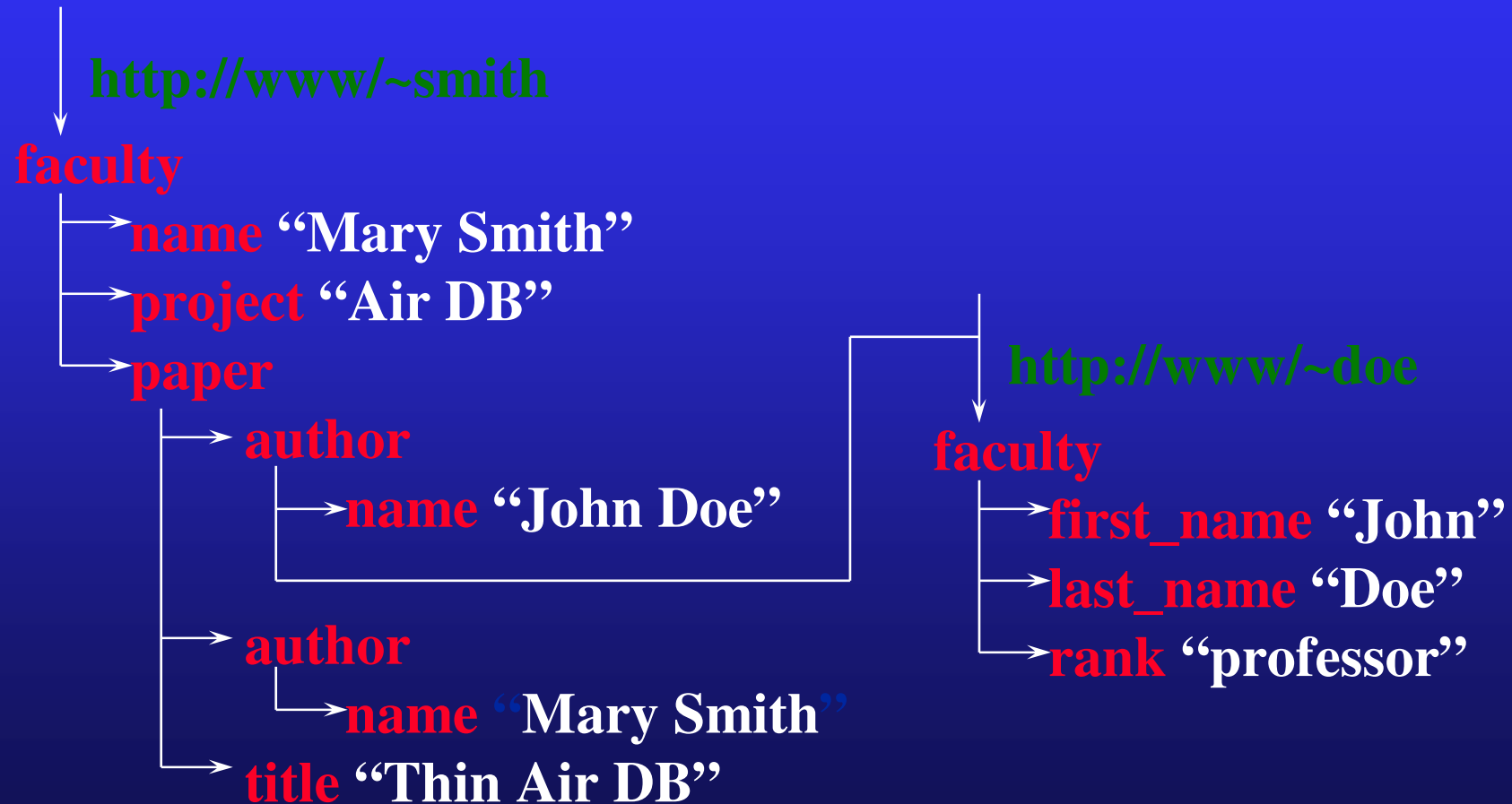
<http://www/~doe, faculty, {&f1,&l1,&r1}>
 <&f1, first_name, “John”>
 <&l1, last_name, “Doe”>
 <&r1, rank, “professor”>



http://www/~doe faculty

- first_name “John”
- last_name “Doe”
- rank “professor”

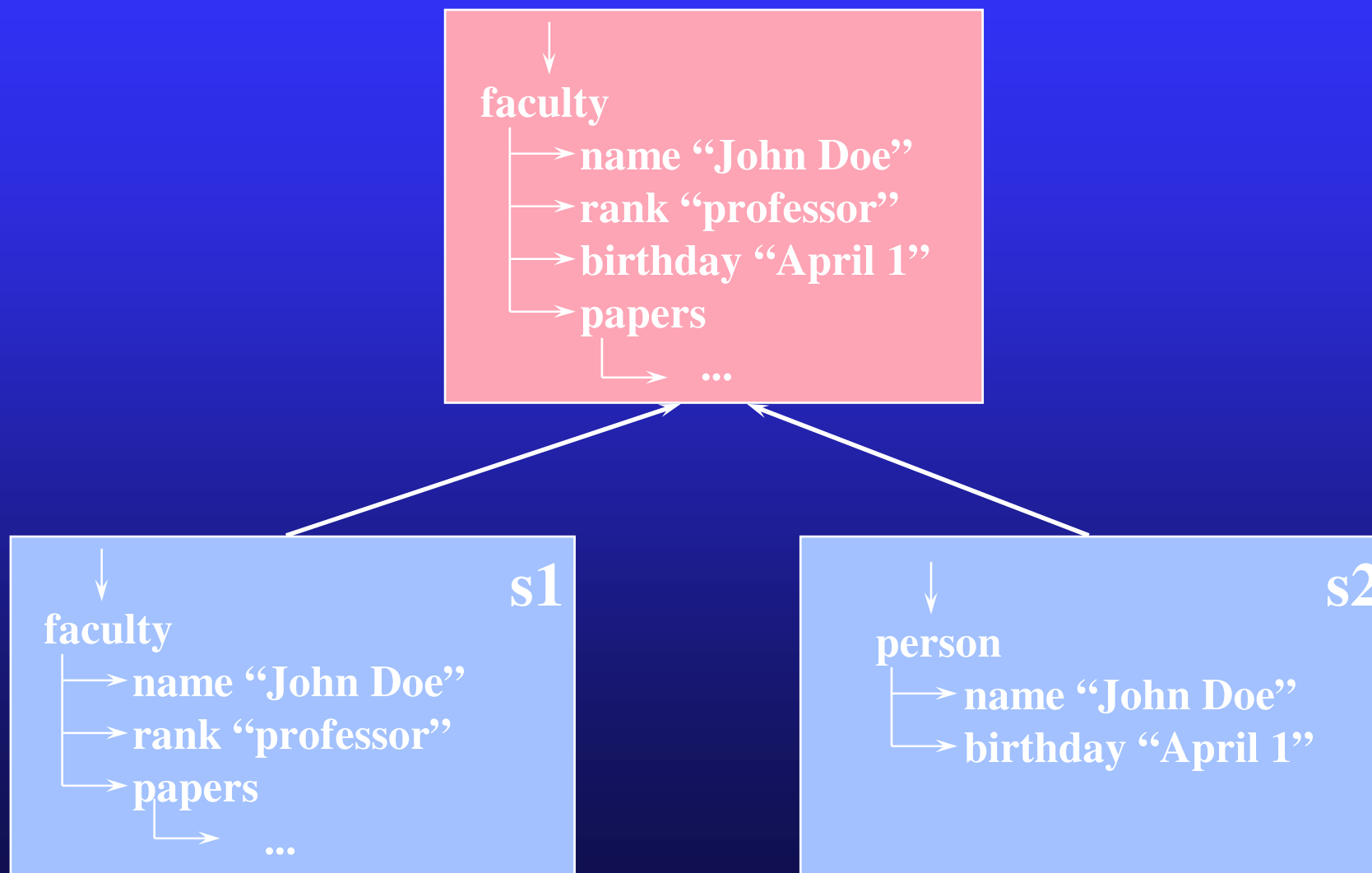
OEM struktúra tetszőleges gráf reprezentációt megenged



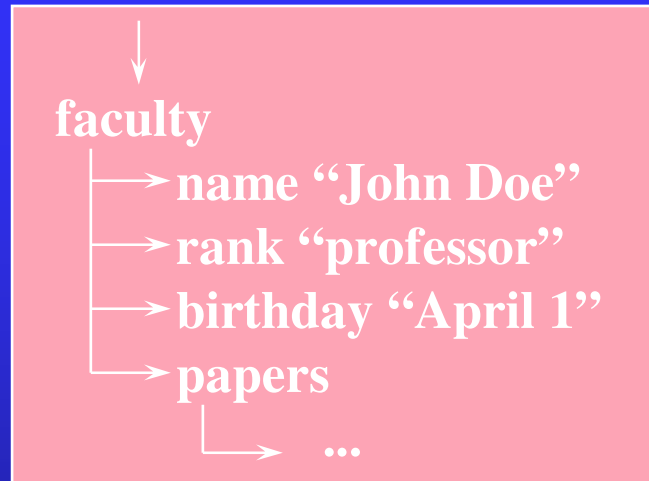
Áttekintés

- Félig-strukturált adat reprezentáció
- ➔ • Mediátor generálás
 - Mediátor specifikációs példák
 - Nyelv kifejezőerő
 - Megvalósítás és teljesítmény
- Wrapper generálás
- Képességeket figyelembe vevő átírás

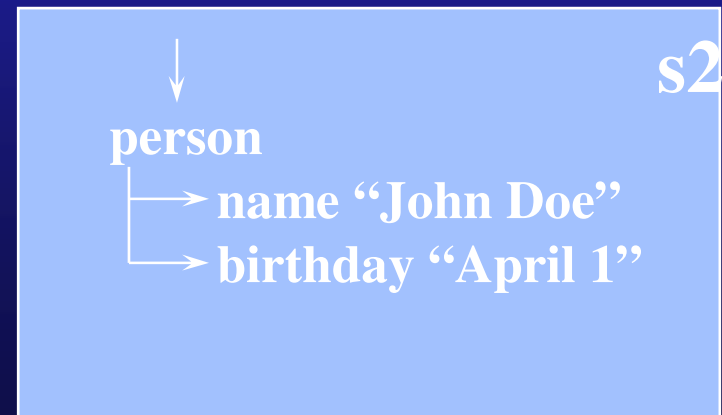
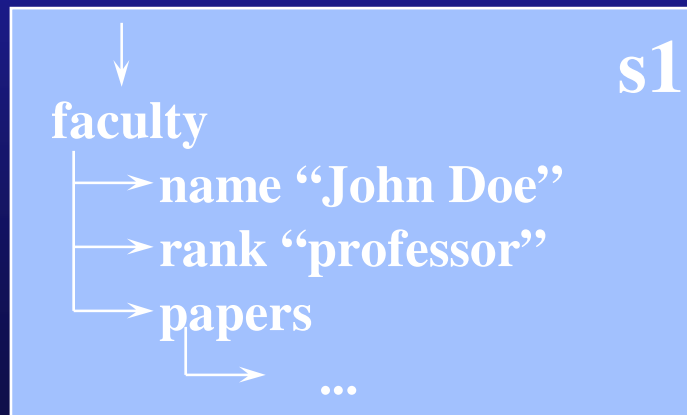
Információ egyesítés



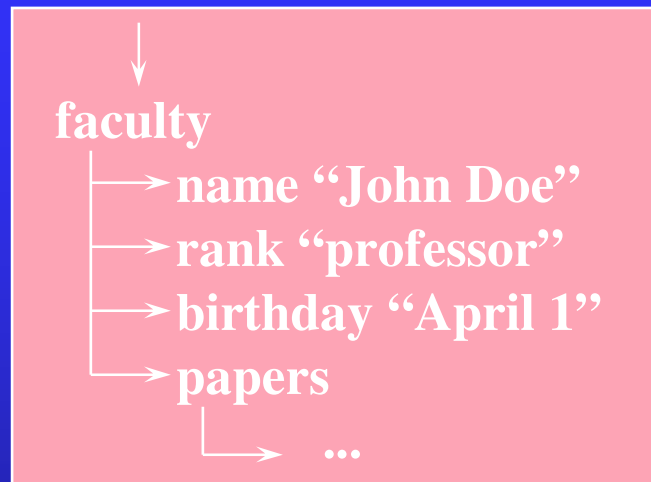
Mediátor specifikáció példa



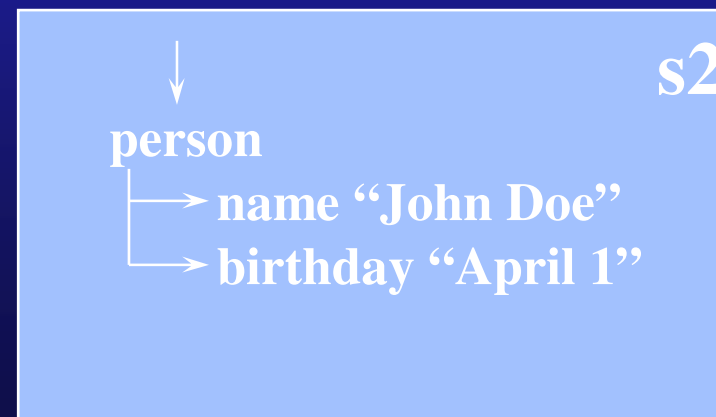
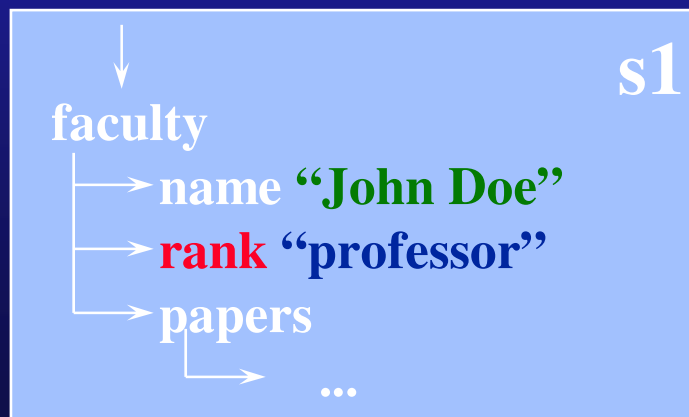
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{person } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s2$



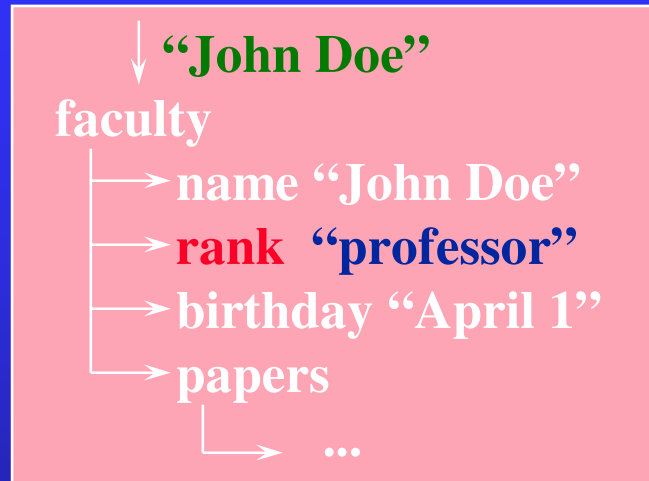
Mediátor specifikáció példa: Szabályok törzsének elemei



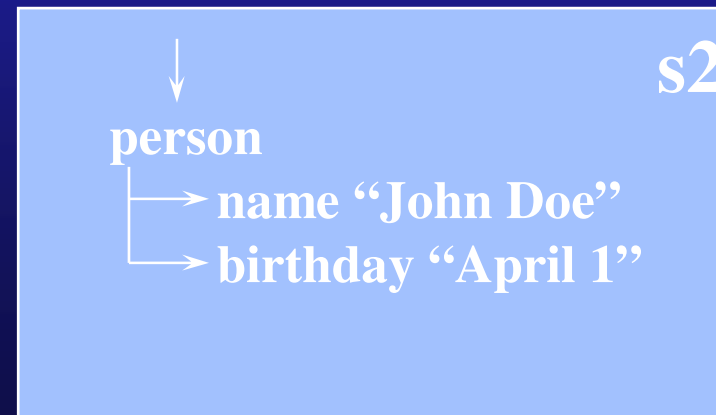
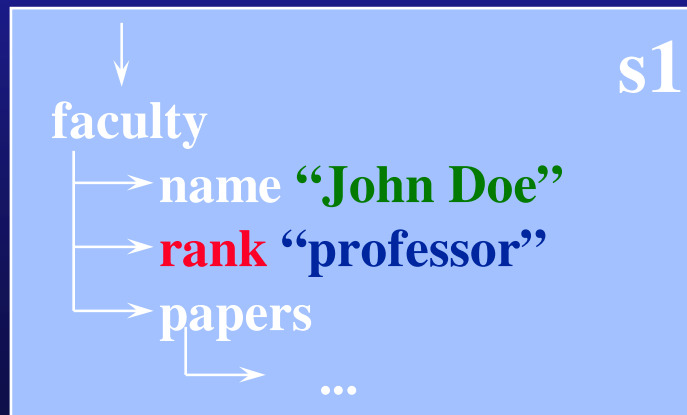
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{person } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s2$



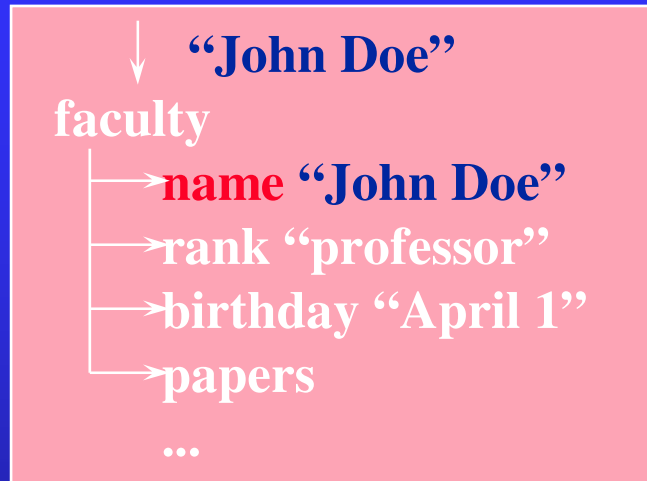
Mediátor specifikáció példa: Szabályok fejének értelmezése



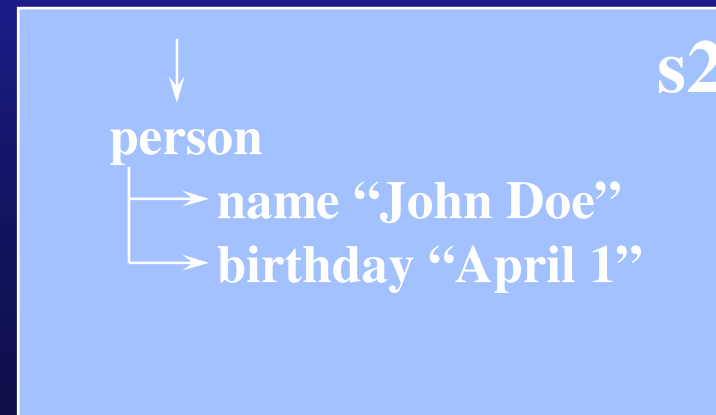
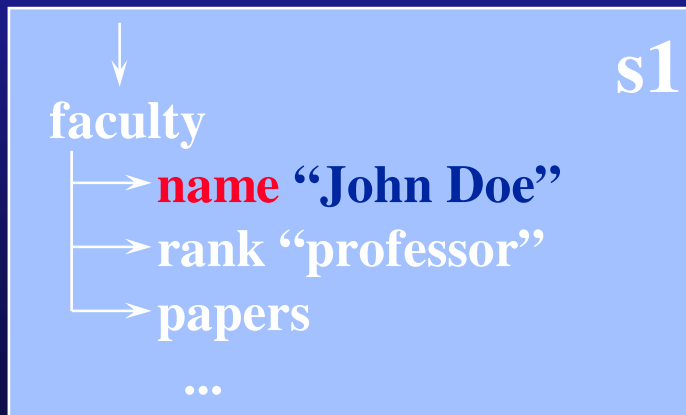
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{person } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s2$



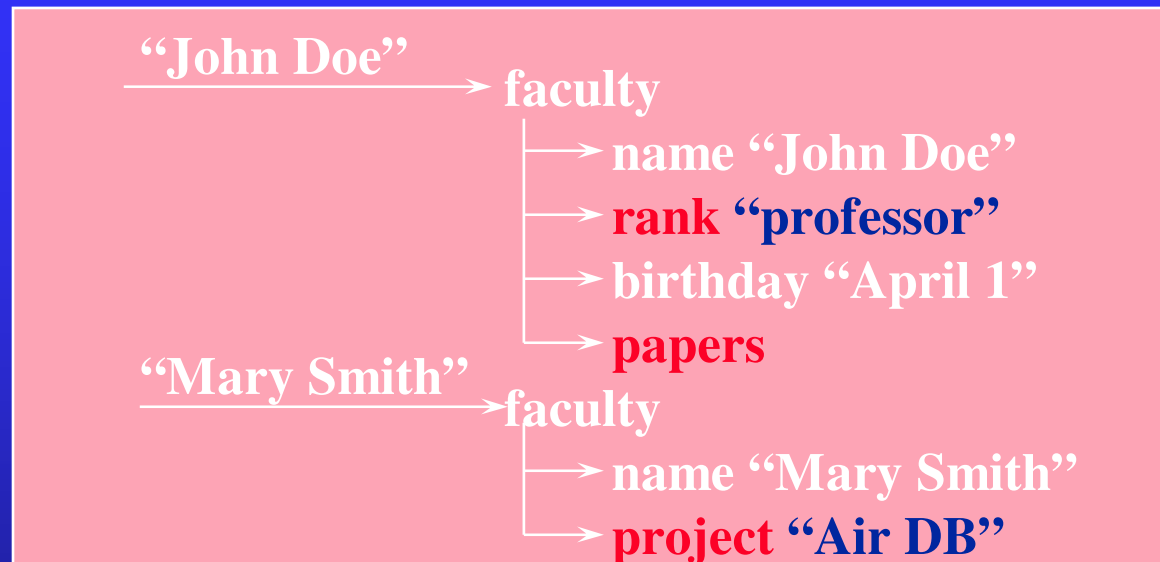
Szemantikusan azonosított objektum hozzáadása



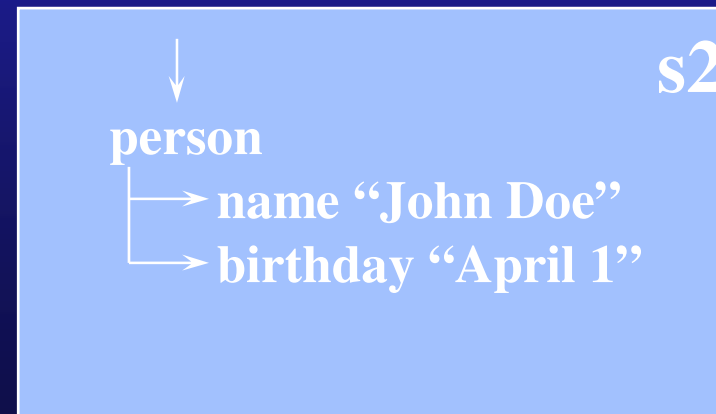
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{person } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s2$



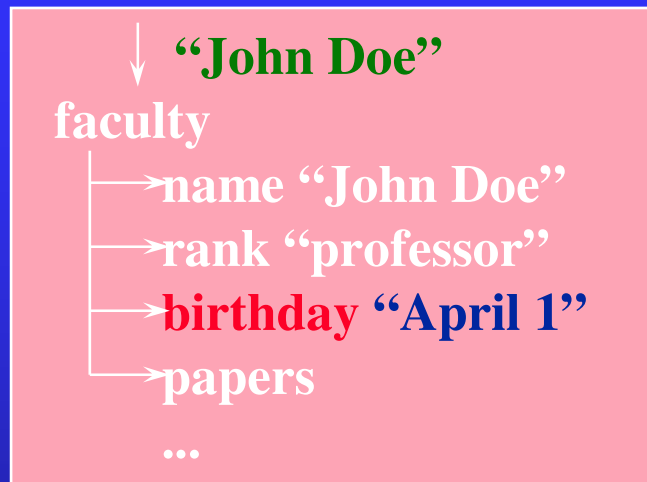
Irregularis és hiányzó információ



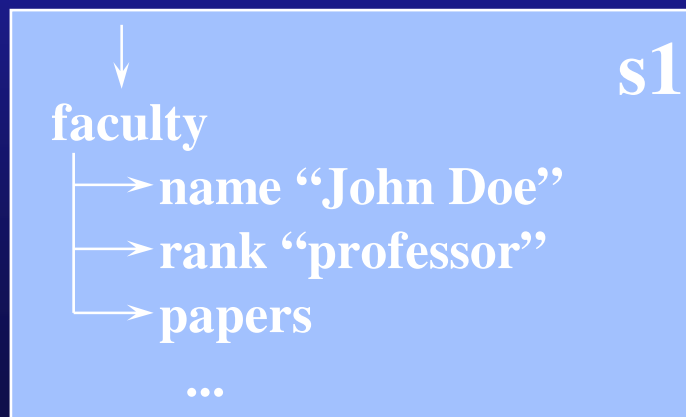
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$



A második szabály további (al)objektumot rendel a nézet objektumaihoz



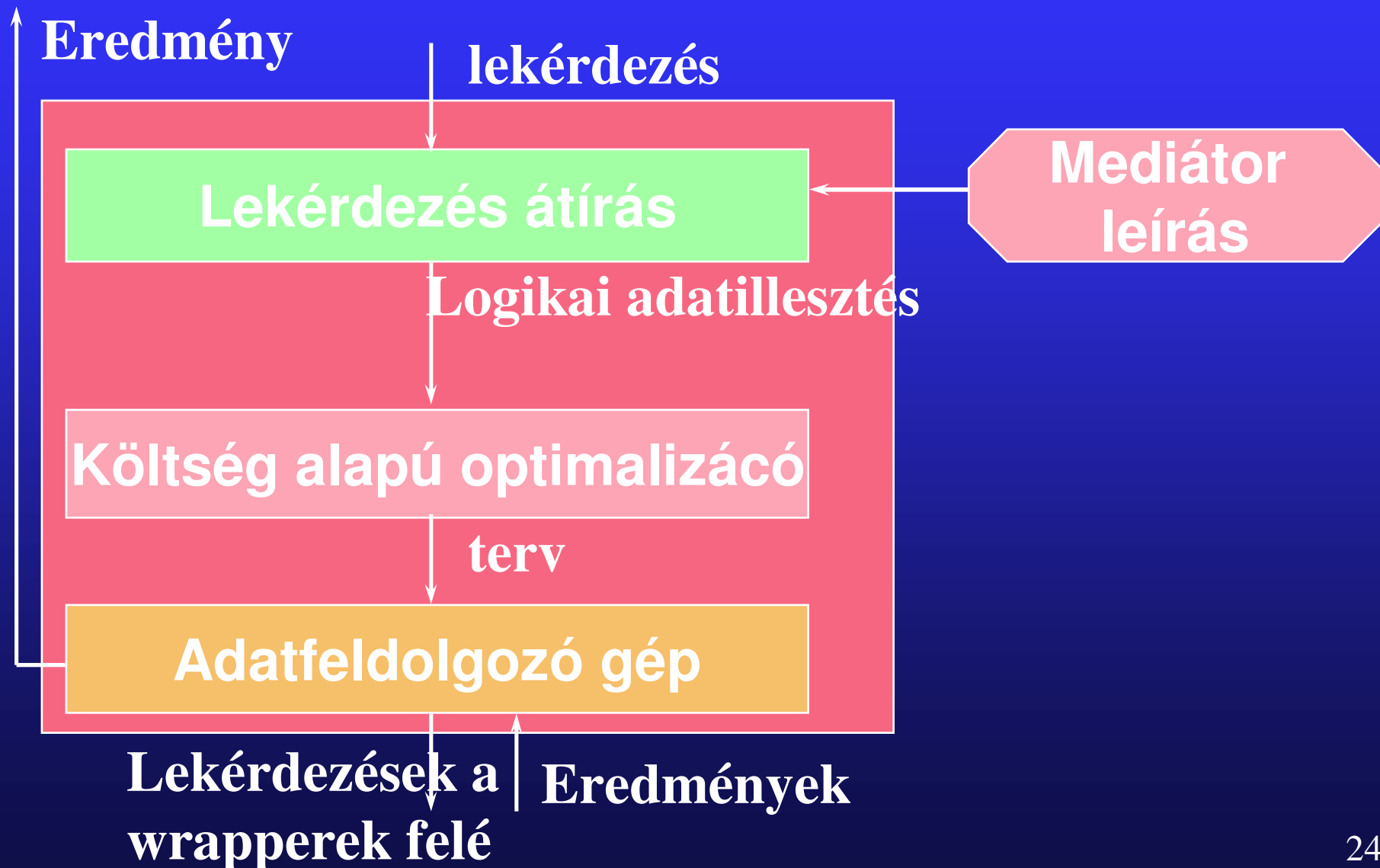
$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{faculty } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle \text{ :- } \langle \text{person } \{ \langle \text{name } N \rangle \langle L \ V \rangle \} \rangle @s2$



A nyelv kifejezőereje

- MSL nyelvvel megoldott problémák
 - Irregularitások
 - Hiányos struktúra információk
 - Kereszthivatkozások kezelése
 - **Inkonzisztens és redundáns adat**
 - **Tetszőleges illeszkedési kritérium**

Mediátor leírás interpreter architektúra



Lekérdezés átírás

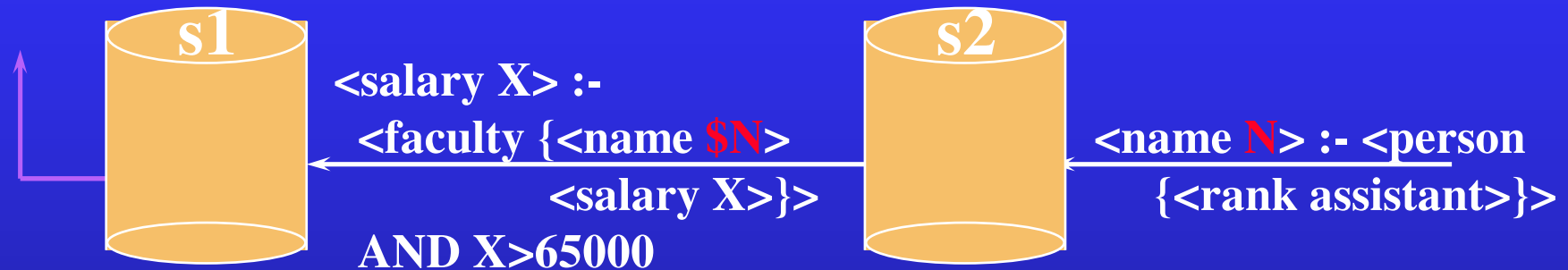
- $\langle N \text{ faculty } \{ \langle \text{salary } S \rangle \} \rangle :-$
 $:- \langle \text{faculty } \{ \langle \text{name } N \rangle \langle \text{salary } S \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle \text{rank } R \rangle \} \rangle$
 $:- \langle \text{person } \{ \langle \text{name } N \rangle \langle \text{rank } R \rangle \} \rangle @s2$
- $\langle \text{well-paid } \{ \langle \text{name } N \rangle \langle \text{salary } X \rangle \} \rangle$
 $:- \langle N \text{ faculty } \{ \langle \text{salary } X \rangle \langle \text{rank assistant} \rangle \} \rangle$
 $\text{AND } X > 65000$

Lekérdezés átírás: feltételek átadása a források fele

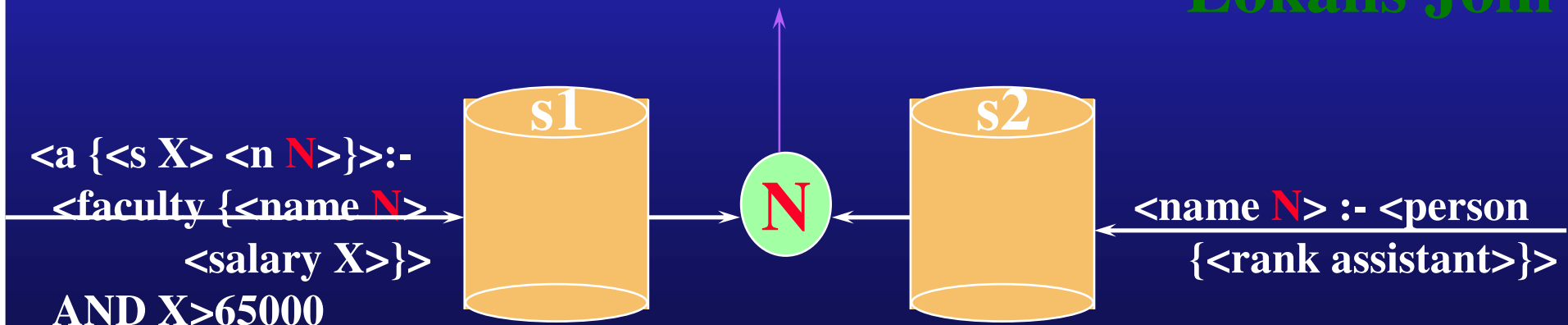
- $\langle N \text{ faculty } \{ \langle \text{salary } S \rangle \} \rangle :-$
 $:- \langle \text{faculty } \{ \langle \text{name } N \rangle \langle \text{salary } S \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle \text{rank } R \rangle \} \rangle$
 $:- \langle \text{person } \{ \langle \text{name } N \rangle \langle \text{rank } R \rangle \} \rangle @s2$
- $\langle \text{well-paid } \{ \langle \text{name } N \rangle \langle \text{salary } X \rangle \} \rangle$
 $:- \langle N \text{ faculty } \{ \langle \text{salary } X \rangle \langle \text{rank assistant} \rangle \} \rangle$
 $\text{AND } X > 65000$
- **logical datamerge program**
 $\langle \text{well-paid } \{ \langle \text{name } N \rangle \langle \text{salary } X \rangle \} \rangle$
 $:- (\langle \text{faculty } \{ \langle \text{name } N \rangle \langle \text{salary } X \rangle \} \rangle$
 $\text{AND } X > 65000) @s1$
 $\text{AND } \langle \text{person } \{ \langle \text{name } N \rangle \langle \text{rank assistant} \rangle \} \rangle @s2$

Lekötések átadása & Lokális Join Tervek

Lekötések átadása



Lokális Join



Lekérdezés dekomponálása ismeretlen eredet figyelembevételével

<N faculty {<L V>}> :- <faculty {<name N> <L V>}>@s1

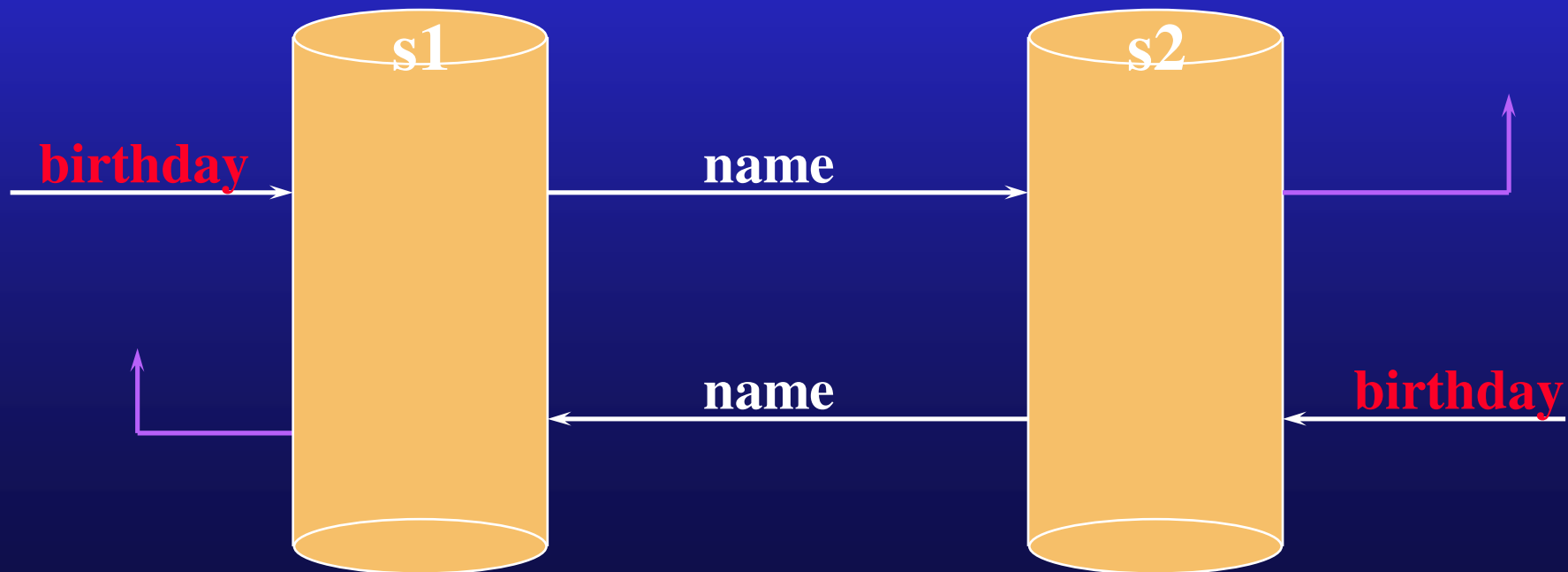
<N faculty {<L V>}> :- <person {<name N> <L V>}>@s2

<X faculty {<S Y>}> :- <X faculty {<birthday “1/20”> <S Y>}>

A terv figyelembe vesz minden lehetséges forrást a **birthday** értékére

$\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{faculty } \{ \langle \text{name } N \rangle \ \langle L \ V \rangle \} \rangle @s1$
 $\langle N \text{ faculty } \{ \langle L \ V \rangle \} \rangle :- \langle \text{person } \{ \langle \text{name } N \rangle \ \langle L \ V \rangle \} \rangle @s2$

$\langle X \text{ faculty } \{ \langle S \ Y \rangle \} \rangle :- \langle X \text{ faculty } \{ \langle \text{birthday "1/20"} \rangle \ \langle S \ Y \rangle \} \rangle$



Áttekintés

Félig-strukturált adat reprezentáció

Mediátor generálás

 **Wrapper generálás**

Képességeket figyelembe vevő átírás

Lekérdezés lefordítás wrapperre



Gyors lekérdezés lefordítás

SELECT * FROM person
SELECT * FROM person
WHERE name="Smith"

Sablon
Interpreter

SELECT * FROM person
{emit "find -all" }
SELECT * FROM person
WHERE name=\$N
{emit "find -n \$N" }

Eredmény
lefordító

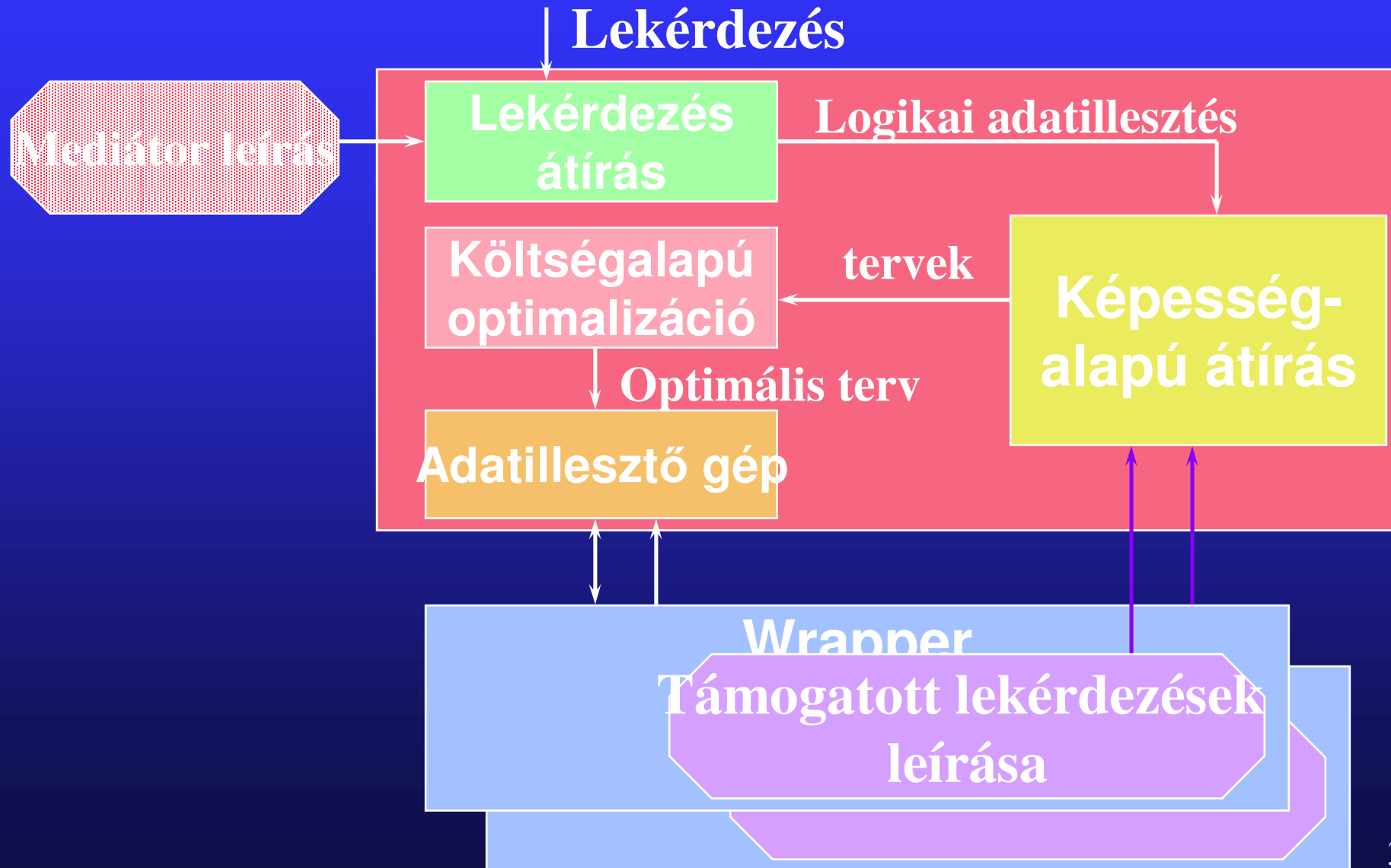
Minden keresése
Példány keresése

Forrás

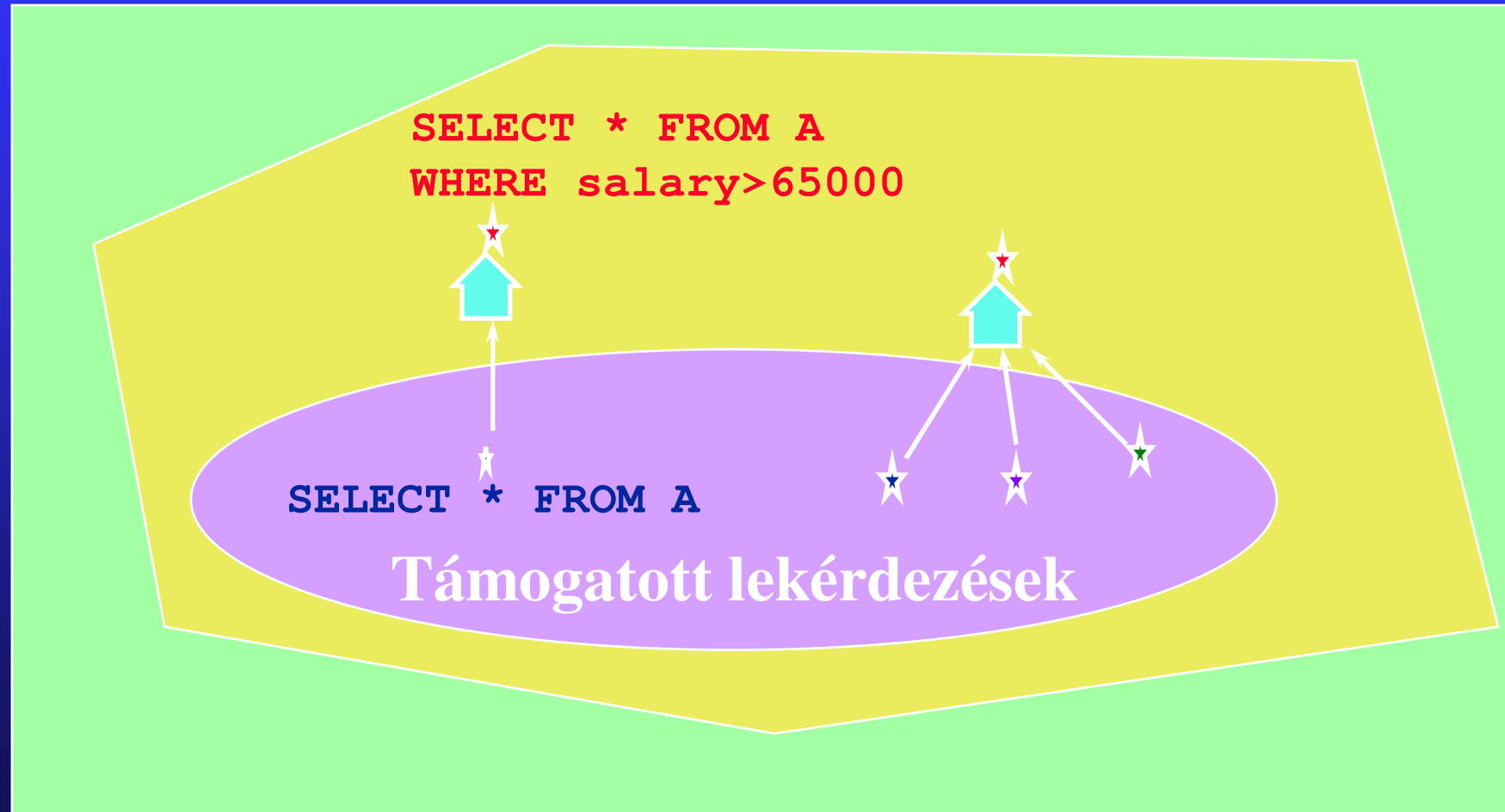
Lekérdezés halmazok leírása

- Rekurzív non-terminálisok használata
- Példa:
 - Munka leírások tartalmazznak w1, w2, stb. elemeket
 - `SELECT subset(person) FROM person`
`WHERE \CJob`
`\CJob : job LIKE $W AND \CJob`
`\CJob : TRUE`

Képeségek figyelembevétele az átírásnál



Átalakító megtalálja a támogatott tervek leírását



Képesség alapú átírás megtalálja a legpontosabb támogatott tervet



Képesség alapú átírás architektúra



TSIMMIS összefoglalás

- Rendszer integráció heterogén forrásokra
- Kihívások és megoldások
 - semistructured data & incomplete schema knowledge
 - Megfelelő lekérdezés definiáló nyelv és lekérdezés lefordítás
 - Eltérő képességű források kezelése
 - Lekérdezés lefordítási módszer
 - Képesség alapú átírás