

# A NAVA adatmodellje és adatgyűjtő protokollja

Szakadát István, BME  
Knapp Gábor, NAVA

## NAVA

- ◆ Technológiai rendszer
  - Általános OA modell
  - Teljesen digitális archívum (HSM, szoftver transzkódolás)
  - Nincs HighRes (csak durva vágás)
  - Tárolás és hozzáférés biztosítása
- ◆ Közgyűjtemény
  - Feldolgozás, visszakereshetőség (meta!)
  - NDA adatszolgáltatás, elérési pontok
- ◆ Kihívás (hardware, software, wetware): (valós) IDŐ
  - Ma (2005-04-18-tól): 7 archivátor, 1000 műsorszám (kb. 800 óra)
  - Feldolgozási sebesség (tapasztalat): 1xHW, 1xSW, 2xWW
  - **Amit csak lehet (adatcserét is), automatizálni kell!**

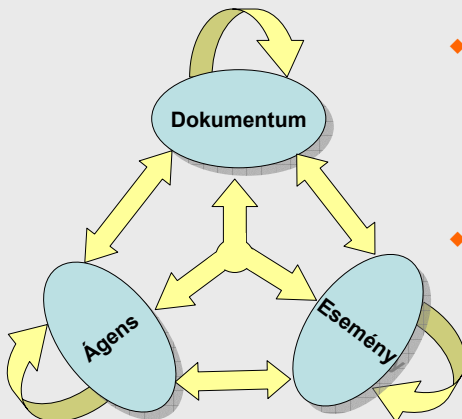
## Keretek

- ◆ NAVA küldetés (NAVA törvény, 2004/CXXXVII)
  - Sugárzott kötelempéldányok gyűjtése (műsorszórásból)
  - Önkéntesen felajánlott anyagok
  - Nyilvánosan elérhető katalógus építése (műsorszolgáltatók adatai alapján)
  - Műsorok megtekinthetővé tétele zárt hálózaton
- ◆ Végrehajtási rendeletek (nemsokára)
  - NAVA szolgáltatásának szabályai (jog)
  - NAVA szolgáltatásának szabályai (technika, minőség)
  - A kötelempéldány (tartalom) szolgáltatásának szabályai
  - **A kötelempéldány (metaadat) szolgáltatásának szabályai**

## Metadatum szolgáltatás

- ◆ Törekvés
  - Hasznos metaadatok (a keresés szempontjából)
  - Gyorsan beilleszthető metaadatok (hamar elérhető)
  - Minél kisebb emberi energia ráfordítás (automatikus rendszer)
  - Minél egyszerűbb és olcsóbb technológia
- ◆ Ennek érdekében cél
  - Egységes, de rugalmas adatbázis séma
  - Elektronikusan generálható, feldolgozható rekordok
  - Nyílt adatgyűjtő protokoll (OAI-PMH)

## NAVA adatmodell



### ◆ Relációs adatmodell

- Kifinomult ágens kezelés
- Hierarchikus dokumentum szerkezet
- Mozgóképre optimalizált

### ◆ DE

- Nemzetközi szabványok
- Dublin Core elvek
- NDA ajánlások
- OAI-PMH export/import

## Modellezési elvek - 1

### ◆ A múlt öröksége

- könyvtári tudás
- tartalmi vs. formai leírás

### ◆ Dokumentum

- általános dokumentummodell (MNFA)
- típus
  - műsorszám vs. audiovizuális dokumentum
  - műsorszám vs. egyéb dokumentum
- rész-egész
  - sorozat-elem (eleme vs. tartalmazása reláció)
  - teljes dokumentum - dokumentumszegmens

### ◆ 3+1\*n individuum (tulajdonnév)

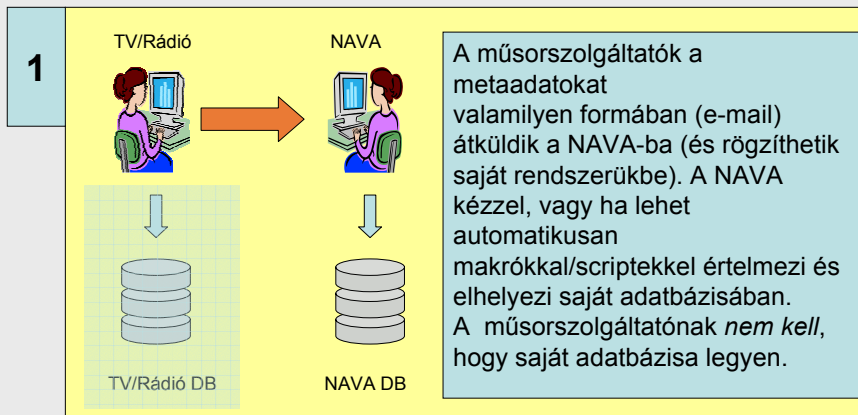
- ágensek (2)
  - személy
  - testület
- hely (1)
- dokumentumcím (n)

## Modellezési elvek - 2

- ◆ Idő
  - egységes eseménykezelés
  - időpontosság
- ◆ Szereplélációk
  - profession (foglalkozás: diszpozíció szerep)
  - character (karakter: játszott szerep)
  - position (beosztás: birtokolt szerep)
- ◆ A jövő lehetősége
  - NDA-integráció
  - típusnév tartomány (namespace kritika)

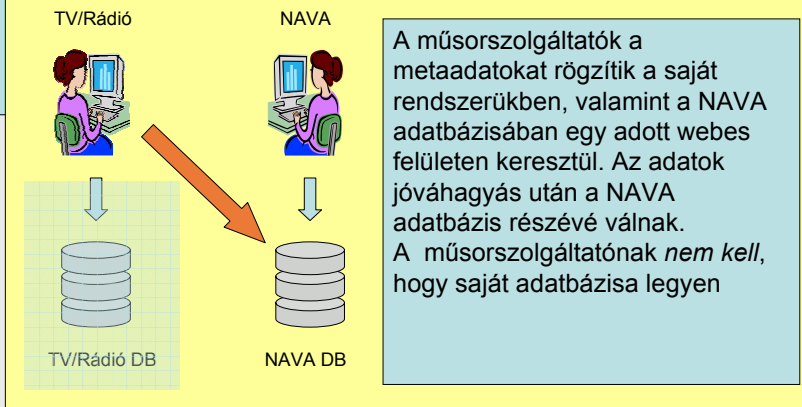


## Off-line adatcsere



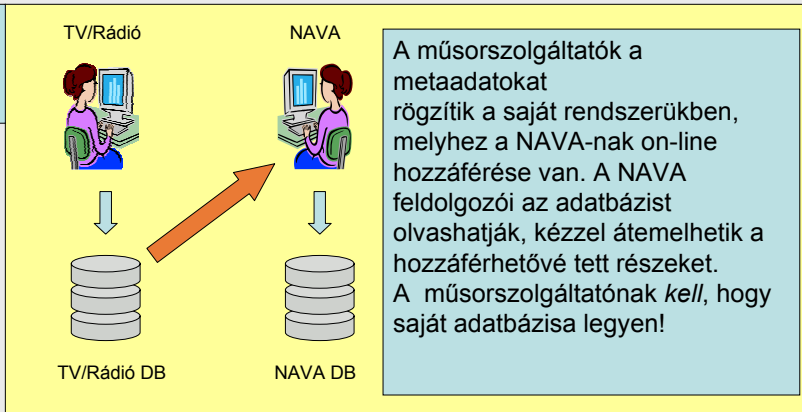
## On-line adatcsere/Push

2



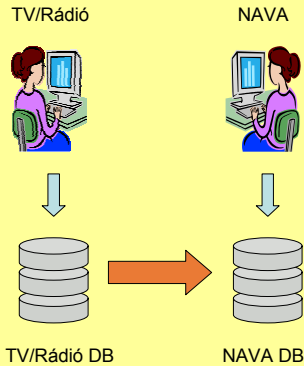
## On-line adatcsere /Pull

3



# Automatizált adatcsere

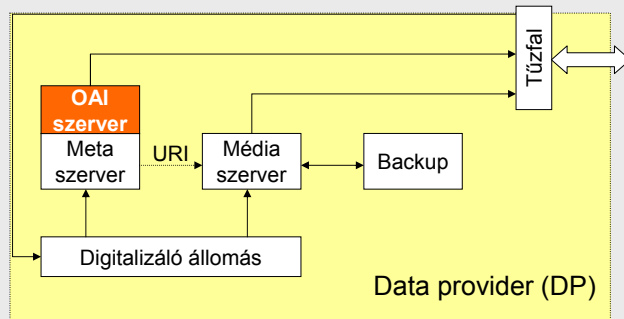
4



A műsorszolgáltatók a metaadatokat rögzítik a saját rendszerükben, melyen fut egy OAI-PMH szerű alkalmazás, melyet a NAVA kliense elér, és a metaadatokat jóváhagyás után a helyére teszi. A műsorszolgáltatónak *kell*, hogy saját adatbázisa legyen!

# OAI-PMH

- ◆ Szabványos protokoll
  - Near-line elérés
  - Tűzfal-tűrő, de biztonságos
  - Olcsó, sok szabad szoftver, tapasztalat
  - NDA ajánlás...
  - Automatikus



## Metaadat-csere rekord felépítése

- ◆ **Rekord:**
  - Minden adatelemből legalább egyet kell tartalmaznia
  - Minden adatelem tetszőleges számban ismétlődhet
- ◆ **Adatelemek:**
  - Adatelem neve (pl. Közreműködő)
  - Minősítő: a megadott érték jelentésének pontosítása (pl. Író)
- ◆ **Minősítők:**
  - Alapértelmezésben az NDA ajánlások szerint

## Elsődleges metaadatok

Adatelem	Minősítő	Meghatározás	Megjegyzés
Csatorna		Nyilvánosságra hozó (műsorszolgáltató) csatorna megnevezése	
Dátum		Az a naptári nap, amikor adott adásmenet alapján a nyilvánossághoz közvetítik a műsorszámot (év, hó, nap), illetve a műsorszám sugárzásának időpontja (óra, perc)	
Adáshossz		A műsorszám hossza (perc)	
Cím	Címtípus	A műsorszámokban szereplő szó, kifejezés vagy egyéb jelcsoport, amely a műsorszám megnevezésére szolgál	A Címtípus a Cím jellegének megadására szolgál (pl. eredeti cím, alcím, fiktív cím stb.). Ha a műsorszám sorozat tagja, azt a Kapcsolat elemnél kell megadni.
Műfaj	Műfajtípus	A műsorszám típusát, műfaját, általánosabb sajátosságait jelölő kifejezés vagy besorolási szám (pl. EBU-TV)	A műfajtípus a különböző osztályozási rendszereket különbözteti meg (ETSI, EBU)
Közreműködő	Szerep	Az a személy vagy testület, aki/amely a műsorszám létrejöttében alkotóként vett részt, vagy valamilyen járulékos szerepet játszott.	A Szerep a közreműködő szerepvállalásának jellegét adja meg (pl. rendező, író, operatőr, vágó, riportalany, fordító, színész, zeneszerző). Célszerű a közreműködőket fontossági sorrendben megadni.
Nyelv	Nyelvhasználat	A műsorszám tartalmának megértéséhez szükséges nyelvek	A Nyelvhasználat minősítő azt adja meg, hogy a megadott nyelv milyen módon fordul elő (pl. eredeti, szinkron, felirat)
Kapcsolat	Kapcsolattípus	Hivatkozás egy kapcsolódó forrásra, kapcsolat más műsorszámokkal	A Kapcsolattípus a kapcsolat jellegét adja meg (pl. sorozat, ismétlés változat)

## Másodlagos metaadatok (opcionális)

Adatelem	Minősítő	Meghatározás	Megjegyzés
Téma	Tématípus	A dokumentum tartalmát leíró tárgyszavak, kifejezések, osztályozási jelzetek.	Javasolt, hogy a tárgyszavakat a műsorszolgáltatók egységesen közös kódszótárból válasszák. A tématípus meghatározza, hogy Fogalomról, Személyről, Testületről vagy földrajzi névről van szó (és ennek megfelelően más-más névtérből származik az adat).
Leírás	Leírástípus	A dokumentum tartalmának, cselekményének, térben, időben történő leírása.	A Leírás típus lehet tartalomjegyzék, annotáció, vagy akár szabadon választott tárgyszavak felsorolása.

- Megjegyzés:
  - Téma (kötött szavas)
  - Leírás (szabad szöveges)

## Köszönjük a figyelmet!

Szakadát István, BME, [syi@axelero.hu](mailto:syi@axelero.hu)

Knapp Gábor, NAVA, [knapp@nava.hu](mailto:knapp@nava.hu)