A HuComTech adatbázis a Nyelvtudományi Intézet nyelvi archívumában

Oravecz Csaba

oravecz.csaba@nytud.mta.hu

MTA Nyelvtudományi Intézet

Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály

Nyelvi archívum

Előzmények

- digitálisan hozzáférhető nyelvi adatok mennyiségének ugrásszerű növekedése (hangzó anyagok, multimodalitás)
- ad hoc tárolási megoldások $\rightarrow$ adattemetők

Kívánalmak

- változatos megjelenési formájú nyelvi adatok hosszútávú megőrzése
- a gyűjtött adatok szervezett és biztonságos tárolása
- hatékony és szabványos hozzáférés biztosítása

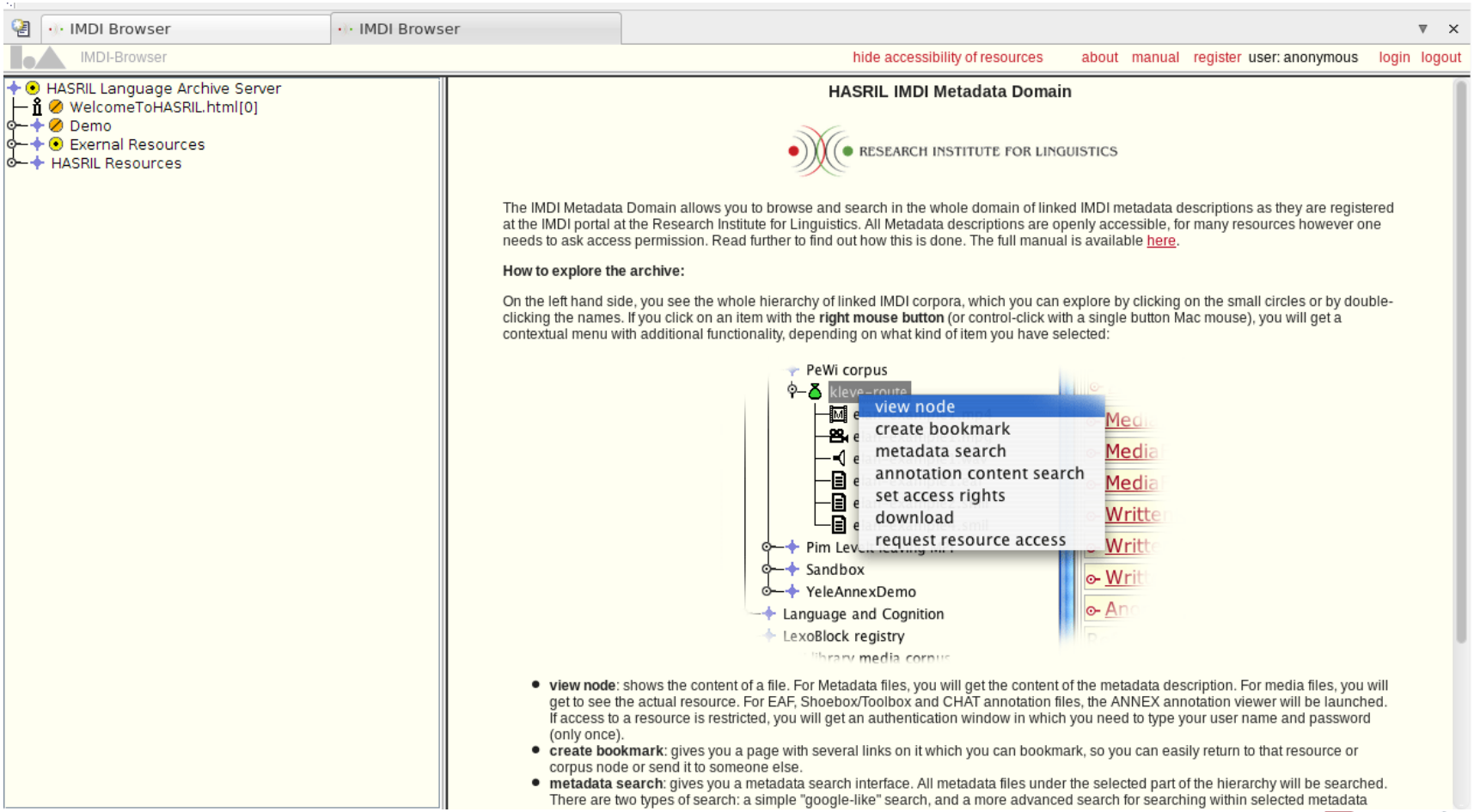
Nyelvi archívum

Megoldás

- kutatási infrastruktúrába illeszkedő adatszolgáltató központ
- metaadatok sztenderdizált módon történő rögzítése
  - IMDI: ISLE Meta Data Initiative
    (ISLE: International Standard for Language Engineering)
  - CMDI: Component MetaData infrastructure
- kereshetőség, hozzáférés szabályozhatósága

MTA NYTI archívum

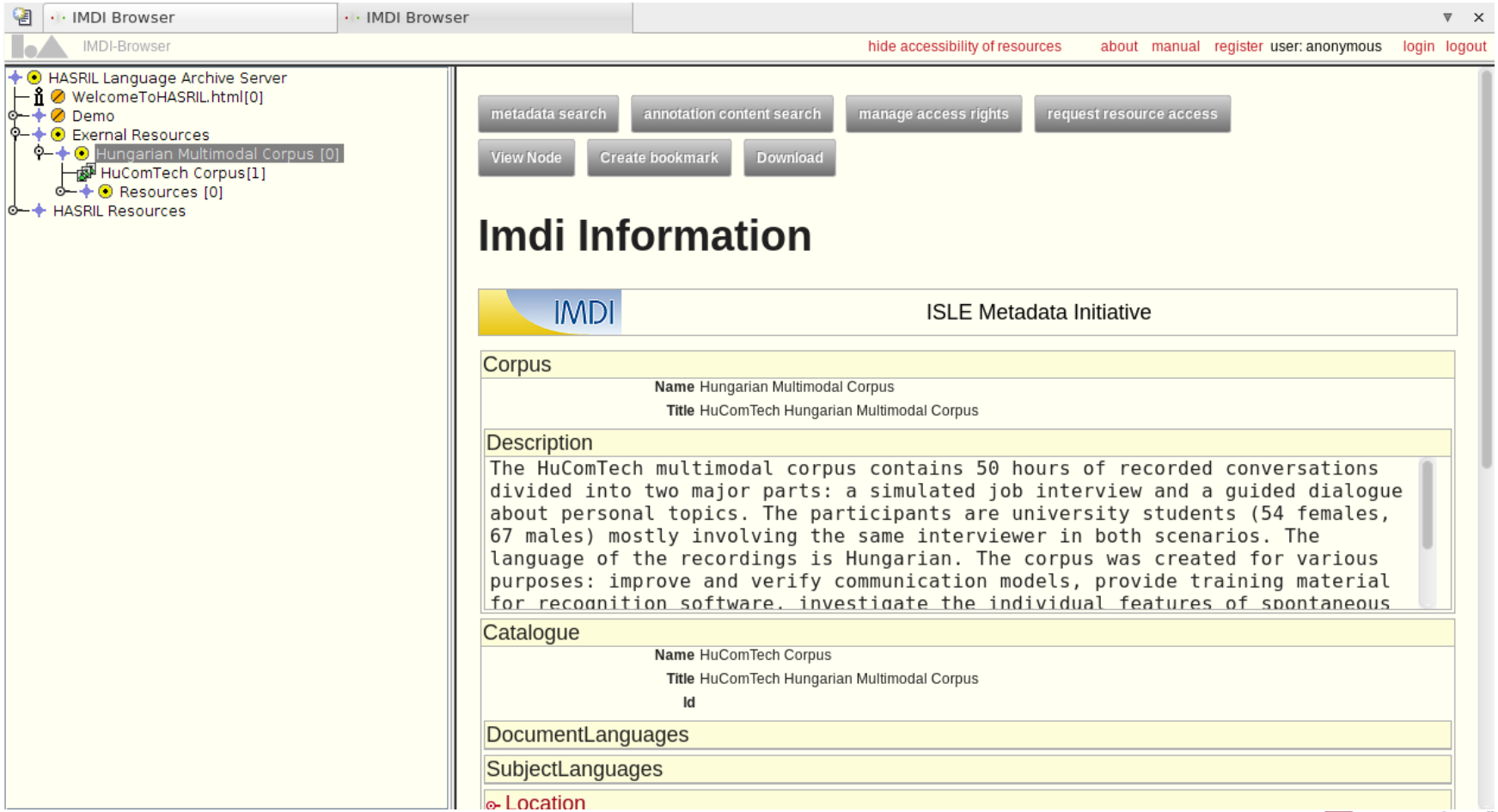
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

HucomTech adatbázis csomópont

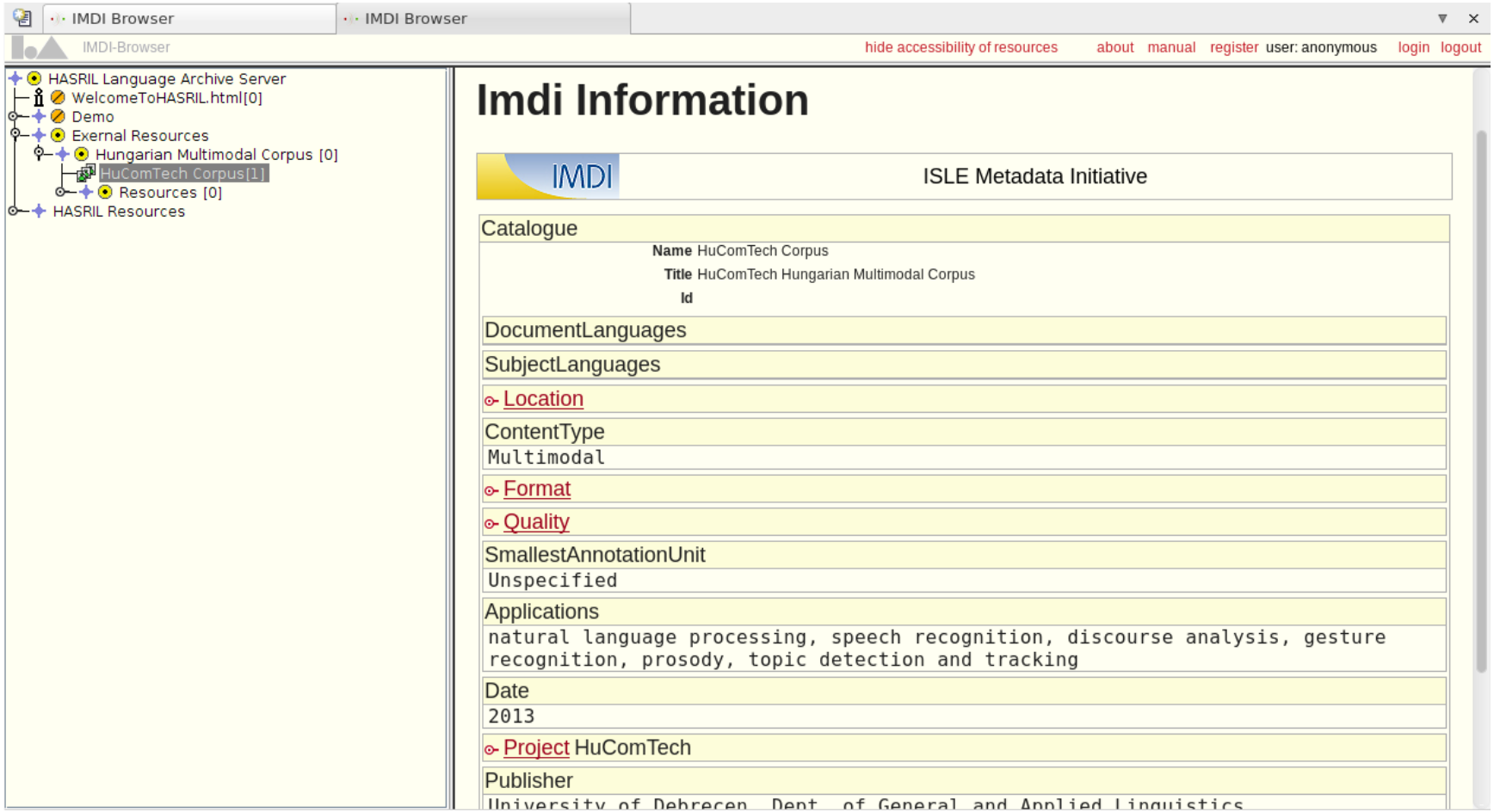
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

HucomTech szabványos metaadatok

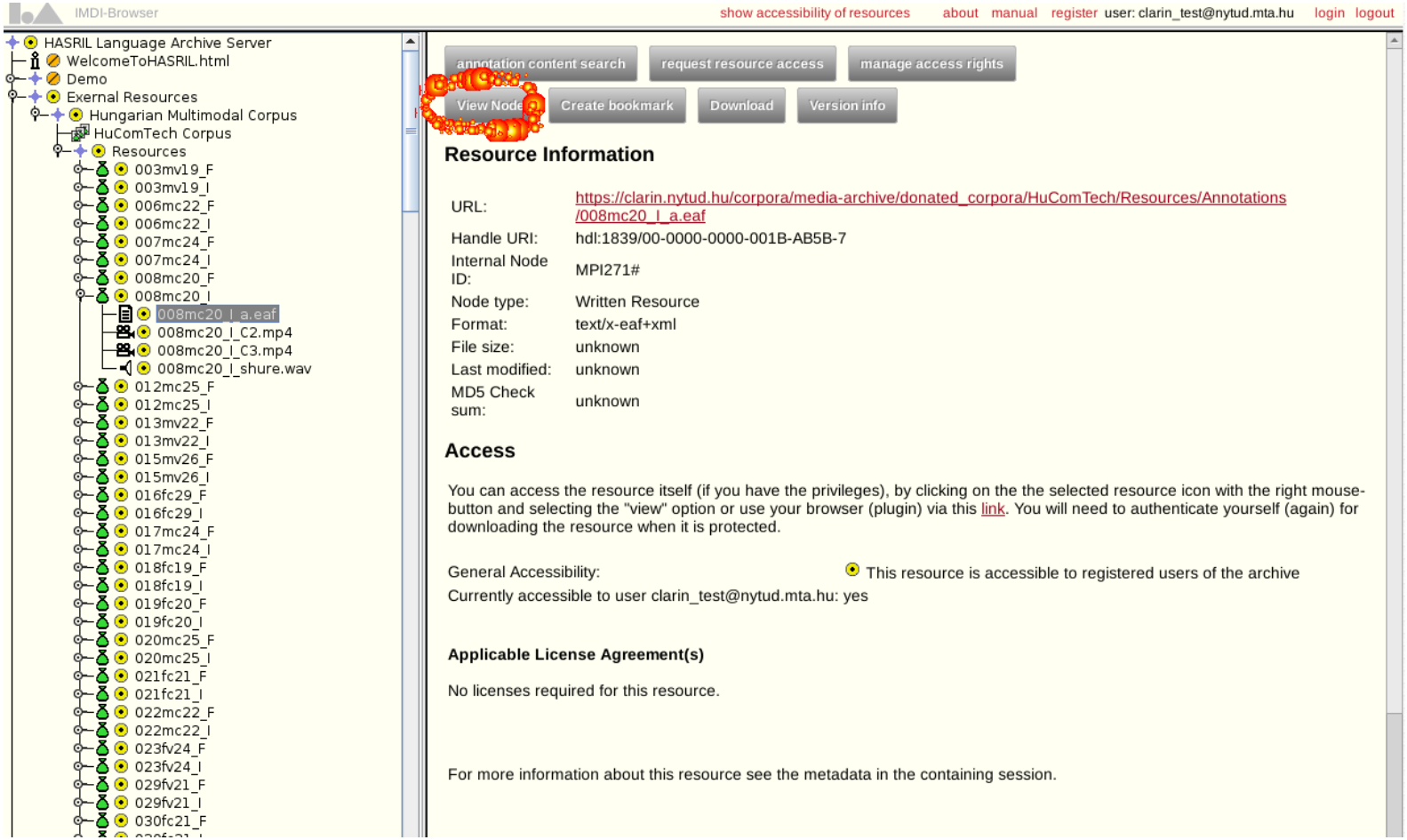
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

Az annotáció vizsgálata

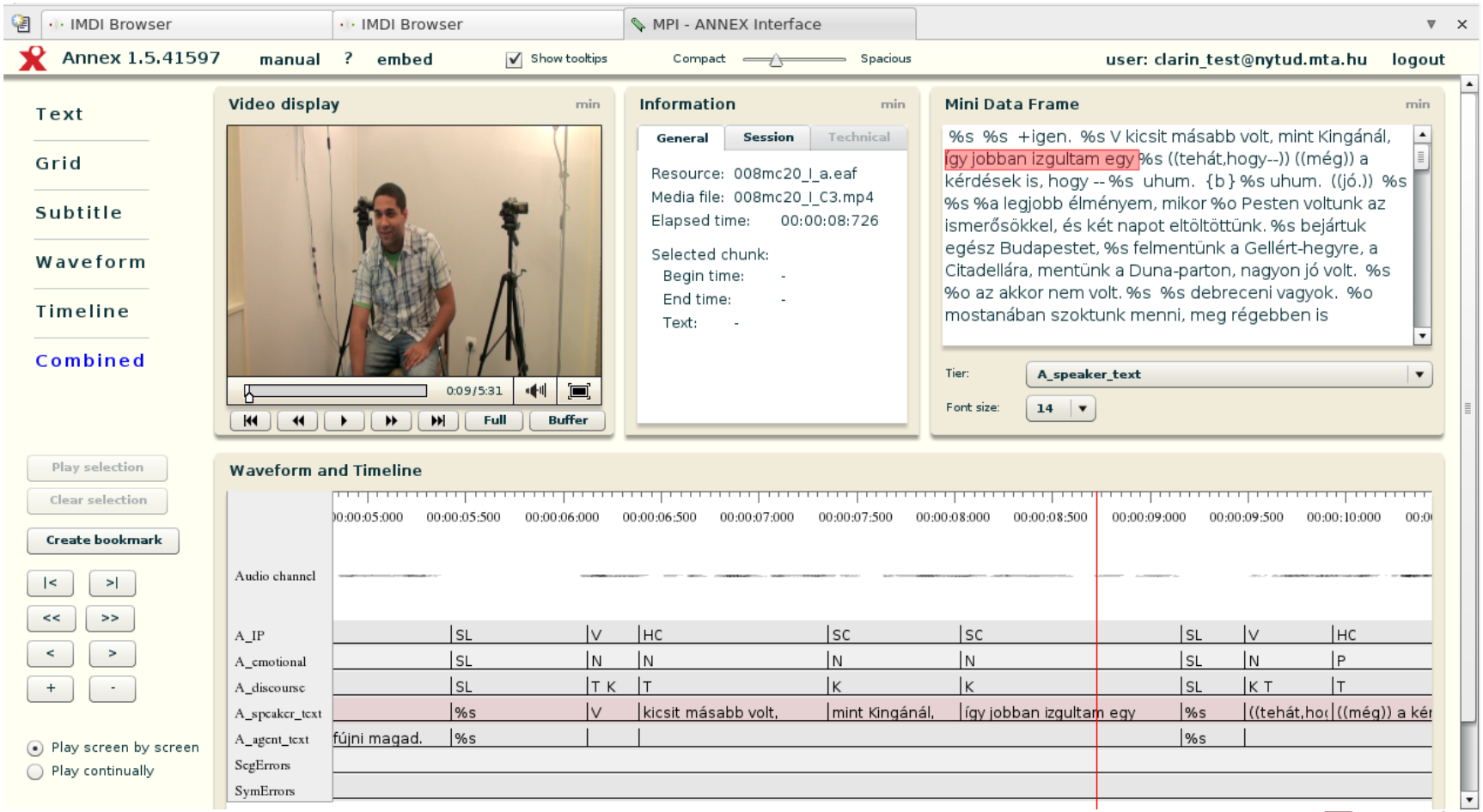
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

ANNEX

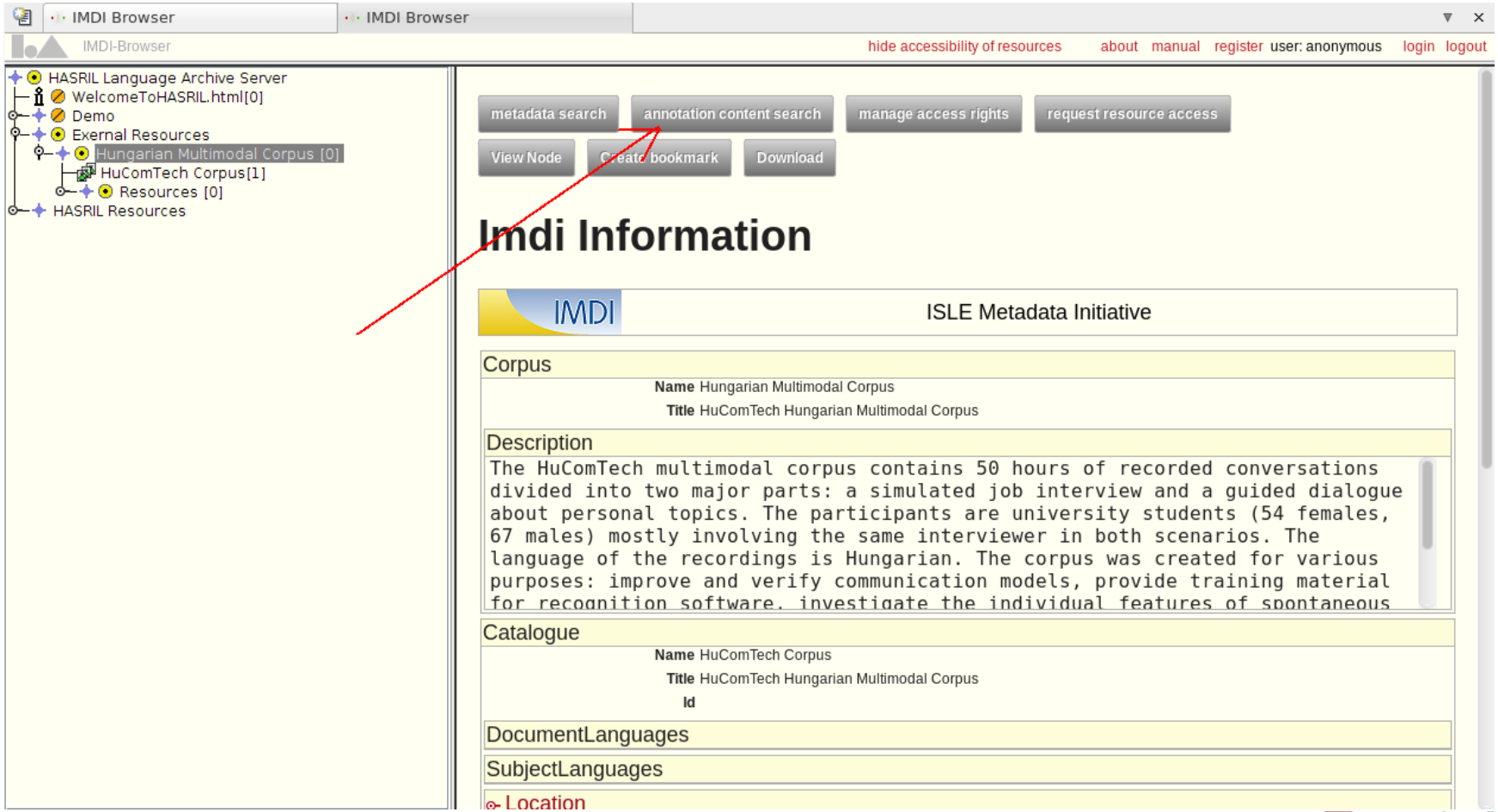
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

Keresés

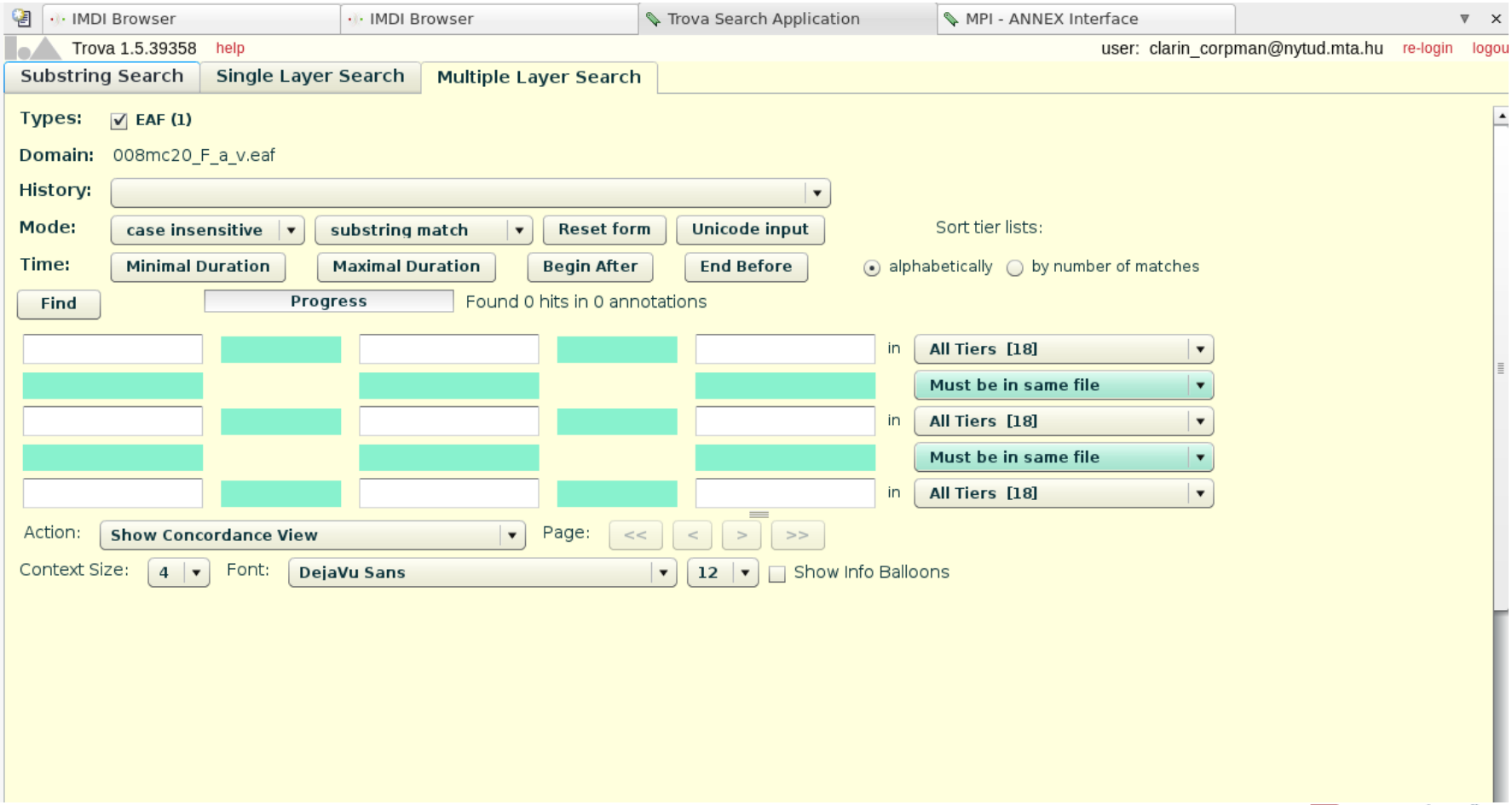
NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



A MAGYAR TUDOMÁNY ÜNNEPE

Trova

NYELVTUDOMÁNYI INTÉZET
MAGYAR TUDOMÁNYOS AKADÉMIA



Elérhetőség

http://clarin.nytud.hu

Köszönöm a figyelmet!