



Digitális bölcsészet

Digital humanities

Kanász-Nagy Zoltán (zkanasznagy@monguz.hu)

Tóth Kornél (ktoth@monguz.hu)

Fogalmi meghatározás

- "A **digitális bölcsészet** olyan tudományos diszciplína, amely számítógépes technológiákat használ, fejleszt bölcsészeti, társadalomtudományi kutatások során. Így hagyományos tudományos kérdésekre keres többnyire matematizálható válaszokat, valamint az új módszerek révén új kérdésfeltevésekkel gazdagítja a bölcsészet- és társadalomtudományokat. A digitális bölcsészet ennél fogva transzdiszciplináris kutatási terület, ahol a számítógép számolási kapacitását irodalomtudományi, kulturális emlékezeti, médiatudományi, történettudományi és nyelvészeti kutatási programokban használja fel ..." (wikipedia)

Történeti áttekintés

- A számítógép és bölcsészettudományok története az 1940-es évekre megy vissza, különös tekintettel [Roberto Busa](#) jezsuita szerzetes tevékenységére, aki az IBM-mel együttműködve elkészítette az *Index Thomisticus*t, azaz egy számítógép generálta konkordanciát [Aquinói Szent Tamás](#) műveihez.
- Az ő tevékenységét követve kezdtek el szövegeket digitalizálni, majd ezekben a szavak keresését, számolását, statisztikai módszereket alkalmazni.
- (wikipedia)

Történeti áttekintés

- A digitális szövegkezelés következő nagy etapa az volt, amikor az xml fájlok adta lehetőségeket kezdték kiaknázni, azaz az xml címkézési struktúrát alkalmazták szöveges dokumentumok metaadatokkal való ellátására. (wikipedia)

Történeti áttekintés

- Ez a tevékenység irányította a figyelmet arra, hogy érdemes szabványokat kialakítani a metaadatolásra, és így jöttek létre metaadatszabványok, amelyek közül a leghíresebbé és legsikeresebbé a **TEI** (Text Encoding Initiative) kódolás vált a bölcsészettudományok körében, hiszen a „*TEI-XML segítségével nemcsak a látható szerkezeti struktúrákat jelölhetjük (fejezet, címsor, bekezdés), hanem a szöveghez fűződő elemzéseket és értelmezéseket is*” (Labádi Gergely: *"Digitális filológia"* -- wikipedia)

Történeti áttekintés

- A digitális bölcsészet (Digital Humanities) megszületésével egy másik folyamat is elindult, mégpedig a nyelvészeti és kulturális beállítódás különválása. A digitális bölcsészet elsősorban kulturális produktumokkal foglalkozik, míg a számítógépes nyelvészet a nyelvi elemzésekkel, korpusznyelvészettel, és ezek gyakorlati alkalmazásával, azaz gépi fordítással, szöveg hanggá, hang szöveggé alakításával, szövegbányászattal. (wikipédia)

Szemantikus web

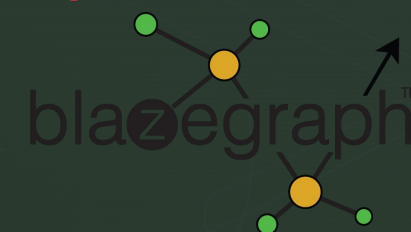
- *A szemantikus web egy közös keretet biztosít az adatok határok nélküli megosztására és újrahasznosítására az alkalmazások, intézmények és közösségek között ...”* A szemantikus webet mint fogalmat először Tim Berners-Lee írta le 2001-ben.
- A szemantikus web adatelemeket és nem weblapokat kapcsol össze. Ez az adatkapcsolt modell újszerű összefüggéseket tár fel az emberek, az üzleti világ, különböző közösségek számára, megkönnyítve a tudás megszerzését és terjesztését.
- A szemantikus web eszközei (**URI, RDF, SPARQL, ontológiák**) jól használhatók a digitális bölcsészet területén.
- Az általunk kínált megoldás egy olyan repozitóriumi szoftvert biztosít, amely fel van készítve a szemantikus adattárolásra / keresésre, adatok továbbítására.

Islandora CLAW

- Az **Islandora CLAW** az Islandora szoftver **Drupal 8** és **Fedora 4** komponensekre épülő változata.
- Az Islandora CLAW egy open source megoldás, amelyet az Islandora közössége fejleszt.
- Szoftvercsomag, legfontosabb komponensei:
 - Drupal 8 CMS rendszer
 - alaprendszer + dedikált modulok
 - Php nyelven megírt mikroszervíz réteg (Crayfish, Crayfish-Commons)
 - szemantikus szinkron, derivatívák generálása
 - Fedora 4 (fcrepo)
 - Blazegraph (bigdata)



Drupal™



Islandora CLAW jellemzői

- egyedi űrlapok készítése (rdf mappelés)
- többnyelvűség (felhasználói felület, rekordok leírása)
- rekordok verziókezelése
- rekordok közötti kapcsolatok profi módon történő kezelése
- Solr integráció (fulltext keresés, facettázás, highlight)
- felhasználó kezelés
- témakezelés
- digitális fájlok kezelése (szöveges, audióvizuális)
- IIF képszerver
- mobilra optimalizált felület

Gyűjteménykezelés

- Az Islandora CLAW gyűjtemény-kezelése követi az alap Islandora verzió hierarchikus szerkezetét. Ezáltal benne tetszőleges mélységű gyűjtemény-hierarchia alakítható ki. Az egyes gyűjtemények más algyűjteményeket vagy pedig fájlokat, szöveges tartalmakat foghatnak egybe. Bármelyik elem ebben a gyűjteményi rendszerben (gyűjtemény, fájl, szöveges tartalom) valamennyi leíró adata (cím, leírás, stb.) több különböző nyelven megadható.

Referenciák -- kapcsolatok

- Lehetőség van egy- és kétirányú referenciák használatára. Sok esetben elég az egyirányú hivatkozás, viszont előfordulnak olyan tartalomtípusok, ahol a hivatkozást szükséges látni mindkét típusú tartalom felől. Pl.: folyóirat számok és cikkek között vagy folyóirat rekord és a kapcsolódó számok között.
- Bármelyik oldal felől történik egy mentés (szám - cikk) a túloldali referencia automatikusan elkészül.

Szemantikus adatkezelés

- Az Islandora CLAW célja, hogy a Drupal tartalomkezelő szoftveren (CMS) kívül egy prezervációs réteget is adjon, melynek révén a Drupalban létrehozott tartalom automatikusan bekerüljön olyan komponensekbe is, amelyek a szemantikus adattárolásra lettek felépítve.
- A Drupal összes tartalma a megfelelő RDF mappings beállításával átszinkronizálható a Fedora és a Blazegraph triplestore adatbázisába.
- Blazegraph (bigdata) összes tartalma SPARQL lekérdezés segítségével lekereshetővé válik.

RDF leképezés (mapping)

- Ahhoz, hogy a Drupal tartalmak (node-ok) átkerülhessenek a Fedora-ba illetve a Blazegraph triplestore-ba meg kell feleltetni, **le kell képezni az egyes tartalomtípusok mezőit a megfelelő ontológiák elemeinek.**
- A Drupal ehhez biztosít webes felületen is használható eszközt az **RDF UI** komponens személyével.
 - csak a schema.org szótárat tartalmazza
 - egyedi illetve más ontológiák használata esetén szöveges konfigurációra van szükség (yaml)

Drupal - Fedora szinkron

Edit Repository Item Collection ☆

[View](#) [Edit](#) [Members](#) [Media](#) [Delete](#) [Devel](#)

[Home](#) » [Collection](#)

▼ GENERAL *

Title *

Alternative Title

[Show row weights](#)

IDENTIFIER

[Add another item](#)

Resource Type

Description

Collection@en

[Home](#) / [ec](#) / [ae](#) / [3f](#) / [d8](#) / ecae3fd8-a3cf-4f68-9e73-535b4166ba7c

Created at

2018-12-05T03:31:56.653Z by bypassAdmin

Last Modified at

2018-12-05T03:31:56.653Z by bypassAdmin

Children 0

Properties

dcterms: **extent**
1 item

dcterms: **title**
Collection

fedora: **created**
2018-12-05T03:31:56.653Z

fedora: **createdBy**
bypassAdmin

fedora: **hasParent**
<http://192.168.6.167:8080/fcrepo/rest/>

fedora: **lastModified**
2018-12-05T03:31:56.653Z

fedora: **lastModifiedBy**
bypassAdmin

fedora: **writable**
true

rdf: **type**
<http://fedora.info/definitions/v4/repository#Container>
<http://fedora.info/definitions/v4/repository#Resource>
<http://pcdm.org/models#Object>
<http://www.w3.org/ns/ldp#Container>
<http://www.w3.org/ns/ldp#RDFSsource>

schema: **author**
http://192.168.6.167:8000/user/1?_format=jsonld

schema: **dateCreated**
2018-12-05T03:31:42+00:00

schema: **dateModified**
2018-12-05T03:31:56+00:00

schema: **sameAs**
http://192.168.6.167:8000/node/2?_format=jsonld

Projektek

- **Magyar Filozófiai Tudástár** -- ELTE DH
- Előadás a Videotoriumon:
<https://kifu.videotorium.hu/hu/recordings/21153/a-magyar-filozofiai-tudastar-bemutatasa>
- A repozitórium tartalma:
 - 250 szöveges dokumentum, könyv, könyvrészlet, cikk
 - 150 hangfelvétel (sok egészen új publikálás)
 - 5 videofelvétel

Metaadatok és tárgyszavazás

- ELTE DH központ: 100 tanszék, nagyon eltérő digitális tartalmak, többnyelvű tárgyszavazás kell;
- Metaadat- és tárgyszórendszer: **The Cambridge Dictionary of Philosophy** ~ 3500 szavas; tárgyszórendszerként alkalmazva --> ontológiába szervezhető; ez lesz a keresés alapja.
- Metaadat-séma: **Schema.org**
- ELTE-DH URL: **<http://repo.elte-dh.hu/>**

Linkek

- Islandora CLAW:
 - <https://islandora.ca/CLAW>
 - <https://github.com/Islandora-CLAW/CLAW>
- Fedora
 - <https://wiki.duraspace.org/display/FF/2018-11-26+SWIB18+-+Fedora+Workshop>
- Drupal
 - <https://www.drupal.org/>
 - <https://www.drupal.org/documentation>

▶

Köszönjük a figyelmet!

