

TÉMATÉRKÉPEK A FELSŐOKTATÁSBAN

TOPIC MAPS IN HIGHER EDUCATION

Tóth Máté

Országos Széchényi Könyvtár – Könyvtári Intézet

Összefoglaló

Az előadás bemutatja, hogy a norvégiai Høgskolen i Oslo könyvtár szakán miként oktatják a tématerkép technológiát. A tématerkép technológia a tartalomkezelésnek egy olyan formája, amely témák közötti kapcsolatokat (asszociációk) definiál és azokhoz külső erőforrásokat, feljegyzéseket (külső és belső előfordulások) rendel. Az ismertetett gyakorlatban egy profitorientált vállalkozás segíti a tématerképek oktatását. Rendelkezésre bocsátják az oktatáshoz szükséges szoftvereket, az Omnigator és az Ontopolyt. Az Omnigator egy tématerkép vizualizáló, az Ontopoly pedig egy szerkesztő program, amellyel az ontológiavezérelt tématerkép-szerkesztés sajátítható el. A diákok két tématerkép szintakszist sajátítanak el (LTM, XTM). Minden tématerkép alkalmazás egy ontológiára épül, a témák egyértelmű címezhetőségét URI-k biztosítják, amelyek között kapcsolatokat írunk le. A tématerképek oktatásával így a szemantikus web működésének alapelvét és az ontológiaszerkesztés alapjait is elsajátítják a hallgatók. A norvég gyakorlat ismertetése példát ad arra is, hogy miként segítheti egy profitorientált vállalkozás az egyetemi képzés céljainak megvalósítását, illetve, hogy miként működhet közre az oktatás egy cég sikerében.

Kulcsszavak

felsőoktatás, tématerkép, ontológia, szemantikus web

Abstract

The paper presents how topic map technology is taught in Oslo University College (Norway) at the Department of Library, Information Science and Journalism. Topic map technology is based on three elements: topics, associations, occurrences. *Associations* are defined between *topics*, while the related information are stored in external and internal *occurrences*. A profit-oriented company supports topic map education in the presented practice. It makes all needed softwares available (Omnigator, Ontopoly). Omnigator is a topic map browser. Ontopoly is a ontology-driven topic map editor. Students learn two topic map syntax (LTM, XTM). As technology is very similar to the semantic web (topics are defined by URIs, ontology contains knowledge, triplets) students understand the main features of the next generation intelligent applications and the basics of ontology editing. The Norwegian practice gives a good example on how profit-oriented companies can support higher education, and how it can contribute in the success of a university.

Keywords

Higher education, topic maps, ontology, semantic web

1. Bevezetés

A tématerkép (topic map) a szemantikus web vízió mögött meghúzódó logika gyakorlatban is széleskörűen alkalmazott technológiája, a szemantikus web jelene. A felsőoktatásban ennek ellenére ez egy rendkívül kis mértékben reprezentált terület. Az előadás a tématerképek oktatásával kapcsolatosan szerzett norvégiai gyakorlati tapasztalatokat közvetíti a hazai felsőoktatás számára, nem titkoltan azzal a szándékkal, hogy kedvet ébresszen a technológia oktatására a hazai intézményekben.

2007 tavaszán az Oslói Főiskolán (Høgskolen i Oslo) három hónapos kutatói ösztöndíj során ismertem meg az ismertetésre kerülő gyakorlatot. A kutatást norvég-magyar államközi megállapodás alapján ösztöndíj formájában a norvég Kutatási Tanács (Forskningsrådet) finanszírozta.

2. A tématerkép technológia

A tématerkép technológia a tartalomkezelésnek egy olyan formája, amely témák közötti kapcsolatokat (asszociációk) definiál és azokhoz külső erőforrásokat, feljegyzéseket (külső és belső előfordulások) rendel. A tématerkép technológiával egy-egy téma környezetét, előfordulásait és asszociációs mezejét lehet feltárni. A szemantikus webhez hasonlóan a technológia lehetőséget teremt arra, hogy a számítógéppel egyszerű következtetéseket hajtassunk végre.

A tématerképek a következő elemekből állnak:

- Téma (topic) – A tématerképekben tárolt információ alapvető egységei a témák. Minden olyan dolgot témának kell tekintenünk, amelyről állítást kívánunk tenni. A témát meg kell nevezni, és biztosítani számára az egyértelmű címezhetőséget. A szemantikus web vízióban az egyértelmű címezhetőséget az URI-k (Universal Resource Identifier-ek) biztosítják, ahogy lényegében a tématerkép technológia esetében is. Minden témához rendelnünk kell egy olyan (külső) erőforrást, amely egyértelműen azonosítja a témát, és minden más témával szemben összetéveszthetlenné teszi. A norvégiai gyakorlatban használt alkalmazások PSI-ket (Public Subject Identifier) használnak a témák azonosítására, de Wikipédia oldalak, intézményi weboldalak, személyek esetében személyes weboldalak vagy email címek is alkalmasak az egyértelmű címezhetőség biztosítására. (Pl. Topic Map http://en.wikipedia.org/wiki/Topic_Map; Tóth Máté thmate@oszk.hu)
- Asszociáció (association) – A témák közötti relációkat az asszociációk írják le. Az asszociációk segítségével állításokat fogalmazhatunk meg két egymással összefüggő témáról. Például a fenti példában megnevezett két témát formálisan a kutatási területe reláció kapcsolja össze: Tóth Máté kutatási területe a Topic Map. Az asszociációk esetében is szükséges lehet az egyértelmű címezhetőség biztosítására, így ezekhez is rendelhetünk azonosítókat.
- Előfordulás (occurrence) – A tématerkép technológia harmadik meghatározó elemei az előfordulások. A tématerképekben az előfordulások azok az elemek, amelyek tényleges információt hordoznak, és nem egy adott dolgot reprezentálnak, mint az asszociációk, vagy a témák. Megkülönböztetünk külső és belső előfordulásokat. Előbbiek a tématerképen kívüli erőforrásokra mutató linkek, utóbbiak a tématerképbe írt információk, magyarázatok, adatok, tények.
- Névtípusok (name types) – Minden téma állhat különböző neveken. Minden témának van egy kitüntetett neve, amely szerepel az asszociációkban és amely az adott fogalom

reprezentánsaként funkcionál. Ezen túl azonban bármennyi névtípust meghatározhatunk. Például beírhatjuk mindennek az angol nevét (tématerkép – topic map); de személyek becenevét is (Kovács István – Kokó), vagy intézményi rövidítéseket (Központi Statisztikai Hivatal – KSH).

- Szereptípusok (role types) – A témák közötti kapcsolatok definiálása során gyakran nagy jelentősége van annak, hogy egy téma milyen szerepben vesz részt egy relációban. Egy személy például kutatóként rendelkezik kutatási területtel, munkavállalóként vesz részt egy munkáltatói viszonyban, és például apaként vesz részt egy szülő-gyermek viszonyban. Ezen szerepek definiálására használják a szerepeket, szereptípusokat.
- Érvényesség, hatókör (scope) – A tématerképekben szükség lehet bizonyos állítások irányának, bizonyos névtípusok érvényességi körének meghatározására. Erre szolgál az érvényesség definiálása. Nem mindegy például egy állításban, hogy ”Tóth Máté kutatási területe a Topic Map”, vagy fordítva, a ”Topic Map kutatási területe Tóth Máté”. Ebben az esetben az irányt határozhatjuk meg scope-ként. Az sem mindegy, hogy a ”topic map” és a ”tématerkép” kifejezések közül melyik az, amelyik angol és melyik, amely magyar nyelvi kontextusban értelmezhető. (Pepper, 2002)

3. Tématerképek és a szemantikus web

A tématerkép technológia ismertetésénél látható, hogy a szemantikus webhez hasonló logika szerint működnek a tématerképek. Az alkalmazások elemi szintjét képezik az egyértelműen (főként URI-kkal) azonosítható témák, amelyekről állításokat teszünk (lényegében az RDF logikája mentén). Mivel egy téma több állításban is részt vesz, a géppel egyszerű következtetéseket is végezhetünk. Az RDF nyelv éppen ezért alkalmas is tématerképek készítésére.

A tématerképekben azonban van egy nagyon fontos szint, az ontológiák szintje, amely az alkalmazás elkészítéséhez és működéséhez szükséges háttértudást tartalmazza. Az ontológiai szinten határozza meg a tématerkép készítője azt, hogy mely témátípus, mely témátípussal lehet relációban. Ezeket az egész alkalmazásra nézve általánosan érvényes tudástartalmakat tároljuk ontológiákban. Például egy irodalomról szóló tématerképben meghatározhatjuk, hogy ”a regény irodalmi mű”. ”Egy irodalmi műnek van egy vagy több szerzője”. ”Minden szerzőnek van egy születési helye” stb. Ezen ismeretek számunkra egyértelmű axiómák, amelyeket a géppel is meg kell ismertetnünk, számára formálisan le kell írunk.

Ahogy a szemantikus webhez való kapcsolódás egyik legfontosabb pontja az ontológiák megléte, a tématerkép és a szemantikus világháló közötti különbségek is leginkább ezen a szinten ragadhatók meg. A tématerképek esetében az ontológia egy kifejezetten szűk témakör meghatározott célra elkészített tudásszervezési rendszere. A tématerképben előforduló téma-, asszociáció-, előfordulás-, szerep- és névtípusok összessége (Garshol, 2007), amelyeket keresésre optimalizáltak, emberi visszakeresésre szolgálnak és amelyek nem formális logikán alapulnak. A szemantikus web ontológiáit általánosabb érvénnyel hozzák létre, sokkal szigorúbb struktúrákat kezelnek. Ezeket gépi visszakeresésre szánták és formális logikán alapulnak.

A különbségek ellenére a tématerképek alkalmasak a szemantikus web működési mechanizmusának gyakorlati demonstrálására. Tématerkép-szerkesztő programok segítségével a diákok maguk is készíthetnek ilyen alkalmazásokat, és közben megérthetik az intelligens alkalmazások mögött húzódó logikát. A következőkben a tématerképek oktatásának egy lehetséges módját mutatom be norvégiai tapasztalatok alapján.

4. Tématérképek oktatása Norvégiában

Norvégiában a tématerkép-alkalmazásokat sokkal szélesebb körben használják, mint hazánkban. Gyakoribb, hogy egy-egy weboldal, online napilap felülete mögött egy tématerkép-alkalmazás húzódik és teszi könnyebbé az adott oldalon való navigációt. Éppen ezért az Oslói Főiskola Könyvtár, Média és Újságírás karán is kiemelt fontosságot kap ez a technológia. 2007 őszén az Oslói Főiskola nemzetközi mesterképzést indított a digitális könyvtári ismeretekről (Tóth, 2008). A 2007 őszi félévben két tanegységet indítottak, egyenként 15-15 kreditpontért. Ezek egyike volt a digitális dokumentumok, amelynek keretében volt szó a tématerkép technológiáról.

A kurzusnak nem egészen a felét teszi ki ez a téma. A kurzus másik felében a digitális könyvtárak fogalma, a webes műfajok, a felhasználó által generált web (web 2.0), tudásszervezési rendszerek, a szemantikus web, metaadat sémák, és a digitalizálás jogi aspektusai jelennek meg (http://www.jbi.hio.no/bibin/digdok/dig_doc_course_plan_rev.htm).

A Digitális dokumentumok kurzus feltételezi az XML alapszintű ismeretét. Hogy ez mindenkinek meglegyen, valamennyi hallgatónak részt kell venni egy előzetes nyári kurzuson, amelyen többek között erről is szó van.

A tématerképekről összesen 15 órában (5 alkalommal háromszor 45 percen keresztül) van szó a következő bontásban:

- 1) Bevezetés (alapfogalmak, az Omnigator szoftver bemutatása, tématerkép alkalmazások. Házi feladat: letölteni az Ontopia oldaláról a szabadon elérhető tématerkép alkalmazásokat, az „OKS Samplers”-t);
- 2) Tématérképek készítése (LTM és XTM nyelvek szintaxisa, az Omnigator működése, házi feladat: LTM nyelven elkészíteni egy tématerképet);
- 3) Tématérképek készítése (további fogalmak, hierarchiák, az Ontopoly bemutatása)
- 4) A tématerképek mechanizmusa (tématérkép szabványok, Tématérkép adatmodell, tématerképek lekérdezése);
- 5) Tématérképek és ontológiák.

A tématerképekről szóló órák nem egymás után következnek, hanem azok be vannak építve a web 2.0-áról, a szemantikus webről és egyéb digitális dokumentumokkal kapcsolatos ismeretek közé. A diákok számára így az adatmodellekről, metaadatokról, ontológiákról szóló órákon tanultak gyakorlatban is összekapcsolódnak a tématerképekkel. A félév során két tématerképet kell a diákoknak elkészíteni, egyet önállóan LTM nyelven, egyet pedig projekt munka keretében 3-5 fős csoportokban XTM nyelven, egy ontológiavezérelt tématerkép-szerkesztő program, az Ontopoly segítségével.

Az LTM (Linear Topic Map Notation) egy kifejezetten az oktatás céljaira kifejlesztett tématerkép-szintakszis. (Garshol, 2006) Előnye az egyszerűség, és a világosság. Az LTM szemben az XTM-mel „szükszavúbb”. Az első LTM nyelven elkészített tématerkép formai követelménye, hogy abban leglább 3-3 téma-, asszociáció- és előfordulástípus legyen, és összesen legalább 50 téma, asszociáció és előfordulás legyen benne. Erre a diákok 2 hetet kapnak.

A tématerképek megjelenítésére az általam megismert gyakorlatban az Omnigator vizualizáló programot használják. Az Omnigator „mindenevő” (omnivorous), így LTM, XTM és RDF nyelveken elkészített tématerképeket is képes vizualizálni, illetve ezeket át tudja konvertálni egymásba. Az Omnigatort szintén oktatási céllal fejlesztették ki: a felület minden háttérinformációt megjelenít. Egy plug-in, a Vizigator segítségével a témák, és a közöttük lévő kapcsolatok vizuálisan is megjeleníthetők. A diákok a saját LTM tématerképeiket is ezzel a vizualizáló programmal jeleníthetik meg.

Egy LTM tématerkép elkészítéséhez nem szükséges, hogy a diákok tisztában legyenek az ontológiaszerkesztéssel. Ennek elsajátítására csak ezen első házi feladat kiértékelése után kerül sor. A második házi feladat egy XTM tématerkép elkészítése az Ontopoly ontológiavezérelt tématerkép-szerkesztő szoftverrel. A szerkesztés során először egy ontológiát kell elkészíteni, azaz definiálni a tématerképben előforduló téma-, asszociáció-, előfordulás-, szerep- és névtípusokat, az azok közötti lehetséges relációkat és azok számossági jellemzőit. Ezt követően kell a tématerkép így elkészített fogalmi kereteit feltölteni adatokkal, konkrét ismeretekkel. A második házi feladatra a diákok több időt kapnak, csoportmunka keretében dolgoznak. Az elvárás itt már 10-10 téma, asszociáció és előfordulástípus definiálása.

Az XTM (XML Topic Maps) szintakszis a tématerképek ISO által is szabványosított XML alapú nyelve. (<http://www.topicmaps.org/xtm/1.0/index.html>) Ennek ismerete (írása, olvasása) elvárás ugyan a hallgatóktól, de a második házi feladat során az XTM fájl az Ontopoly használata közben a háttérben készül. Az Ontopoly ingyenes verzióját szintén az oktatás számára fejlesztették ki, ugyanakkor ennek a programnak létezik üzleti célra is hasznosítható verziója. Az Ontopolyval két részletben készül a tématerkép, először az ontológia, utána jön az adatok feltöltése. Az ontológiában meghatározott fogalmi kereteket a számítógép számon kéri a tématerkép készítőjén, azaz figyelmezteti, hogyha kötelező adatokat nem töltött fel. Az Ontopoly lényege, hogy a legördülő menük szintjére viszi az ontológiaszerkesztést, így a diákoknak nem a bonyolult szintaksziszra, hanem a fogalmi viszonyok és a relációk helyességére kell koncentrálniuk.

Mindkét házi feladat esetében a diákok teljesen szabadon választhatnak, hogy milyen ismereteket szeretnének feldolgozni tématerkép formájában. 2007 őszén Beethoven szimfóniákról, angol detektívregényekről, turisztikai látványosságokról, borokról, Donald kacsá barátairól, kutyafajtákról készültek tématerképek. Lényegében bármit fel lehet dolgozni ilyen formában. A diákok kifejezetten szeretik, hogy egy kedvelt témával foglalkozhatnak heteken keresztül.

The screenshot shows a web browser window with the title "[Omniator] Tokaj-Hegyalja - Microsoft Internet Explorer". The address bar shows the URL: http://www.ontopedia.net/omniator/models/topic_nontopoly.jsp?tm=Wine.xml&id=1036. The page content is titled "Tokaj-Hegyalja" and is categorized as "Region". It displays several sections of information:

- Untyped Names (1)**: Tokaj-Hegyalja
- Subject Identifiers (1)**: http://psi.ontopedia.net/Tokaj_Hegyalja_wine_region
- Hungarian name (1)**: Tokaj-Hegyalja
- Hierarchy (Located in)**: Hungary > Tokaj-Hegyalja > ...
- Associations (16)**:
 - Contains: Bodrogkeresztúr, Bodrogkisfalud, Mezőzombor, Mád, Rátka, Szegi, Tarcál, Tokaj, Tállya
 - Located in: Hungary
 - Made here: Tokaji Aszú, Tokaji Eszencia, Tokaji Furmint
- Internal Occurrences (1)**:
 - Description: A historic wine region in Northern Hungary.
- External Occurrences (4)**:
 - Official site: <http://www.tokajhegyalja.hu>
 - Picture: http://en.wikipedia.org/wiki/Image:Tokaj_-_Hegyalja-06.jpg
 - Webpage: <http://en.wikipedia.org/wiki/Tokaj-Hegyalja>
 - Website: <http://tokajbor.lap.hu>

1. ábra Tokaj-Hegyaljával kapcsolatos információk a „Borok” tématerképben az Omniator felületén

5. A vállalati szektor szerepe az oktatásban

A tématerképek oktatását a vállalati szektor aktív közreműködése teszi sikeressé. A norvégiai gyakorlatban egy oslói székhelyű, tématerképek üzleti alkalmazásaival és oktatásával foglalkozó, cég működött közre az oktatásban. A cég szakembere tartotta az órákat. A félév során használt ingyenes szoftvereket is ők fejlesztették ki, így minden tekintetben első kézből származó információkhoz juthattak a diákok. A félév alatt szervezett fizetős képzéseken a szabad helyeket a főiskolai hallgatókkal töltötték fel, és a – szintén általuk szponzorált – ”Tématerkép használói kör” előadásaira is meghívást kaptak a diákok. A kör tagjaival való ismeretséget a cég által szponzorált vacsorán mélyíthették el.

Az oslói tapasztalatok alapján az együttműködés mindkét fél számára gyümölcsöző lehet. A felsőoktatási intézmény számára előnyt a szoftverek, a közreműködő szakember jelentenek. A diákok gyakorlatorientált képzést kapnak, megismerkedhetnek későbbi potenciális munkáltatóikkal. A képzésben szerepet vállaló cég tématerképekre való igényt generál, ezzel potenciális megrendelőket kap; a saját maga által kifejlesztett, vagy használt szoftverekre kiképzett munkaerőt termel, akiből válogathat a későbbiekben; házi feladatokként elkészített tématerképekkel konkrét alkalmazásokat kap, amelyekkel be tudja mutatni a technológiát sokféle kontextusban.

6. Lehetőségek, kihívások hazai környezetben

A tématerképek Magyarországon mind a felsőoktatásban, mind pedig az üzleti alkalmazások terén kiaknázatlan lehetőségeket rejt. Tekintettel a Norvégiában használt szoftverek ingyenességére a gyakorlat egy az egyben is alkalmazható hazai környezetben. Ezek nyelve angol, ami bizonyos nehézségeket okozhat. A TM4L tématerkép-szerkesztő szoftver magyar nyelven is hozzáférhető, így az infrastrukturális feltételei egy anyanyelven való programnak is megvannak. (Kovács, Tóth, 2008) Lehetőséget jelent, hogy vannak már hazai vállalatok és hazai szakemberek is, akik érdeklődnek a tématerképek iránt, így a humán erőforrás és a vállalati szektorral való együttműködésnek a lehetősége is megvan. Egyelőre a felsőoktatási intézmények részéről hiányzik az érdeklődés.

A tématerkép-szerkesztés oktatásával elsősorban azoknak a felsőoktatási képzőhelyeknek kellene foglalkozniuk, ahol a szemantikus web, a tartalomkezelés, a tudásmenedzsment, illetve az ontológiák, fogalmi hierarchiák bármilyen alkalmazása felmerül.

Szintén szükség lenne arra, hogy a tématerképek iránt érdeklődő vállalatoknak sikerüljön erőteljesebb igényt generálni a technológia alkalmazására, amelyre a következő területeken lenne komoly lehetőség: oktatás, elektronikus könyvtári dokumentumkezelés, vállalati tudásmenedzsment, vállalati iratkezelés. Oktatásban való alkalmazásra jó példa az Országos Széchényi Könyvtár Hálózatos irodalom projektje. (<http://mekmester.oszk.hu:8181/itm/tmv/index.htm>) Ez egy középiskolások számára készült interaktív felület, amelyben a középiskolai irodalom tananyagot dolgozták fel egy szabadon böngészhető formában. Az oldal összefüggéseiben mutatja be a magyar irodalmat, így a diákok könnyebben megértik a tananyagot. Elektronikus könyvtári alkalmazásra a Neumann Ház katalógus projektje lehet példa (Ignéczi, 2004). Vállalati tudásmenedzsmentre és iratkezelésre javaslat szinten van egy tématerkép alapú rendszer. (Fülöp et al, 2007)

7. Összegzés, konklúzió

A tématerkép technológiát Magyarországon nagyon szűk körben alkalmazzák, pedig mind üzleti, mind pedig oktatási célra nagyon hatékonyan használható. A felsőoktatásban a szemantikus webbel, ontológiákkal, tudásmenedzsmenttel kapcsolatos ismeretek oktatását hatékonyan egészíti ki egy gyakorlati tématerképekről szóló képzés. A vállalati szektorban és a felsőoktatásban való intenzívebb alkalmazás fokozottabb igényt generálna a tématerképekre, közvetetten pedig az intelligens alkalmazásokra, amely a vállalati kultúra erősödését, a tudatosabb tudásgazdálkodást segítené.

Irodalomjegyzék

- [1] Fülöp Gyula, Kormos János, Kovács Zoltán Csaba, Lencse Zsolt (2007) The Role of Managerial Notes in Knowledge Sharing: Helping the Management of Unstructured Information by Using Topic Maps = Proceedings of the 8th European Conference on Knowledge Management 2007. Barcelona, Spain Sept. 6-7.
- [2] Garshol, Lars Marius (2006) The Linear Topic Map Notation, Definition and Introduction, version 1.3. Ontopia AS 2006./06/17 <http://www.ontopia.net/download/ltn.html>
- [3] Garshol, Lars Marius (2007) Towards a Methodology for Developing Topic Maps Ontologies = Leveraging the Semantics of Topic Maps. Berlin – Heidelberg, Springer, 2007. 20–31.
- [4] Ignéczi Lilla (2004) Tématerkép belső fejlesztéssel – Vizuális tezaurusz a magyar Internet katalógusában = Tudományos és műszaki tájékoztatás, 2004. (51. évf.) 7. sz. http://tmt.omikk.bme.hu/show_news.html?id=3649&issue_id=452
- [5] Kovács Zoltán Csaba, Tóth Máté (2008) A tématerkép technológia bemutatása, a TM4L szerkesztő magyarosításának folyamatán keresztül, MicroCAD 2008 International Scientific Conference, Miskolc, 20-21 March 2008 Miskolci Egyetem, Konferencia kiadvány Section O: Applied Information Engineering, 75-82.
- [6] Omnigator: The Topic Map Browser, User's Guide Ontopia, 2007/08/24 <http://www.ontopia.net/omnigator/docs/navigator/userguide.html>
- [7] Pepper, Steve (2002) The TAO of Topic Maps, Finding the Way in the Age of Infoglut, Ontopia AS, 2002. <http://www.ontopia.net/topicmaps/materials/tao.html>
- [8] Tóth Máté (2008) Nemzetközi mesterképzés a digitális könyvtárakról In: Könyv, könyvtár, könyvtáros, 2008. (17. évf.) 3. sz. 45-50. p.
- [9] XML Topic Maps (XTM) 1.0 <http://www.topicmaps.org/xtm/1.0/index.html>