

## WEBBÁNYÁSZATI MÓDSZEREK ALKALMAZÁSA INTERNET ALAPÚ TANANYAGOK MINŐSÉGVIZSGÁLATÁRA

Izsó Lajos

Budapesti Műszaki Főiskola, Tanárképző és Mérnökpedagógiai Központ  
1081 Budapest, Népszínház u. 8.  
izsolajos@erg.bme.hu

---

*Absztrakt: Az internet alapú tananyagok – bár ellentmondásosan és időnkénti visszaesésekkel – erőteljesen terjednek. Ez az erőteljes terjedés gyakran azzal jár, hogy az intenzív mennyiségi növekedéssel az anyagok (szakmai, didaktikai, használhatósági/szoftver-ergonómiai) minősége nem tart lépést. Növekvő igény tapasztalható ezért ezen tananyagok megfelelő minőségére. Ennek következményeként számos hagyományosnak tekinthető – különböző kérdőívekre, skálákra, interjúkra épülő – próbálkozást láthatunk Internet alapú tananyagok minőségének mérésére. Elismerjük ezen hagyományos megközelítések értékeit és szükségességét, úgy véljük azonban, hogy az internet alapú interaktív tananyagok minőségét ezeknél radikálisan jobban meg lehet ragadni alkalmas webbányászati technikákkal nyert adatok segítségével. A BME Alkalmazott Pedagógia és Pszichológia Intézete – a Budapesti Műszaki Főiskola Tanárképző és Mérnökpedagógiai Központja aktív részvételével – egy kutatócsoportot hozott létre azzal a céllal, hogy az ún. „felhasználó-középpontú” filozófiára épülő Clementine adatbányász programcsomag és annak WebMining CAT modulja segítségével – a tanuló és a tananyag interakciója során képződő naplófájlok kifinomult elemzése útján – kiválasztott konkrét internet alapú interaktív tananyagok minőségét az eddigieknél megalapozottabb megközelítésekkel, módszerekkel jellemezze. Ez az előadás és az azonos című kapcsolódó oktatási webbányász szekció ennek a kutató munkának az első eredményeit mutatja be.*

---

### 1. Bevezetés

Az elmúlt néhány évben az internet alapú tananyagok – a gyakran tapasztalható ellentmondások és átmeneti visszaesések ellenére – jelentős mértékben terjedtek. Az oktatási webhelyek gyakran csupán közvetítik a különböző formátumú (doc, ppt, pdf, stb.) tananyagok hozzáférését, letöltését. Újabban azonban már határozottan nő azon webhelyek száma is, amelyek valódi interaktív tananyag-feldolgozási lehetőségeket biztosítanak (pl. anyagfeldolgozás, példamegoldás, tantárgytesztek kitöltése interaktív visszajelzéssel, stb.). Az internet alapú tananyagok erőteljes terjedését kísérő gyakori negatív tapasztalat ugyanakkor az, hogy ezzel a mennyiségi növekedéssel az anyagok (szakmai, didaktikai, használhatósági/szoftver-ergonómiai) minősége nem tart lépést. Ennek egyenes következménye az a növekvő igény, hogy ezen tananyagok minőségét elméleti és gyakorlati szempontból egyaránt megalapozott módon tudjuk megítélni, jellemezni. Számos próbálkozást láthatunk ezért újabban ilyen tananyagok minőségének a – lehetőség szerinti objektív – mérésére (a legtöbb ilyen megközelítés különböző kérdőívekre, skálákra, interjúkra vagy strukturált megfigyelésekre épül). Mindezek mellett megállapítható, hogy a probléma még távolról sincs megoldva.

### 2. Webbányászat: az üzleti szféra internet alapú szegmensében kidolgozott igen erőteljes speciális adatbányászati megközelítés

Webbányászat: az üzleti szféra internet alapú szegmensében kidolgozott igen erőteljes speciális adatbányászati megközelítés, az adatbányászat alkalmazása a webhelyek forgalomelemzésének speciális területére. A módszer a webhelyek üzemeltetőjének a szerverén képződő, a forgalomra jellemző naplófájlok (az ún. logfájlok) kifinomult elemzésére épül és az adott webhelyek hatékonyságának a fokozását tűzi célul.

A különböző típusú naplófájlok (ún. weblogok) elvben a felhasználók tevékenységére vonatkozó igen gazdag információforrások. Itt a probléma a szokásosnak éppen a fordítottja: itt nem kevés, hanem túlságosan is sok információ áll rendelkezésünkre. A naplófájlok mérete ugyanis az üzleti szférában – a látogatók számától és elemzett időszak hosszától függően – gyakran a több száz MB vagy a GB nagyságrendjébe esik. Ezt a nagy mennyiségű – és nagy mértékben pusztán technikai jellegű – információt kell célszerű módon megszűrni, átalakítani, feldolgozni úgy, hogy a kezdeti technikai információból a felhasználó viselkedését és motivációját megalapozottan jellemző információkat nyerjünk.

A gyakorlati webbányászati célokra korábbi tapasztalataink alapján a *Clementine WebMining* eszközöket javasoljuk alkalmazni.

A webbányászatnak a *Clementine*-ban követett ún. „felhasználó-középpontú” filozófiája ugyanis teljes mértékben megfelel a korszerű marketing, ergonómia és pedagógia felfogásának. Ez az újszerű felfogás – szemben a hagyományos technikai ihletésű „lap-középpontú” megközelítéssel – a felhasználó céljait, szándékait helyezi az előtérbe és ennek alárendelve tervezi meg a rendszer szolgáltatásait, igen sikeresen.

Példaként következzen néhány hasznosnak bizonyult elemzési megközelítés, amelyeket kifinomult – gyakran különböző tanulóalgoritmusokat is tartalmazó – célra irányult szoftver modulok (a *Clementine* esetében ún. *streamek*) – támogatnak.

- A látogatások és a látogatók szegmentációja: *Visit and User Segmentation (E-ChannelUser RFM Classifications, User Mode Determination, Visit Branding)*.
- A látogatók viselkedésének elemzése: *Web Site Activity and User Behavior (Visit Activity Variances, Identifying Undesirable Behavior, Lifetime Conversion Tracking, Points of Abandonment, User Activity Focus, Visit Activity Funnels, Navigational Usage)*.
- A látogatók leggyakoribb aktivitási sorozatainak azonosítása: *Activity Sequence Analysis (Most Common Activity Sequences, Eventstream Visualisation)*.
- A látogatók vásárlási, döntési hajlamának meghatározása tanuló algoritmusok segítségével: *Propensity Analysis*.
- A látogatók magasabb szintű szegmentációja: *Advanced User Segmentation*.
- Online hirdetési tevékenység hatásvizsgálata: *Targeting Online Promotional Activity*, illetve *Campaign Performance Measurement*.

Látni fogjuk később, hogy ezen *streamek* egy része haszonnal adaptálható az internet alapú tananyagok világára.

### **3. Alkalmazzuk a webbányászatot internetes tananyagokra!**

Az oktatás világa maga nem lett volna képes a webbányászati eszközök kifejlesztéséhez szükséges erőforrások biztosítására. Miután azonban ezeket az eszközöket az üzleti szféra igényeinek a kielégítésére már kifejlesztették, adott annak a lehetősége, hogy az oktatás világára adaptáljuk. Mindkét világpiacevezető adatbányász programcsomag (az *SPSS Clementine* és a *SAS Enterprise Miner*) rendelkezik kiegészítő webbányász eszközökkel.

A BME Alkalmazott Pedagógia és Pszichológia Intézetében (APPI) – a Budapesti Műszaki Főiskola Tanárképző és Mérnökpedagógiai Központja aktív részvételével – létrejött egy kutatócsoport, amely a Clementine eszközökkel megvalósítható oktatási célú gyakorlati adatbányászati alkalmazások lehetőségeit vizsgálja. A kutatócsoportnak jelenleg kb. 14-16 aktív tagja van, akik zömmel a BME APPI és a BMF TMPK munkatársai mellett más pedagógusképző társintézmények (BDF, EKF, SZE) szakemberei, illetve diplomázó hallgatói közül kerülnek ki.

Az interakció professzionális vizsgálata lehetővé teszi a tananyag tényleges használati módjainak az azonosítását, amiből megalapozottan következtetni lehet a tananyag hatékonyságára, illetve a hatékonyság konkrét akadályaira.

A webbányászati technikák ezért a tananyag fejlesztésének a folyamatában – mintegy az ipari gyártásközi minőségellenőrzés analógiájára – az egzakt közbenső „mérések” lehetőségét nyújtják: a fejlesztés minden „mintavételezés” után attól függő irányokat vehet, hogy a tanulók a megelőző lépésben hogyan fogadták a tananyag számukra rendelkezésre bocsátott verzióját. A menet közben azonosított szakmai, didaktikai, használhatósági, szoftver-ergonómiai (leggyakrabban a navigációval vagy az információmegjelenítéssel kapcsolatos), stb. problémákra ilyen módon gyors áttekintési, javítási javaslat adható és az így kialakított verzió egy újabb ciklusban tesztelhető.

Ez a megközelítés annyiban radikálisan más a korábbiakhoz képest, hogy itt a tanulók és a tananyag vonatkozásában nem egyfajta mintavételezésről van szó, hanem valamennyi tanuló valamennyi interakcióját elemezhetjük a teljes tananyagban a billentyűleütések és egérgattintások szintjének megfelelő finomságú felbontásban. Az elemzés tehát nem a célpopulációból vett többé-kevésbé reprezentatív mintákra épül, hanem a célpopulációra nézve teljes körű.

A sikeres webbányászati elemzés feltétele, hogy a különböző modellekkel nyert eredményeket az elemző képes legyen a tanulói/felhasználói szintű élményekkel, illetve tapasztalatokkal összekapcsolni. Ez azt jelenti, hogy csak az lehet valóban sikeres, aki mind a webbányászati elveket és eszközöket, mind az elemzett konkrét tananyagot, mind pedig a tananyagot használó tanulókat eléggé mélyen ismeri.

Végül, a kétségtelenül jelentős előnyök felsorolása után szólni kell ennek a megközelítésnek két jelenleg még fennálló akadályáról. Az egyik az adatbányász, illetve webbányász programcsomagok – üzleti szférához szabott – igen magas ára. A másik pedig az a tény, hogy a rendszer használatának elsajátítása komoly szellemi erőfeszítéseket igényel. Hosszabb távon azonban mindkét akadály vonatkozásában optimisták vagyunk. Egyfelől kifejezetten oktatási – nem profitszerző – célokra általában lehetséges a webbányász termékek forgalmazóitól jelentős kedvezményeket, illetve engedményeket kapni, másfelől az informatikai és matematikai kultúra terjedésével – ami napjaink információs társadalmában ha lassan is, de kétségtelenül végbemegy – a webbányászati elvi alapjainak és gyakorlati technikáinak az elsajátítása az újabb tanár nemzedékek számára már egyre könnyebb lesz.