



# FerSML szurkolói avatárok a könyvtárban

BÁTFAI Norbert – BÁTFAI Mária Erika

## Bevezetés

A FerSML labdarúgás szimulációs platform szurkolói avatárja kicsi XML dokumentum, melyből automatikusan nyerhetünk ki olyan információkat, melyek összessége bibliográfiai tételek tekinthető. Felmerül a kérdés, hogy a bibliográfiai adatokból ki/mi építse fel a bibliográfiai rekordokat. Az egész rendszer megszervezésének szintjén pedig az az alapkérdés merül fel, mint ami a hivatkozáskézelő szoftverek egyik alapvető megkülönböztetője: centralizált vagy elosztott legyen? Leegyszerűsítve: a bibliográfiai tételek érdemi kezelése egy távoli, központosított gépen vagy a helyi gépen történik-e. Előbbi hivatkozáskézelő szoftverre példa a CiteULike vagy a RefWorks, utóbbira a JabRef vagy a Zotero [12]. Elképzelésünknek a könyv-

tári funkciójú rendszerekkel mutatott analógiája indokolja, hogy a szurkolói avatár adatbázis tervezésénél megvizsgáljuk a tételek kezelésének azon irányait, amelyeket a könyvtárak az elmúlt évtizedtől napjainkig kifejlesztettek és használnak a saját céljaikra. Megpróbálunk tovább is lépni azzal, hogy a szurkolói avatárokat összekapcsolt nyílt adatokként (linked open data) a szemantikus világháló részévé tesszük.

A szurkolói avatár adatbázis egy közösségi erőforrás-megosztás alapú kezdeményezés, ezért a választott forma különösen nagymértékben befolyásolhatja erőfeszítéseink sikerét. Konkrétan, hogy a labdarúgást szerető és értő szurkolók körében fel tud-e épülni valóban az általunk elképzelt szurkolói rendszer.

## Háttér: a labdarúgáshoz kapcsolódó munkák

A könyvtáros és az informatikus szempontjából a FerSML (Footballer and Football Simulation Markup Language, Labdarúgó és Labdarúgás Szimulációs Jelölőnyelv) egy Relax NG kompakt szintaxist [10] használó új XML nyelv, amivel leírjuk a futball csapatokat. A leírás célja, hogy ezek az állományok olyan információkat tároljanak, amikkel a FerSML platform további része, a szimulációs szoftver sikerrel tudja a valódi futballmérkőzéseket szimulálni. A mi terminológiánkban ezek a leírások, azaz a szóban forgó XML állományok az avatárok. Elképzelésünket egy korábbi publikációnkban [1] vezettük be. A szimulációk sikeressége alatt az értendő, hogy a szimulált és a valóságos mérkőzéseken a különböző megfigyelt jellemzők (például egy tornán a lőtt összes gólok száma) eloszlásainak egyezéséről szóló nullhipotézist nem kell elvetnünk [2, 3]. A vázolt FerSML platform (a kölcsönösen együtt fejlődő a jelölőnyelv és szimulációs szoftver) a GNU Általános Nyilvános Engedélye (GNU GPL v3) alatt fejlesztett nyílt forráskód alapú kezdeményezés [4, 5].

Tegyük fel, hogy a hazai csapat szurkolói avatárjai rendelkezésre állnak, de nincsenek meg a vendégcsapat avatárjai. Erre a problémára adnak egyfajta közösségi választ a szurkolói avatárok, amelyről egy korábbi publikációnkban írtunk [6]. A következő néhány link szurkolói avatár állományokra mutat be:

[http://www.inf.unideb.hu/~nbatfai/Debrecen\\_Basel/FerSML.Basel.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Debrecen_Basel/FerSML.Basel.avatar.xml)

[http://www.inf.unideb.hu/~nbatfai/Debrecen\\_Basel/FerSML.Debrecen.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Debrecen_Basel/FerSML.Debrecen.avatar.xml)

[http://www.inf.unideb.hu/~nbatfai/Basel\\_Debrecen\\_kezdo/FerSML.Debrecen.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Basel_Debrecen_kezdo/FerSML.Debrecen.avatar.xml)

[http://www.inf.unideb.hu/~nbatfai/Basel\\_Debrecen\\_kezdo2/FerSML.Debrecen.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Basel_Debrecen_kezdo2/FerSML.Debrecen.avatar.xml)

[http://www.inf.unideb.hu/~nbatfai/Basel\\_Debrecen\\_kezdo3/FerSML.Debrecen.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Basel_Debrecen_kezdo3/FerSML.Debrecen.avatar.xml)

[http://www.inf.unideb.hu/~nbatfai/Debrecen\\_Lovech/FerSML.Lovech.avatar.xml](http://www.inf.unideb.hu/~nbatfai/Debrecen_Lovech/FerSML.Lovech.avatar.xml)

A FerSML projekt társadalmi beágyazottságát és a közösségépítést a magyar érdeklődők körében [7, 8] a blogok karbantartásával tervezzük megoldani, ahol népszerűsítő formában ismertetjük kutatásaink legfrissebb eredményeit is. Jól mutatja ezt a [9] közleményben közölt eredmény és a [8] blog kapcsolódó posztjainak kapcsolata.

## Hasonló, a labdarúgáshoz nem kapcsolható munkák

A digitális könyvtárak és a szemantikus web kapcsolata fejlődő kutatási terület [20, 21, 22, 23]. A szurkolói avatárok kapcsán általunk felvetett irányok egyikével, a szemantikus webbel kapcsolatos erőfeszítéseinkkel teljesen rokon munka olvasható [19].

## Eredmények

A szurkolói avatár adatbázis megszervezésére több megoldási lehetőség kínálkozik, tekintsük ezek közül néhányat abból a szempontból, hogy elkülönül-e a szurkolói avatárok funkcionális (azaz szimulációbeli) és bibliográfiai feldolgozása. A felmerülő néhány tipikus esetet előbb nagy vonalakban, majd részleteiben és egymásra ható kölcsönhatásaikban vizsgáljuk meg a következő pontokban.

A. Gyümölcsöző lehet az avatárok bibliográfiai kezelését összekötni funkcionális, azaz a szimulációban történő felhasználással. Ez konkrétan azt jelenti, hogy magába a szimulációs programba építünk be valamilyen, az avatárból történő konverziót, aminek a kimenete lehet például egy COinS (ContextObject in SPAN) [11] alapú XHTML, vagy MODS XML, vagy akár egyszerűen BibTeX. Továbbá megkönnyíti helyzetünket, hogy az avatár állományok maguk is XML dokumentumok, így tetszőleges konverziót tudunk kis munkabefektetéssel, az XSLT transzformáció megadásával, akár még magából a szimulációs Java program (hu.javacska.xml csomagjából) is elvégezni. Miért gyümölcsöző ez a szem-

lélet? Mert az avatár-fejlesztő magának építi fel az adatbázisát, amit biztosan sikerrel tud felhasználni a további szurkolói avatárjai elkészítésében, illetve a eredményei publikálásában.

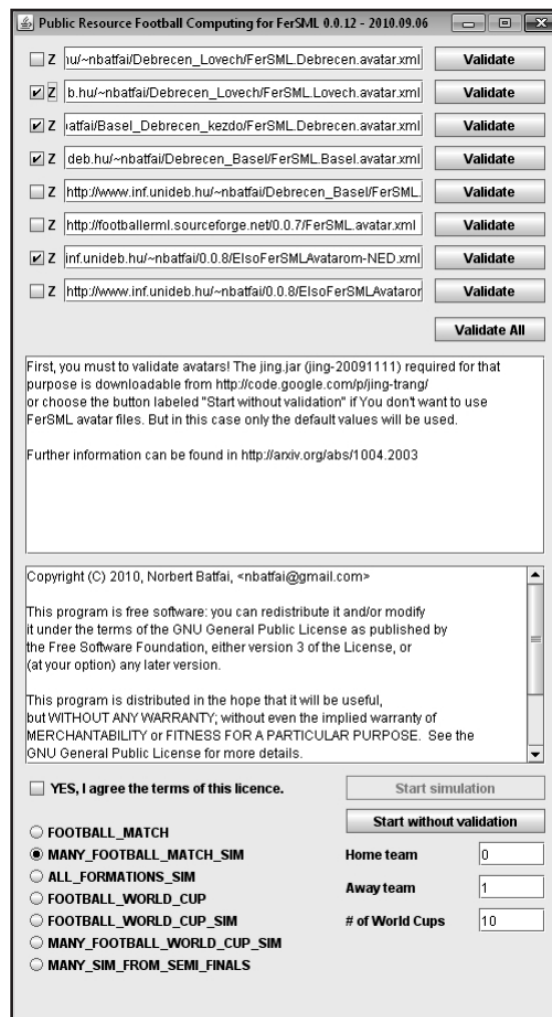
B. Ennek ellenkezője az az eset, amikor a szurkolói avatárok funkcionális és bibliográfiai feldolgozása elkülönül. Ez sokkal inkább egy centralizált feldolgozási irányt jelent.

C. Az is elképzelhető, hogy az avatár-fejlesztők munkájának támogatásához egy külön kliens programot fejlesztünk ki, amit az avatárok építésére és egyben bibliográfiai kezelésükre specializálunk. Ez az A.) esettel rokon, különbség, hogy a funkcionális használattól, azaz a szimulációtól elkülönül.

D. Továbbá az is felmerült, hogy a szurkolói avatárokat, mint nyílt adathalmazokat (open data set) a szemantikus világháló részévé tegyük. Ezt az A.) és a C.) esetekkel is kombinálhatjuk.

Az A.) eset egy tipikusan elosztott rendszert vetít előre, ahol az avatárokat készítő, kezelő avatár-fejlesztő, aki tipikusan nem rendelkezik könyvtáros ismeretekkel, saját bibliográfiáját vagy adatbázisát építheti fel. Ebben az esetben végeztünk néhány bevezető mérést: az 1. ábra mutatta szimulációs szoftver felületén a bemenő avatárok megadása mellett feltüntetettünk egy „Z” (mint Zotero) betűvel címkézett opcionális választási lehetőséget. Ezt kiválasztva a validálás után a program a szóban forgó avatárhoz elkészít egy egyszerű XHTML állományt, amibe befűzi az avatárból készített CoinS tagot, s ezt automatikusan betölti a böngészőbe, ahonnan a CoinS tagon keresztül az avatár bibliográfiai adatai a Zoteróba tölthetőek. Ez látható a 2. ábrán. A belinkelt, tipikusan távoli avatár állomány egy további kattintással a Zoteróba, azaz a helyi gépre is tölthető, ennek az újonnan létrejött bibliográfia bejegyzésnek a mellékleteként. Megjegyezhetjük, hogy a CoinS alapú megoldás helyett választhatnánk a talán még egyszerűbb beágyazott RDF alkalmazását is, így jár el például a Kempelen Farkas Felsőoktatási Digitális

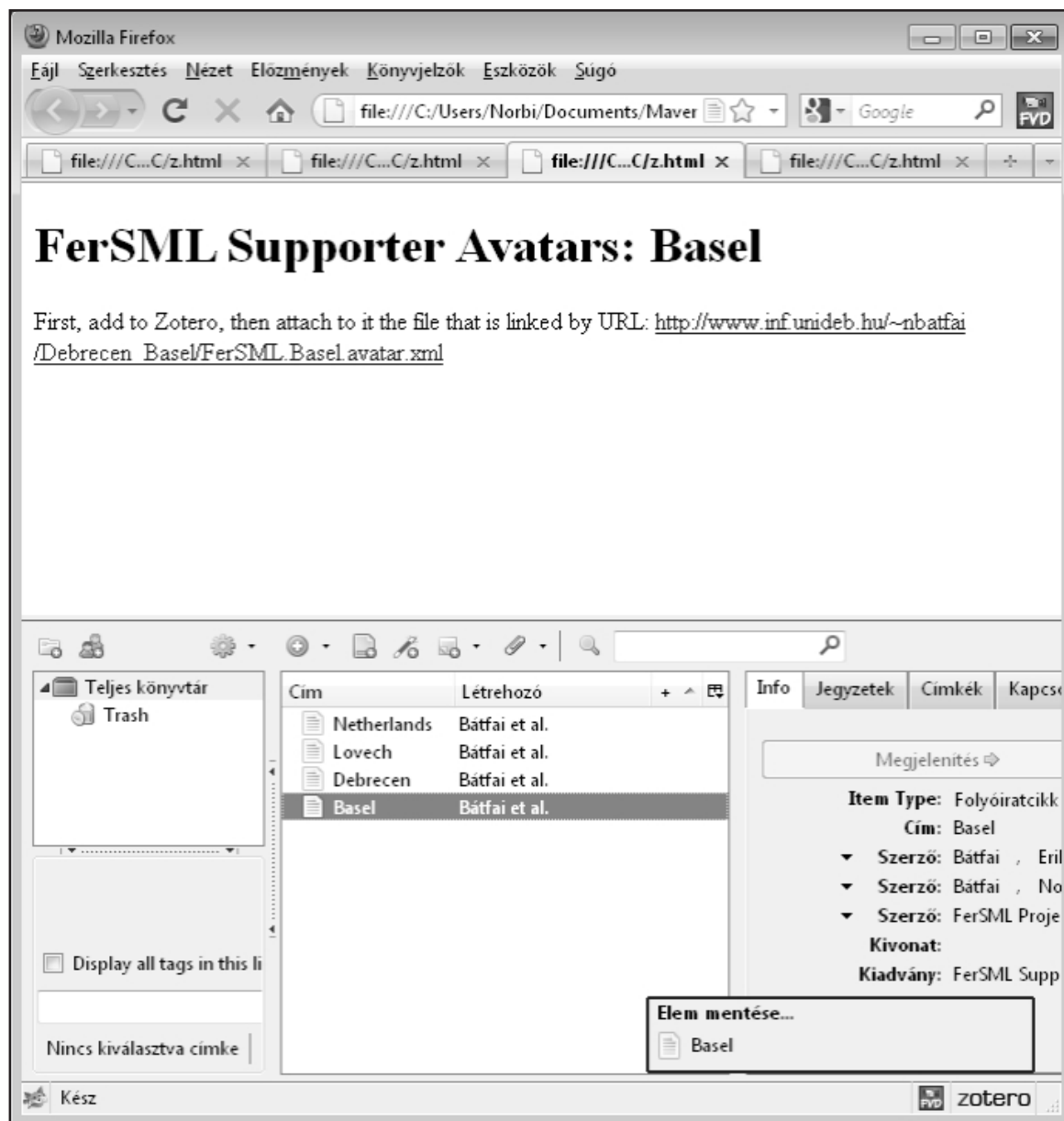
Tankönyvtár (<http://www.tankonyvtar.hu>) vagy éppen a YouTube videó-megosztó portál.



1. ábra

A Zoteróba tölteni kívánt szurkolói avatárok kiválasztása.

A B.) eset is nagy további mozgásteret enged. Vegyünk itt is egy izometrikus esetet! A Debreceni Egyetemen tipikusan keletkezni fognak szurkolói avatárok, hiszen egyrészt ez a fejlesztők munkahelye, másrészt hallgatókat vonunk be a szurkolói avatárok készítésébe, például nyári ösztöndíjprogram keretében. Adódik, hogy könyvtáros írja le ezeket a helyi ILS rendszerben, de az sem elképzelhetetlen, hogy bekerüljenek az egyetem DSpace alapú elektronikus archívumába.



2. ábra

*A Bajnokok Ligája selejtezőkörében a Lokit legyőző FC Basel szurkolói avatárjának (pontosabban bibliográfiai tételének, majd avatárjának) betöltése a Zoteróba.*

Érdeemes-e, kell-e ehhez egy integrált könyvtári rendszer? A válasz nem. Egyrészt ezért nem, mert a szurkolói avatárok kezelésekor egy ILS alapvetően nem használható ki, hiszen nincs kölcsönzés, gyarapítás, folyóiratok; alapvetően a humán leírás is felesleges emberi erőforráspazarlás, mivel maguk az avatárok tartalmazzák

a bibliográfia adatokat. Másrészt azért nem, mert az avatárokat be is kell gyűjteni, mielőtt a 856-os mezők halott linkeket tartalmaznának. E két ok miatt vagy a Zotero vagy egy DSpace alapú **elektronikus archívum** lenne a megfelelő választás. Viszont mivel a besorolási adatokat mindenképpen érdemes egységesíteni, a teljes

automatizálás sem üdvös, különben a visszakereshetőség komoly kihívássá válik.

A C.) eset izgalmas alternatíva. Egyrészt mert az átlagos szurkolótól nagyon távol állhat, hogy „kedvenc szövegszerkesztőjével” a Relax NG kompakt szintaxisával megadott nyelvtan szerint érvényes szurkolói avatárt (XML állományt) állít össze. Másrészt ez a célkliens egyben fájl-megosztó funkciókkal is ellátható (p2p jellegű megosztásra gondoljunk). Az előbbi felvetés konkrétan azt jelenti, hogy a szurkoló csak a megfigyeléseit adja meg és a kliens program készíti, validálja a szurkolói avatárokat. Utóbbi pedig, hogy az elkészített avatár automatikusan meg is osztható a közösséggel. Ez egy tipikusan elosztott rendszer lehet.

Kétségtől a D.) eset tartalmazza a legizgalmasabb felvetéseket. Hiszen adott a lehetőség, hogy a szurkolói avatárokat összekapcsolt adatokként (linked data) a weben publikáljuk. Mivel nem nagy munkával le tudjuk automatikusan generálni a szurkolói avatárokból a szükséges RDF (Resource Description Framework) XML (vagy akár N3 vagy Turtle) állományokat. Ebben az irányban is végeztünk néhány bevezető mérést. Az 1. táblázatban a 2010-es világbajnokság Spanyol nemzeti tizenegyet jellemző avatár fájljából próbaképpen készített RDF XML állomány részletét mutatjuk meg. A bibliográfiai jellemzéshez Dublin Core [13], a személyek leírásához FOAF [14] formát alkalmaztunk, miközben ODP [15] kategóriákat is használtunk.

#### 1. táblázat

*A 2010-es világbajnokság Spanyol nemzeti tizenegyet jellemző avatár fájljából próbaképpen készített RDF XML állomány eleje.*

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:taxo="http://purl.org/rss/1.0/modules/taxonomy/"
  xmlns:sa="http://footballerml.sourceforge.net/sa/">
  <!-- ESP -->
  <rdf:Description rdf:about="http://footballerml.sourceforge.net/supporter_avatars/.../-ESP.xml">
    <dc:title>Spain 2010 World Cup supporter avatar</dc:title>
    <dc:subject>
      <rdf:Description>
        <taxo:topic rdf:resource="http://www.dmoz.org/Sports/Soccer/UEFA/Spain/National_Team/" />
        <rdf:value>Spanyol labdarúgó válogatott</rdf:value>
      </rdf:Description>
    </dc:subject>
    <dc:source> http://footballerml.sourceforge.net/supporter_avatars/.../-ESP.xml
    </dc:source>
    <dc:creator>
      <foaf:Person rdf:about="http://www.inf.unideb.hu/~nbatfai/">
        <rdf:value>nb</rdf:value>
        <foaf:mbox rdf:resource="mailto:batfai.norbert@inf.unideb.hu"/>
      </foaf:Person>
    </dc:creator>
```

```

<dc:description>Ez a 2010-es VB játészó spanyol csapat avatárja. Példaként raktam össze,
    hogy segítse a leendő avatár fejlesztők első próbálkozásait.
</dc:description>
<dc:date>2010-07-17</dc:date>
<dc:publisher>Supporter Avatar Database</dc:publisher>
<dc:rights>GNU General Public License</dc:rights>
<sa:avatars>
  <rdf:Description rdf:ID="Spain">
    <sa:starting_team>
      <rdf:Bag>
        <rdf:li>
          <rdf:Description rdf:ID="Spain_squad_number6">
            <sa:footballer_avatar>
              <foaf:Person rdf:about="http://dbpedia.org/page/Andrés_Iniesta">
                <rdf:value>6</rdf:value>
              </foaf:Person>
            </sa:footballer_avatar>
            <sa:shutting_goal>0.6</sa:shutting_goal>
            <sa:gaining_bal>0.1</sa:gaining_bal>
          </rdf:Description>
        </rdf:li>
      </rdf:Bag>
    </sa:starting_team>
  </rdf:Description>
</sa:avatars>

```

A (továbbiakban kialakítandó) RDF formába konvertált állományokat a Jena [17] segítségével könnyedén N3 formába másolhatjuk, amit már bemenetül tudunk felkínálni a telepített Joseki [18] szervernek. A 2. táblázat egy olyan SPARQL (SPARQL Protocol and RDF Query Language) lekérdezést mutat, amely a 10 százaléknál erősebb támadójátékot mutató egyé-

ni játékosokat keresi meg. (Jobban átgondolt FerSML RDF konverzió után az ilyen lekérdezések nem lesznek indokolatlanul bonyolultak, de jelen fázisban nem iteráltuk ennek a konverzióknak a meghatározását. Az első, példaképpen az 1. táblázatban mutatott RDF alapján dolgoztunk mechanikusan.)

## 2. táblázat

*A bevezető mérés során kiadott egyik példa SPARQL lekérdezés.*

```

PREFIX sa:    http://footballerml.sourceforge.net/sa/
PREFIX dc:    http://purl.org/dc/elements/1.1/
PREFIX foaf:  http://xmlns.com/foaf/0.1/
PREFIX rdf:   http://www.w3.org/1999/02/22-rdf-syntax-ns#
PREFIX rdfs:  http://www.w3.org/2000/01/rdf-schema#
SELECT ?Mezszám ?Játékos ?Valség ?Fejlesztő
WHERE {
  ?j sa:shutting_goal ?Valség .
  ?j sa:footballer_avatar ?Játékos .
  ?Játékos rdf:value ?Mezszám .
  ?csapat ?k ?bag .
  ?bag rdfs:member ?j .
  ?a sa:avatars ?csapat .
  ?a dc:creator ?f .
  ?f foaf:mbox ?Fejlesztő .
  FILTER (?Valség > „0.1”)
}

```



A 3. táblázatban látható, hogy a lekérdezés eredményeképpen a játékosok mezszámát, magát a játékost, a támadójátékát jellemző relatív

gyakoriságot és az avatár-fejlesztő email címét nyomtattuk ki.

**3. táblázat**  
A példa SPARQL lekérdezés eredménye.

Mezszám	Játékos	Valség	Fejlesztő
13	<a href="http://dbpedia.org/page/Alexander_Frei">http://dbpedia.org/page/Alexander_Frei</a>	0.2	mailto:ebatfai@lib.unideb.hu
6	<a href="http://dbpedia.org/page/Andrés_Iniesta">http://dbpedia.org/page/Andrés_Iniesta</a>	0.6	mailto:batfai.norbert@inf.unideb.hu
8	<a href="http://dbpedia.org/page/Nigel_de_Jong">http://dbpedia.org/page/Nigel_de_Jong</a>	1.0	mailto:batfai.norbert@inf.unideb.hu

## Értékelés

Jelen munkánkban összegeztük azokat az elképzeléseinket, melyek alapján elkezdjük majd a szurkolói avatárok kapcsán a közösségépítést. Reméljük, ehhez ez a munka is hozzájárul, egyrészt a téma népszerűsítésével, másrészt a szurkolói avatárok jövője szempontjából gyümölcsözni képes szakmai kapcsolatok kialakításának megkezdésével.

Most azon dolgozunk, hogy az I., azaz a Debrecen tartalmazó csoport mérkőzéseinek leírásával a futballt értők és szeretők be tudjanak kapcsolódni a szurkolói avatárok készítésébe. Ezt támogatandó, egy webes felületet kínálunk majd fel, ahol a szurkolók egy egyszerű táblázatos formában adhatják meg megfigyeléseiket, melyekből elkészítjük a FerSML XML és RDF állományokat. További terv, hogy az RDF formátumba konvertált adatokat a nálunk futó Joseki szerver SPARQL végpontján keresztül is elérhetővé tesszük [16], amivel a szurkolói avatárok a szemantikus web szerves részévé válhatnak.

## Irodalom

1. BÁTFAI, Norbert: Footballer and Football Simulation Markup Language and related Simulation Software Development. = *Journal of Computer Science and Control Systems*, vol. 3. 2010. no. 1. pp. 13–18.
2. BÁTFAI Norbert: Bevezető számítások a labdarúgás szimulációs jelölőnyelv kialakításához. = *Híradástechnika*, 65. évf. 2010. 5–6 sz. 16–20. p.
3. BÁTFAI, Norbert. (2010) The Soccer Force, *ArXiv e-prints*, 2010., <http://arxiv.org/abs/1004.2003>
4. BÁTFAI, Norbert. (2010) Football(er) Simulation Markup Language, *The project page*, <https://sourceforge.net/projects/footballerml/>
5. BÁTFAI, Norbert (2010): Footballer and Football Simulation Markup Language and related Software, *The project webpage*, <http://footballerml.sourceforge.net/>
6. BÁTFAI, NORBERT – BÁTFAI, Erika: (2010) Distributed Supporter Avatar Database In World Football, *submitted*.
7. BÁTFAI Norbert (2010): Labdarúgás szimulációk a gyakorlatban, *blog*, <http://fersml.blog.hu/>
8. BÁTFAI Norbert (2010): Nemzeti Szurkolói Avatár Adatbázis,



### Köszönetnyilvánítás

A publikáció elkészítését részben a TÁMOP 4.2.1./B-09/1/KONV-2010-0007 számú projekt támogatta. A projekt részben az Új Magyarország Fejlesztési Terven keresztül az Európai Unió támogatásával, az Európai Regionális Fejlesztési Alap és az Európai Szociális Alap társfinanszírozásával valósult meg.

- blog*, <http://nsza.blog.hu/>
9. BÁTFAI Norbert – BÁTFAI Erika (2010) Nemzeti labdarúgó bajnokságok összehasonlító vizsgálata. [megjelenés alatt]
  10. RELAX NG Compact Syntax Tutorial, (2003) <http://www.relaxng.org/compact-tutorial-20030326.html>
  11. OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML, (2005) <http://ocoinfos.info/>
  12. Zotero, <http://www.zotero.org/>
  13. The Dublin Core Metadata Element Set (2007), <http://www.ieft.org/rfc/rfc5013.txt>
  14. FOAF Project homepage (2007), <http://rdfweb.org/foaf/>
  15. Open directory project, <http://www.dmoz.org/>
  16. BIZER, Chris – CYGANIAK, Richard – HEATH, TOM (2007): How to publish linked data on the web, <http://www4.wiwiwss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
  17. Jena, a Semantic Web Framework for Java, <http://jena.sourceforge.net/>
  18. Joseki, a SPARQL Server for Jena, <http://joseki.sourceforge.net/>
  19. MALMSTEN, Martin (2008), Making a library catalogue part of the semantic web, DCMI '08. In: Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications, pp. 146–152., <http://dcpapers.dublincore.org/ojs/pubs/article/download/927/923>
  20. BURKE, MARY (2009), The semantic web and the digital library. = Aslib Proceedings: New Information Perspectives, pp. 316–322, DOI: 10.1108/00012530910959844,
  21. HUNTER, Jane –CHOUDHURY, Sharmin (2004), A semi-automated digital preservation system based on semantic web services, JCDL '04. In: Proceedings of the 4<sup>th</sup> ACM/IEEE-CS joint conference on Digital libraries, pp. 269–278, <http://doi.acm.org/10.1145/996350.996415>
  22. MARKETAKIS, Yannis – TZANAKIS, Makis –TZITZIKAS, Yannis (2009): PreScan: towards automating the preservation of digital objects, MEDES '09. In: Proceedings of the International Conference on Management of Emergent Digital EcoSystems, pp. 404–411, <http://doi.acm.org/10.1145/1643823.1643898>
  23. CAMPBELL, Laura E. (2009), Recollection: integrating data through access, ECDL'09. In: Proceedings of the 13<sup>th</sup> European conference on Research and advanced technology for digital libraries, pp. 396–397.



## A Google Books projekt újabb állomása

Az európai könyvtárak újabb nagyágyújával, a British Libraryvel kötött együttműködési megállapodást nemrég a Google. A megállapodás értelmében 250 ezer 1700 és 1870 között készült, szerzői jogvédelem alá nem eső könyv és időszaki kiadvány digitalizálását végzik majd el.

A digitalizálás költségeit a Google vállalja, cserébe a British Librarynek teljes egészében nyilvánossá kell tennie ezeket a dokumentumokat. A 250 ezer dokumentum között tudományos és szépirodalmi művek, könyvek és időszaki kiadványok válnak használhatóvá a tudományos igényű érdeklődők számára. A projekt végére kb. 40 millió oldalnyi új dokumentum lesz elérhető ingyenesen a Google Books szolgáltatásában.

Jelenleg a Google Books adatbázisában mintegy 13 millió digitalizált könyv található.

(Forrás: [http://www://mmonline.hu/cikk/tobb\\_szazezer\\_konyvet](http://www://mmonline.hu/cikk/tobb_szazezer_konyvet))