

## **Digitalizált Törvényhozási Tudástár az Országgyűlési Könyvtárban**

**Interjú Ambrus Jánossal, az Országgyűlési Könyvtár főigazgatójával**

2011. április 22. dr. Horváth Eszter, Budapest

**Mintegy 2 millió oldalnyi jogforrás, történeti országgyűlési dokumentum, jogtudományi, politikatudományi és történettudományi szakirodalom válik hamarosan hozzáférhetővé az Országgyűlési Könyvtár „Digitalizált Törvényhozási Tudástár” elnevezésű projektje eredményeképpen. A létrejövő portál tartalomszolgáltatásával hatékonyan kívánja segíteni a jogalkotási munkát és a jogalkotás háttérében folyó szakmai tevékenységet is többek között. A projektről Ambrus Jánost, az Országgyűlési Könyvtár főigazgatóját kérdeztük.**



**Mit takar a „Digitalizált Törvényhozási Tudástár” kifejezés, s mi vezette a Országgyűlési Könyvtárat arra, hogy belevágjon a projektbe?**

A kifejezés részben egy projektet, részben pedig az annak eredményeképpen létrejövő digitális gyűjteményt, portált jelenti. A projekt, amely az Új Magyarország Fejlesztési Terv Elektronikus kormányzat operatív programjának támogatásával valósul meg 2 millió oldalnyi könyvtári dokumentum digitalizálását, adatbázisba szervezését és portálon való szolgáltatását tűzte ki célul. Maga a DTT portál – a szerzői jog előírásainak figyelembe vételével, – ingyenes hozzáférést kínál majd mindazokhoz a jogforrásokhoz, országgyűlési dokumentumokhoz és szakirodalomhoz, melyek a jog, a politikatudomány és a legújabb kori történettudomány iránt érdeklődők számára nélkülözhetetlenek; s teszi ezt digitális formában.

A tudástár létrehozásának több oka van. Az egyik nagyon általános. Napjainkban nemcsak az információval foglalkozó szakemberek, hanem az átlagos könyvtárhasználók, tanulmányaikhoz, kutatásaikhoz vagy akár csak a mindennapi munkájukhoz információt keresők is mind erősebben érzékelik a digitális tartalmak terjedését, s annak szükségét, hogy az információ ott és akkor legyen elérhető, ahol és amikor szeretnénk. Világszerte növekszik az internetre felkerülő tartalmak mennyisége. Mivel a könyvtárhasználók részéről is egyre erősebb elvárás, hogy otthon vagy az iskolában, a számítógép mellett ülve tudjanak egyre több szöveget lehívni, a könyvtárak számára egyértelmű a feladat. Létre kell hozni ezeket a digitális tartalmakat, s elérhetővé kell tenni őket az érdeklődők számára. Részben emiatt gondoltuk úgy, hogy a könyvtár állományában, tulajdonában lévő könyvek meghatározó részét digitalizáljuk. A másik ilyen igény az a konkrétabb parlamenti igény volt. A törvényhozási folyamatban számos forrásra, háttéranyagra, szakirodalomra van szükség, amelyek digitális formában sokkal könnyebben és gyorsabban elérhetők a parlamenti felhasználók számára is. A Digitalizált Törvényhozási Tudástár projekt, s az annak eredményeképpen létrejövő portál ezen kihívásoknak kíván megfelelni.

**Hogyan tudja segíteni a Tudástár a jogalkotási tevékenységet?**

A jogalkotási munkának rendkívül komoly szakmai háttérigénye van. Ezt részben a jogszabályt beterjesztők végzik el a jogszabály előkészítésekor. Másrészt a javaslat országgyűlés elé kerülésekor, a bizottságok, a képviselők, a plenáris ülés színvonalas munkájához, szükség van bizonyos háttérinformációkra. Ezeket többnyire szakértők „szállítják”, de az ő ismereteikhez nélkülözhetetlen háttérbázist nyújtanak a könyvtárakban (folyóiratokban, könyvekben) fellelhető anyagok, melyek digitalizált formában könnyebben hozzáférhetők.

**Milyen egyéb gyakorlati haszna lesz még a Tudástárnak a jogi szakma számára?**

Az egyetemi oktatásban, de akár csak a konkrét, mindennapi jogi munkához is sokféle dokumentumot használnak a jogászok. Oktatók, kutatók, joghallgatók, gyakorló szakemberek egyaránt. A kétmillió

oldal részeként az Országgyűlési Könyvtár hatszázezer oldalnyi jogi folyóiratot is feldolgoz. Ezek többsége történeti folyóirat, s egy ma folyó perben, vagy eljárásban nyilván csekélyebb gyakorlati haszna van egy 1906-os jogelméleti folyóiratnak (bár nem tartom kizártnak, hogy van olyan ügy, ahol pont ilyesmi segíthet). Az oktatásban és a képzésben ezeket az anyagokat viszont egész biztosan használnák, hiszen ma is használják. A digitalizálásra ugyanis elsősorban azokat az anyagokat válogattuk össze, melyeket a könyvtár olvasói, tapasztalatunk szerint, rendszeresen igényelnek.

### **Milyen nagyságrendű munka a Tudástár elkészítése? Mekkora apparátusra, milyen technikai háttérre van szükség?**

Ez a munka nagyon összetett és bonyolult. Minden részletét, nehézségét a munka kezdetekor még nem láttuk, nem láthattuk teljes terjedelmében. Sok mindenről mesélhetnék, próbálok a súlyponti részekre koncentrálni. Az egyik ilyen a dokumentumok kiválasztása. Nyilvánvaló, hogy olyan dokumentumokat kell kiválasztani az egyes szakterületeken – döntően jogtudományról van szó –, amelyek a felhasználók érdeklődésére leginkább számot tartanak. Felesleges lenne olyasmit digitalizálni, amit 30 év alatt egy vagy két olvasó néz(ett) meg. A másik szempont a válogatásnál inkább gyakorlati. Csak olyan dokumentumokat tudunk digitalizálni, amelyeknek az állapota ezt megengedi. Azt, hogy ezek melyek, nem lehet előre megmondani, hiszen csak az egyes dokumentum kézbevételekor győződhetünk meg arról, hogy nem hiányoznak-e abból oldalak, hogy az olvasók nem rongálták-e meg az adott könyv vagy folyóirat lapjait úgy, hogy azt már ha akarnánk se tudnánk digitalizálni. Vannak ugyanakkor olyan speciális technológiai sajátosságok, melyek szintén ellehetetleníthetik a digitalizálást. Régen például sok dokumentumot nagyon vékony papírra nyomtattak. Ezen a szöveg átüt, így a digitalizált végeredmény olvashatatlaná válik, vagy ha a szöveg képként értelmezhető is, a karakterfelismerésre már nem ad lehetőséget. Céljaink közt pedig az is szerepel, hogy az olvasók ne csak a képet lássák, hanem keresni is tudjanak a szövegben. Ehhez azonban azt az ún. OCR technológiával „karakterfelismertetni” kell. Hasonlóan, nem könnyű a digitalizált végeredményt olyan gyűjteménybe rendezni, hogy abban az olvasók könnyen megtalálják majd, amit keresnek. Ha azt szeretnénk, hogy az olvasó adattárként tudjon tájékozódni például a fent említett hatszázezer oldalnyi anyagban – azaz szerzőre, vagy akár egy cikkre vissza tudjon keresni –, ehhez a hatszázezer oldalnyi anyagot nem elég digitalizálni, hanem cikkenként is fel kell dolgozni. A könyvtár adatbázisaiban számos írás már feldolgozott, de a folyóiratállomány jó része még utólagos „adatolásra” szorul. Ez szintén hatalmas feladat és jelentős része ennek a munkának.

### **Vannak –e európai, hazai példák amelyek segíthetik a munkát?**

Számos hazai és nemzetközi példa van az ilyesfajta digitalizálásra, de rendkívül sokféle módszerrel és metodikával. Ha a nálunk folyó munkát a Magyar Elektronikus Könyvtárhoz hasonlítom, vagy azt tekintem példának, akkor azonnal szembetűnik egy alapvető különbség. Ott gyakorlatilag „leíródik” az anyag. Nem a dokumentumot digitalizálják, s nem annak elektronikus képe jelenik meg, hanem maga a beírt, begépelte szöveg, mondjuk József Attila össze költeménye, vagy Mikszáth összes művei. Ez teljesen más metodika, így az a módszer, amit ott alkalmaztak nem feltétlenül ad támpontot nekünk. Ezért megpróbálunk egy saját módszertant kialakítani.

Az egész folyamat háttérében az áll, hogy ma Magyarországon nem létezik korszerű és hozzáférhető elgondolás arról, hogy melyek azok az alapidokumentumok, amelyeket az egyes tudományterületeken első lépésben kellene digitalizálva hozzáférhetővé tenni. Mindenki nagyon félve vállalja csak a döntés meghozatalát arról, hogy – bármely okra hivatkozva is, de kimondja – ezt hozzáférhetővé kell tenni, azt pedig nem. Ha ebben a kérdésben a döntéshozókban több határozottság és vállalkozó szellem lenne, akkor – bár ez nyilván vitákat generálna és fog is generálni, – mégiscsak közelebb jutnánk ahhoz, hogy egy adott tudományterületen definitív módon meg tudjuk mondani, hogy mit is kellene feldolgozni. Ez az anyag, amit mi most a projekt keretében feldolgozunk, gyakorlatias válasz erre a problémára. Mi nyilván csak a saját állományunkból tudunk válogatni. Van olyan dokumentumcsoport – ilyenek például a megyei közlönyök –, aminek a feldolgozásáról azért mondtunk le most az első körben, mert itt az Országgyűlési Könyvtárban csak egy töredékével rendelkezünk. Jelenleg sajnos utópisztikus olyan kooperációban gondolkodni, amellyel ezt a problémát rendezni tudnánk. Ez majd a jövő útja lehet. A könyvtárak között az ilyen ügyekben nélkülözhetetlen lesz az együttműködés.

### **Milyen nehézségekkel kell szembenézni a munka során?**

Érintettük itt már a digitalizálás során felmerülő problémákat.

Hasonló érdekes, nagy kihívást jelentő probléma a szerzői jogok kérdése. A jogszabályokon kívül, melyek többnyire szabadon hozzáférhetők a művek általánosságban a szerző halálát követő 70 évig szerzői jogvédelem alatt állnak. Ez a gyakorlatban azt jelentheti, hogy egy 1910-es folyóirat esetében sem biztos, hogy a cikkeknek nem a többsége az, ami szerzői jogvédelem alatt áll. Itt számos nemzetközi tapasztalat, példa van. A gyakorlat azt mutatja, hogy aki nagyon biztosra akar menni, csak olyan anyagot dolgoz fel, ami intakt a szerzői jogi problémától. Ott általában megállnak az 1880-as, 1890-es éveknél. Legutóbb az Osztrák Nemzeti Könyvtár főigazgatója nyilatkozott a témában. Ők a közel 400.000 kötet digitalizálásakor az 1880-as évnél megálltak a feldolgozással, mert a későbbi műveket kockázatosnak látták feldolgozni a szerzői jog miatt. Mi megpróbálunk nem megállni 1880-nál. Ebből az elhatározásunkból ismét nagyon sok feladat adódik. Egy konkrét példa a jogász olvasók számára: Szladits Károly 1956-ban halt meg, így az ő esetében például még a fiatalon publikált művei is jogvédettek.

### **Hogy fog kinézni a portál? Mit képzeljen maga elé a majdani felhasználó?**

Hát, erről annyira könnyű mesélni, mint egy illatról. Azt, hogy a készülő portál, mint vizuális élmény milyen lesz, még mi sem látjuk pontosan. A funkciókról tudok néhány dolgot elmondani. Az a szándékunk, hogy tematikus csomópontokat alakítunk ki. Konkrétan: a jogtudomány esetében például jogágakra bontjuk a feldolgozott műveket és ezen belül még egy szintig biztosítunk majd tájékoztató, böngészési lehetőséget. A digitális objektumokhoz rendelt metaadatok révén lehetőség lesz közvetlenül rákeresni szerzőre, címre, az OCR technológia pedig a teljes szövegben való keresést teszi lehetővé. Elképzelhető ugyanakkor olyan eset, amikor ez inkább probléma lesz a felhasználónak. Például az „alkotmányjog” kifejezésre több tízezer találatot is kaphat tíz vagy százezer oldalon. Így az olvasóknak is meg kell tanulni majd azt a technikát, amivel a portálon hatékonyan tudnak keresni. A jelenleg folyó digitalizálási munka nem előzmény nélküli. Korábban feldolgoztuk már a parlamenti almanachokat és a parlamenti házszabályokat, 2006-2008 között pedig a parlamenti (képviselőházi és felsőházi) naplókat. Ezek az interneten már elérhetők, a technikai megoldásokkal azonban, sajnos, nem mindig vagyunk elégedettek. A házszabályok esetében például csak a dokumentumokat lehet megnézni az interneten, nincs szöveges felismerés, azaz nem lehet szövegrészletre keresni. Van egy olyan elképzelésünk, hogy ezeket az anyagokat (közel egymillió oldal) is beemeljük a Tudástárba, de a keresetőséget más technikával oldanánk meg. Van tehát a jelenleg folyó munkának egy komoly előzménye már itt a könyvtáron belül.

### **Hol tart most ez a folyamat és mikorra várható a Tudástár elkészülése?**

A dokumentumok leválogatását, állapotfelmérését, az egyes művek szerzői jogi vizsgálatát, a digitalizálás és a teljes munkafolyamat aprólékos megtervezését (2010) követően beszélgetésünk idején a közbeszerzés-nyertes BravoGroup/DocuScan konzorcium munkája révén már folyik a 2 millió oldal digitalizálása, a metaadatok készítése, az adatbázis-építés, s elkezdődtek a portál kialakításával kapcsolatos munkálatok is. A projektet 2011 őszére fogjuk befejezni; eredményéről, természetesen, hírt adunk majd.

### **Köszönöm a beszélgetést, további jó munkát kívánok!**

-----  
Ambrus János könyvkiadó, könyvtáros. A Bölcsészettudományi Karon szerzett tanári és könyvtárosi diplomát; az 1973-ban alapított esztétika szak első végzős évfolyamának tagjaként végzett az ELTE-n. Ezt követően az MTA Filozófiai Intézetének munkatársa, 1989-től a T-Twins Kiadó alapítója és igazgatója, 1996-tól az Országgyűlési Könyvtár főigazgatója.

Forrás: <http://www.jogiforum.hu/interju/82#axzz1MWGsWOGF>

Letöltés dátuma: 2011. május 16.