

Csernoch Mária

## Túl az Óperencián. Idegen nyelvű digitális könyvtárak állományai felhasználói szemmel

*A digitális könyvtárak széles körű elterjedése, szolgáltatásaik jellege, elérhetősége nagymértékben megváltoztatta a könyvtárhasználó közösség összetételét, könyvtárhasználati szokásait. A hagyományos értelemben vett, elsősorban fizikai jellegű korlátok megszűntek, jelentkeztek azonban újak, amelyeket az informatikai eszközök – hardver és szoftver egyaránt – és az informatikai írástudás jelenléte, esetleg hiánya eredményezhet. A digitális könyvtárak által az olvasóközönséghez, felhasználókhöz eljuttatott tartalmak és azok megjelenési formája rendkívül sokféle lehet. Az ezek közötti eligazodás, tájékozódás, a találatok értékelése mindenképpen szükséges a digitális könyvtárak nyújtotta szolgáltatások hatékony kihasználása érdekében, amelyekhez felhasználói oldalról megközelíthető gyakorlati IKT (Információs és Kommunikációs Technológia) ismeretekre is szükség van.*

### Bevezetés

A könyvtárak fejlődési folyamatában a gyűjtemények adottságai minden időben befolyásolták a könyvtár használóinak, olvasóinak összetételét<sup>1</sup>. Ez az olvasótábor a digitális gyűjtemények esetében jelentős mértékben megváltozott, hiszen a könyvtár átlépte azokat a korlátokat, amelyek évszázadokon át meghatározták működését [1]. A digitális könyvtárak olyan tartalmakat, szolgáltatásokat kínálnak, főként az olvasni vágyók és az irodalmi, nyelvészeti kutatásokat végzők számára, amelyek a hagyományos könyvtári keretek között nem, vagy csak komoly nehézségek árán voltak elérhetők. Elsősorban a hagyományos értelemben vett fizikai korlátok azok, amelyek egyrészt megszűntek, átalakultak, másrészt az e-formátumok új megjelenését eredményezték. Mindezen változásokat figyelembe véve adódik, hogy a digitális könyvtárak jellege, az általuk képviselt és közvetített értékek természetes módon megváltoztatják könyvtárhasználati, olvasási, nyelvtanulási, nyelvhasználati szokásainkat, más népekhez, kultúrákhoz kötődésünket.

### Elérhetőségek, korlátok

#### *Hagyományos fizikai korlátok*

A fizikai korlátok alatt értjük egyrészt a földrajzi és az időbeli korlátokat. A földrajzi korlátok jelenthetik az olvasó és a nyomtatott formátumú könyv közötti távolságot, amely egyes esetekben kilométerekben, ezer kilométerekben mérhető, míg más ese-

tekben a távolságok rövidebbek, de az olvasóra, a könyvtárra, vagy a könyvre vonatkozó valamilyen korlátozás miatt nem hidalhatók át. Az időbeli korlátok főként a könyvtár nyitvatartási idejére vonatkoznak, de ide érthetjük a könyvtáros elfoglaltságából adódó időbeli korlátokat is.

Fizikai jellegű korlátok közé sorolhatók a kommunikációs, elsősorban nyelvi jellegű akadályok. A hagyományos könyvtárakban a nyelvi kommunikációban a szervi adottságaik miatt hátrányban élők csak rendkívül alacsony számban jelentek meg. Ezzel szemben a digitális könyvtárak olyan új lehetőségeket biztosítanak az olvasás módjának megváltozásával, kibővítésével, hogy a vakok és gyengén látók, a diszlexiával küzdők aktív olvasókká válhatnak.

Az olvasási képességeikben nem korlátozottaknak elsősorban az idegen nyelvi korlátok csökkentése, megszüntetése jelentheti a digitális könyvtárak előnyeit. A hagyományos könyvtárakban a nyelvi korlátok jelentkezhetek az olvasó és a könyvtáros közötti kommunikáció formájában, de valójában a könyvtári gyűjtemény nyelvének a korlátozottsága az, amely feloldható a digitális könyvtárak elterjedésével. A digitális könyvtárakon keresztül például a nyelvtanulásnak egy új formája kezd kibontakozni, lehetőség nyílik arra, hogy a tankönyvalapú nyelvtanulást kiegészítse, akár le is váltsa az irodalmi műveken keresztül történő nyelvtanulást. Ennek módszerei még nem kidolgozottak, de elvitathatatlán tény, hogy a digitális könyvtárakon, digitális könyvesboltokon keresztül a könyvek szín-

te bármilyen nyelven elérhetővé váltak, lehetőség nyílt azok számítógépes feldolgozására.

### **E-korlátok**

A fizikai elérhetőség magában foglalja a szöveg megjelenési formáját is, tehát azt, hogy a papíralapú formátumot a fájlformátum váltotta le. A szövegek megjelenési formájának megváltozása a hordozó eszköz változását vonja maga után, annak minden előnyével és hátrányával. A szövegek fájlformátumban történő megjelenése egy, a fizikai korlátokat is érintő érdekes paradoxonhoz vezetett. Ami nem létezik, arra nehéz azt mondani, hogy valaminek a korlátja. Így vagyunk az e-dokumentumokkal, könyvekkel is. A számítógépek megjelenése előtt nem mondhattuk, hogy bármiféle akadályt is jelentett az e-dokumentumok hiánya. A számítógépek elterjedésével azonban egyértelművé vált, hogy a szövegek számítógéppel segített feldolgozásának akadályát jelentette a digitalizált szövegek, a digitális könyvtárak nem létezése. Tekinthejük tehát úgy, hogy ez a fizikai korlát is megszűnt.

A szövegek korábbi, papíralapú formátuma az olvasási képességek jelenlétén túl a mozgásukban nem korlátozottaktól nem feltételezett semmiféle eszközhasználati műveltséget, kultúrát. Ezzel szemben az e-formátumú dokumentumok használata feltételezi valamiféle számítógép vagy ahhoz csatlakoztatható – akár vezetéken keresztül, akár vezeték nélkül – eszköz jelenlétét, ezen túl az IKT eszközök használatát, a kezelési módok ismeretét, tehát valamilyen szintű informatikai írástudást. A nem megfelelő eszköz jelenléte, a megfelelő eszköz esetleges hiánya, az eszköz használati módjának elsajátítása jelent és jelenthet akadályokat. Várható azonban, hogy folyamatosan nő a számítógépes műveltséggel, számítógéppel, egyéb IKT eszközökkel, hálózati hozzáféréssel rendelkezők aránya, és ennek következtében csökken az e-dokumentumok elérésében akadályoztatottak száma [2]. Egészséges emberek számára az e-formátumú könyvek a hordozóikkal együtt olyan lehetőségeket jelentenek, amelyek kiegészítik a könyvek korábbi lehetséges elérhetőségi módjait. A fogyasztókkal élőknek azonban az e-könyv lényegesen többet nyújt egy választási lehetőségnél, még abban az esetben is, ha nem kommunikációs képességeikben korlátozottak. Megszűnik, megszűnhet minden olyan akadály, amely a könyvhöz való hozzájutásban, a könyv fizikai valójában, például lapozásában, nehézséget okozhat.

További korlátozást jelenthet az állomány nem megfelelő formátuma. Nem szabad megfeledkezünk arról, hogy a szövegek különböző fájlformátumban történő megjelenése nem minden esetben jelent valódi szöveget. A laponként, képként digitalizált könyvek megjelenítésére szolgáló egyik lehetséges megoldás a kétrétegű megjelenítés, melynek lényege, hogy a felhasználó az eredeti szkennelt oldalképet látja, mögötte/alatta ott van egy másik rétegben a szöveg, így lehet keresni a könyvben. A másik, kevésbé szerencsés megoldás, amikor a fájlok tartalmukat tekintve szövegek, fájlformátumukat tekintve azonban nem képek, még abban az esetben sem, ha ezek a képek utólag egy programmal összeszerkesztésre kerültek egy dokumentumon belül, létrehozva ezzel a módszerrel egy álszöveges dokumentumot. Mindkét megoldásra láthatunk példát a *Gallica* és a *Google Books* fejezetekben. Felhasználói oldalról szemlélve, a valódi és az álszöveges formátumok között az az elsődleges különbség, hogy míg a valódi szöveges állományok teljes egészében karakter szinten kereshetők, addig a csak képeket tartalmazó ál-állományok nem rendelkeznek ezzel a tulajdonsággal.

A szöveges formátumok előállítására vagy a hagyományosnak tekinthető gépeléssel történik vagy szkenneléssel. Ám ebben a második esetben a digitalizálás során nem csupán a könyvek képformátumú előállítása történik, hanem szövegfelismerés is, valamilyen OCR (Optical Character Recognition) programmal.

A digitális könyvtárakban elérhető különböző típusú fájlformátumok említésekor nem feledkezhetünk meg a hangoskönyvekről. Alapvetően kétféle hanganyagot lehet találni a digitális könyvtárakban: felolvasott és számítógép-generált<sup>1</sup> hangfájlokat. A számítógép-generált hanganyag első hallásra ijesztőnek tűnhet, de felbecsülhetetlen értékű azok számára, akik látásukban korlátozottak. A felolvasott szövegek előállítása lényegesen több emberi erőforrást és időt igényel, mint a számítógép-generált hang. A felolvasott hanganyagok minőségében is lehetnek azonban eltérések. Érdemes összehasonlítani a *Project Gutenberg* és a *Magyar Elektronikus Könyvtár (MEK)* színészek által felolvasott hanganyagait – példaként meghallgathatók a „The Goose That Laid The Golden Eggs”<sup>2</sup> és a „Kék pille”<sup>3</sup> hangfájlok. A MEK hanganyaga a *Magyar Vakok és Gyengénlátók Országos Szövetségének* támogatásával jött létre, többnyire kiváló minőségű hanganyagok feltöltésével, míg a Project Gutenberg hanganyagainak előállításában többsé-

gében hangfelolvasásban gyakorlatlan önkéntesek segédkeznek.

Jelen írásban, a digitális könyvtárak elemzése során különös hangsúlyt kapnak a különböző fájlformátumok, azok lehetséges felhasználási módjai, esetleges korlátai.

### **Mennyiségi változások**

Az e-könyvek, a digitális könyvtárak az elérhető szövegek mennyiségének megváltozásával is átalakítják gondolkodásunkat. A speciális esetektől eltekintve, amelyekben a nem elérhetőséget vagy a korlátozott elérhetőséget felváltotta a digitális könyvtárak „határtalansága”, a mennyiségi változás nem feltétlenül jelenti az aktív olvasók számának jelentős emelkedését, nem feltétlenül jelenti az emberek olvasási szokásainak változását, eltekintve az adathordozók megváltozásából adódó eltérésektől, mint például az olvasási sebesség és az információszerezés hatékonyságának változása [3].

Ezzel szemben a digitális szövegek jelenléte az irodalmi és nyelvészeti kutatásokban olyan, eddig nem látott teret nyitott, amely megváltoztatta, megváltoztathatja ezeket a tudományterületeket, illetve új kutatási területek jelentek, jelenhetnek meg. Alapjaiban megváltozott például a szótárkészítés tudománya, a korpusznyelvészet önálló tudományrá fejlődött, és egyre inkább helyet kapnak az olyan jellegű nyelvészeti kutatások, amelyek szövegek statisztikai elemzésével foglalkoznak. Ezek a tudományterületek mind megkövetelik a szövegek elektronikus formátumban és nagy mennyiségben való jelenlétét.

A számítógéppel támogatott szövegfeldolgozás eredményeként napjainkra számos „ipari” alkalmazás is napvilágot látott, amelyek között mindenképpen említésre érdemes a fordító- és helyesírásellenőrző-programok megjelenése, azok megbízhatóságának növekedése. Az idegen nyelvek oktatásában használatos nyelvkönyvek készítésének gyakorlata is átalakult, felhasználva a korpusznyelvészet eredményeit.

### **Olvasóközönség vagy felhasználó**

A könyvek, szövegek elektronikus formában történő megjelenése, azok felhasználhatósága megváltoztatta, megváltoztatja a hagyományos értelemben vett olvasóközönség meghatározását is. Az elektronikus felületen elérhető dokumentumok

megjelenése, elterjedése, népszerűsége megváltoztatta olvasási és tanulási szokásainkat [4] [5] [6] [7], és ez maga után vonta a digitális könyvtárak olvasóközönségének összetételében és szokásaiban bekövetkezett változásokat is.

Ebben a szövegkörnyezetben sokkal inkább beszélhetünk felhasználókról, mintsem olvasóközönségről, mivel a digitális könyvtárak dokumentumai számos esetben nem a hagyományos értelemben vett olvasási céllal kerülnek letöltésre.

Nem beszélhetünk hagyományos értelemben vett olvasásról, amikor a felhasználó számítógéppel segített szövegfeldolgozáshoz kívánja használni a saját gépére letöltött dokumentumokat. Nem tekinthető hagyományos értelemben vett olvasásnak az a tevékenység sem, amely könyvtári és könyvesbolti látogatások során vált ismertté: könyvek átlapozása, „beleolvasgatás”. Az e-gyűjteményekhez kötődően, a teljes könyvek elérhetősége mellett megjelent egy új lehetőség, amikor könyvek csonkított változatai kerülnek az online felületre, bemutató jelleggel, könyvtári, szerzői, kiadói támogatással. A teljes szövegeken elérhető „beleolvasgatás” tehát kiegészült egy leginkább „irányított beleolvasgatás”-nak nevezhető tevékenységgel.

Jelen írás elsősorban az ebben az új értelemben vett szövegfelhasználóknak szól. Ők fizikai valójukban többségében ugyanazok az olvasók, akik korábban a hagyományos könyvtárakat látogatták, de az e-környezetben megváltozott szerepük, és ehhez az új szerephez meg kell tanulni alkalmazkodni, meg kell tanulni kihasználni az új környezet nyújtotta lehetőségeket. Azt kívánjuk bemutatni, hogy az egyes szabad hozzáférésű digitális könyvtárak milyen lehetőségeket, milyen elérhetőségeket biztosítanak, milyen körülmények között lehet a kiválasztott könyvtárat hatékonyan használni annak érdekében, hogy a keresett könyvet a megfelelő formátumban el tudjuk érni.

### **Digitális könyvtárak állományai felhasználói oldalról**

#### ***A válogatás szempontjai***

Jelen írás elsődleges célja magyar anyanyelvű olvasók számára idegen nyelvű, szabad felhasználású digitális könyvtárak, e-gyűjtemények bemutatása, ezek könyvekhez, szövegekhez kötődő szolgáltatásainak, lehetőségeinek összehasonlítása. Az e-gyűjtemények általános bemutatása, azok jellemzőinek leírása azonban nem célunk. A témák

ban érdeklődők számára ajánlom *Koltay Tibor* és *Ördögné Kovács Mónika* [8], valamint *Gobinda* és *Sudatta Chowdhury* [9] írásait, illetve az utóbbi kiadványról a Koltay Tibor által készített ismertetőt [10]. Nem célunk továbbá az e-könyvek általános ismertetése sem, ehhez érdemes *Kerekes Pál* frissen megjelent „E-book kalauz” című művét kézbe venni [11], amely többek között részletesen foglalkozik az e-könyvek jellegével, azok elérhetőségével, az olvasáshoz használható eszközök sokféleségével.

A bemutatásra kerülő gyűjtemények köre egy szigorú válogatás eredménye. Olyan gyűjteményekre esett a választás, amelyek bárki számára hozzáférhetők, amelyek bárkinek lehetővé teszik a szövegek elérését, ha rendelkezik megfelelő adatátviteli kapcsolattal. A gyűjtemények kiválogatásánál a következő szempontokat vettük figyelembe: ha van, akkor, legyen a gyűjtemény az adott ország nemzeti könyvtáré; ezen túl vizsgáltunk nagy, népszerű gyűjteményeket, válogattunk azok bőséges szolgáltatásaiból; valamint azt néztük meg, hogy az egynyelvű és a többnyelvű gyűjtemények milyen lehetőségeket kínálnak.

### Találatok értékelése

A gyűjtemények és a találatok értékelésekor nem szabad azonban elfeledkeznünk arról, hogy milyen szándékkal indítottuk a keresést. Csak akkor tudjuk hatékonyan használni a kiválasztott e-dokumentumot, ha az megfelel céljainknak, ha a felhasználás módja összeegyeztethető a dokumentum formátumával, és nem utolsósorban, ha nincsenek az elérhetőségnek fizikai akadályai. Ezek a feltételek magukban foglalják, hogy nem lehet egy könyvtárról, egy találatról egyértelműen eldönteni, hogy az jó-e vagy nem, mindig a felhasználó szándékai azok, amelyek meghatározzák a találat használhatóságát. Egy számítógéppel generált hanganyag lehet például egy kezdő nyelvtanuló számára teljesen használhatatlan és érthetetlen. Ugyanakkor ugyanez a hanganyag egy látásában korlátozottnak a világ kitérülését jelentheti, egy magasabb szintű nyelvtanulónak segítséget nyújthat beszélt nyelvi szövegek megértésében, amelyekre órarendi keretek között nincs lehetőség. Hasonló módon, bármennyire is értékes műhöz jutottunk, ha a szöveg csak online elérhető, internetkapcsolat hiányában a találatunk értéktelenné válhat. Egy szerző életművének elemzéséhez, szövegek komparatív analíziséhez sem elegendők a csak és kizáró-

lag online elérhető szövegek, hasonlóképpen a csonkított változatok sem. Ilyen esetekben csak azok a találatok tekinthetők teljes értékűnek, amelyeknél a felhasználó az általa megcélzott szövegeket hiánytalanul megtalálja és letöltheti a saját gépére.

A bemutatásra kerülő digitális könyvtárak és gyűjtemények nyelvek szerinti csoportosításban követik egymást, ha meghatározható a szolgáltatás felületének elsődleges nyelve. Találunk azonban olyan gyűjteményeket is, amelyek felületi nyelve IP-specifikus, vagyis alapértelmezésben igazodik a felhasználó földrajzi helyéhez.

## Digitális könyvtárak szolgáltatásai

### Angol felületi nyelvű digitális könyvtárak

Az angol felületi nyelvű könyvtárak a legelterjedtebbek, olyan nagy számban állnak rendelkezésre, hogy valamennyi teljes körű bemutatása jelen írás keretein belül lehetetlenség. A választás az 1. ábrán felsorolt és logójával is bemutatott öt gyűjteményre esett: Project Gutenberg (PG, <http://www.gutenberg.org>), Project Gutenberg Australia (PGA, <http://www.gutenberg.net.au>), ManyBooks (<http://www.manybooks.net>), Classic Bookshelf (<http://www.classicbookshelf.com>) és Aesop's Fables (<http://www.aesopfables.com>). Ezek közül is a Project Gutenberg digitális gyűjtemény kerül részletes bemutatásra, amely már keletkezésének időpontjával is figyelemfelkeltő, mivel 1971-ben indult útjára, abban az időben, amikor még a mai értelemben vett személyi számítógépek és hálózati elérhetőségek nem léteztek (1. ábra A, *Projekt Gutenberg fejezet* és [12]).

### Gutenberg-alapú gyűjtemények

Napjainkra számtalan, a Project Gutenberg<sup>4</sup> gyűjteményhez köthető, ún. Gutenberg-alapú digitális gyűjtemény létezik, amelyek közül mindenképpen említésre méltó a Project Gutenberg Australia<sup>5</sup> és a ManyBooks<sup>6</sup> digitális gyűjtemény (1. ábra B–C), amelyek szolgáltatásaikban, fájlformátumaikban, a könyvek elérhetőségében – a szerzői jogi törvényekben levő eltérések miatt –, speciális gyűjteményeikkel tartalmas kiegészítői az eredeti Project Gutenberg gyűjteménynek. További kapcsolatok és Gutenberg-alapú könyvtárak elérhetők a hetedik hivatkozáson.

### Speciális gyűjtemények

Válogatásunk tartalmaz még további két különleges gyűjteményt is, amelyek említésével az volt a célunk, hogy felhívjuk olvasóink figyelmét a speciális gyűjtemények jelenlétére, jelezve, hogy az általános célú gyűjtemények bőséges tartalmán és szolgáltatásain kívül ezek a weblapok további, főleg tartalomspecifikus szolgáltatásokat nyújthatnak.

A Classic Bookshelf<sup>8</sup> (1. ábra D) sajátossága egy olyan szolgáltatás, amely lehetőséget kínál az olvasófelület beállítására az aktuális igényeknek megfelelően: háttérszín, karakterszín, sormagasság, margó stb., valamint könyvjelző elhelyezésére. A felület ötvözi a monitoron történő megjelenítés és az e-bookolvasók előnyeit.




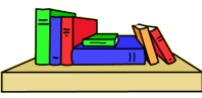

Az Aesop's Fables gyűjtemény<sup>9</sup> (1. ábra E) nagyon jól szemlélteti, hogy egy speciális gyűjtemény milyen tartalmi szolgáltatásokat nyújthat, szemben az általános jellegű digitális könyvtárakkal, kihasználva azt, hogy egyetlen témára fókuszál. A speciális tartalmi szolgáltatásokon túl, hasonlóan a Classic Bookshelf gyűjteményhez, a felületi megjelenés is lehet egyedi, szöveghez kötött. Az Aesop's Fables gyűjteményben található olyan meséket, amelyekhez közvetlenül elérhető egy gyenge minőségű, ám felolvasott hanganyag. Az előre elvégzett párosításnak köszönhetően az írott és hallott szöveg párhuzamosan jelenik meg, igaz, nem emelik ki az éppen olvasott szövegrészt. A sajátos tartalomhoz igazodva, a tartalomjegyzékben a szöveg és a hanganyagok elérhetőségén túl minden mese mellett megjelenik a hozzá tartozó tanulság is.

### Project Gutenberg

Napjainkra a Project Gutenberg gyűjtemény az egyik legkorábban alapított, leghosszabb ideje működő, legnagyobb és legnépszerűbb, bárki számára ingyenesen, szabadon elérhető angol felületi nyelvű digitális könyvtárrá nőtte ki magát (1. ábra A). Fontos hangsúlyozni, hogy míg a Project Gutenberg gyűjtemény alapértelmezett felületi nyelve angol, nem kizárólagosan angol nyelvű könyvek találhatóak benne; tizennégy nyelven ötvennél több, míg további negyvenhat nyelven ötvennél kevesebb könyv érhető el.

További érdekessége a Project Gutenberg gyűjteménynek, hogy teljesen önkéntes alapon szerveződik: ők végzik a könyvek digitalizálását, átolvasását, javítását, a hanganyagok elkészítését, és önkéntes alapon működik a finanszírozás is. Ennek megfelelően örömmel fogadnak minden olvasói hibabejelentést, ezeket figyelembe véve javítják a hibákat és indokolt esetben felülírják a korábbi hibás verziót. A tartalommal kapcsolatos felajánlásokon túl szívesen fogadnak adományokat is a könyvtár működéséhez szükséges költségek fedezéséhez.

A projekt fő profilja a szerzői jogi szabályozás hatályán kívül eső könyvek szöveges formátumban történő digitalizálása és megjelentetése, de az e-könyvek mellett találni a gyűjteményben hanganyagokat – felolvasott és számítógép által generált fájlokat –, álló- és mozgóképeket, zenét, kottákat, egyéb hanganyagokat (például beszédek), számgyűjteményt (például a  $\pi$  értékét, néhány szám négyzetgyökét, kromoszómaszámokat). A gyűjteményt vagy annak egyes részeit CD-re, illetve DVD-re írható formában is megjelentetik<sup>10</sup>.

A	B	C	D	E
				
Project Gutenberg (1971)	Project Gutenberg Australia	ManyBooks	Classic Bookshelf	Aesop's Fables

1. ábra Angol felületi nyelvű általános célú, Gutenberg-alapú digitális könyvtárak (A–C), valamint speciális célú és felhasználású könyvtárak (D–E).

A Project Gutenberg elsősorban könyvek digitalizálásával foglalkozik, jelenleg több mint 39 000 digitalizált dokumentumot tartalmaz<sup>11</sup>. A könyvtárban elérhető valamennyi könyv szövegformátumú. A szkenneléssel és szövegfelismeréssel készült állományoknak két típusát különböztetik meg: a számítógép-generált szövegek tartoznak az egyik csoportba, míg az ellenőrzött szövegek alkotják a másik csoportot. A számítógép-generált szövegeken a szövegfelismerés után nem történik semmiféle ellenőrzés, hibajavítás, tehát várhatóan tartalmaznak szkennelési, leginkább szintaktikai jellegű hibákat, és formázási következetlenségek jelenhetnek meg. Ezzel szemben az ellenőrzött szövegek legfőbb jellemzője, hogy a szövegfelismerés után megtörténik egy, akár többfázisú átolvasás, hibajavítás, és a szövegen belüli formátumok egyezésítése, amit a már említett önkéntesek végeznek. Természetesen a többszöri átolvasás sem zárja ki a hiba lehetőségét, különös tekintettel arra, hogy nem szakemberek végzik az ellenőrzést, de az átolvasott, ellenőrzött szövegek már igen jó minőségűek.

A Project Gutenberg egyik nagy előnye sok más digitális könyvtárral szemben, hogy a tartalma letölthető. Hagyományosan szövegfájlok, .txt és .htm kiterjesztésű fájlok letöltése lehetséges. Napjainkra azonban, igazodva a megváltozott olvasási szokásokhoz és az újonnan megjelent olvasóeszközökhöz, szélesítették az elérhető fájlformátumok skáláját. Ennek következtében e-bookolvasókon és más mobil eszközökön használható formátumokban is elérhetők, illetve letölthetők a szövegek, beleértve a .pdf formátumot is, amely mind asztali számítógépen, mind a hordozható eszközökön olvasható.

A szövegekhez számos esetben hanganyagok is találhatóak, igaz, a szöveges és hangformátumok összepárosítását az olvasóra bízzák, ami a következtelen elnevezések miatt esetenként nem kis feladat. Az elérhető hanganyagok, hasonlóan a szövegekhez, általában többféle formátumban is letölthetők. A Project Gutenberg hanganyagainak előállítását is önkéntesek végzik, ami ismételten azt jelenti, hogy nem szakemberek mondják fel a szöveget és nem feltétlenül stúdióköörülmények között, de lényegesen jobb minőségűek, mint a számítógép által generált beszéd. Mindezeket összevetve, ha megtaláljuk a hanganyagokhoz a megfelelő szöveget, akkor együtt hallgatva és olvasva remek segítség lehet a nyelvtanulóknak, az olvasási nehézségekkel küzdőknek.

## Francia digitális könyvtárak

Két nagy francia digitális könyvtár kerül említésre a jelen írásban: a Francia Nemzeti Könyvtár gyűjteménye, a Bibliothèque nationale de France, Gallica<sup>12</sup> (<http://gallica.bnf.fr>) és a Project Gutenberghez hasonlóan, teljesen önkéntes alapon működő Ebooks Libres & Gratuits<sup>13</sup> (<http://www.ebooksgratuits.com>) vállalkozás (2. ábra).

A	B
 <p>Bibliothèque nationale de France, Gallica</p>	 <p>Ebooks Libres &amp; Gratuits</p>

2. ábra Francia digitális könyvtárak. A Bibliothèque nationale de France Gallica digitális könyvtára (A) elsősorban könyvek képként történő online megjelenítését támogatja, míg az Ebooks Libres & Gratuits (B) a könyvek szöveges formátumokban történő tárolását és letöltését teszi lehetővé

### Gallica

A Gallica elsősorban könyvek képformátumban való digitalizálását és szolgáltatását végzi, a képformátumú megjelenítés mellett azonban az utóbbi években bővítette kínálatát és szolgáltat már számítógép-generált szöveges formátumokat is. A szöveges formátum választásakor minden egyes alkalommal megjelenik a figyelmeztető üzenet, melyben tájékoztatják az olvasót arról, hogy megközelítőleg hány százalékos pontossággal történt meg a szöveg felismerése. A Gallica gyűjtemény elsősorban a dokumentumok online megtekintését támogatja és mindössze két alapvetően különböző nézetet ajánl fel: a képként és a szövegeként történő megjelenítést. Mindkét megjelenítési forma esetében lehetőségünk van a szövegen belüli keresésre, mivel a képként történő megjelenítés esetében is a hátsó rétegen megtalálható az OCR-rel felismertett szöveg (l. *E-korlátok fejezet*). Megjegyzném azonban, hogy a szövegen belüli keresés igazából nem a dokumentumon, hanem a teljes gyűjteményen belüli keresést takarja.

E két megjelenítési formán túl egyes művek esetében felkínál letölthető formátumokat, valamint egy megvásárolható verziót is. A letölthető .pdf formá-



tum a nem kereshető, az *E-korlátok fejezetben* álszövegeknek nevezett dokumentumok csoportjába tartozik, amely a könyv lapjait képként jeleníti meg. A felkínált .txt formátum valójában egy .htm kiterjesztésű dokumentum, ami valóban tartalmazza a szöveget, de ezen túl minden két lap találkozásánál megjelenik a szövegben egy „download ModeText.vue.download 8 sur 342” letöltési információ, jelezve, hogy hányadik oldal letöltése következik és hány oldalas a teljes könyv, valamint a soron következő lap fejléce oldalszámmal együtt.

A szövegekhez számítógép-generált hanganyag is elérhető<sup>14</sup>, ami a lapokon található valamennyi karaktert felolvassa, beleértve az oldalszámot, a teljes fejléctet stb., hasonlóan a szövegfájl tartalmához.

### Ebooks Libres & Gratuits

A Gallica gyűjteménnyel szemben az Ebooks Libres & Gratuits gyűjtemény szöveges formátumokban letölthető irodalmi művek megjelenítésére specializálódott. Elsősorban francia nyelvű műveket tartalmaz, de fellelhető a gyűjteményben számos eredeti, nem francia nyelven írt szöveg is. A felkínált fájlformátumok támogatják a személyi számítógépekre, az e-bookolvasókra és a mobil eszközökre való letöltést is. A letöltött szövegek nagyon igényes, a digitalizálást követő utómunkákra vallanak. Hasonlóan a Project Gutenberghez, az Ebooks Libres & Gratuits is szívesen fogadja a felhasználói megjegyzéseket és adományokat. A gyűjtemény további érdekessége, hogy minden egyes találatnál olvasható egy rövid összefoglaló a műről.

### Német digitális könyvtárak

A német nyelvű digitális könyvtárak közül jelen cikkben három kerül említésre: az indulás előtt álló Német Nemzeti Digitális Könyvtár: a Deutsche Digitale Bibliothek<sup>15</sup> (<http://www.deutsche-digitale->

[bibliothek.de](http://bibliothek.de)), a Project Gutenberg német kiadása: a Projekt Gutenberg-DE<sup>16</sup> (<http://gutenberg.spiegel.de>) és a Zeno<sup>17</sup> (<http://www.zeno.org>) (3. ábra).

### Deutsche Digitale Bibliothek

Hosszas várakozás és előkészületek előzik meg a Deutsche Digitale Bibliothek indítását. Többszöri határidő-módosítást követően a nyitás időpontja 2011 végére, majd 2012 elejére toldott. A könyvtár tartalmáról, az elérhetőségi, letölthetőségi lehetőségekről jelen pillanatban nem sokat lehet tudni, mindezzel meg kell várni a tényleges nyitást. Az ígéretetek szerint nemcsak szövegalapú tartalmakkal töltik fel a könyvtárat, hanem szobrok 3D-s megjelenítése, képek, filmek szolgáltatása is várható. Izgatottan várjuk a nyilvános szolgáltatás elindítását.

### Projekt Gutenberg-DE

A Projekt Gutenberg-DE a Spiegel gondozásában jelenik meg, a Spiegelonline Kultur menüjén, vagy a közvetlen linken keresztül érhető el. A legnagyobb német nyelvű elektronikus könyvtár, amely közel 6000 irodalmi művet tartalmaz („die größte elektronische Sammlung Deutschsprachiger Literatur”). A Project Gutenberg német nyelvű kiadásában a digitalizálás nagy részét a fenntartó, a Spiegel Online végzi, de hasonlóan az eredeti angol nyelvű projekthez, fogadják és várják az önkéntesek munkáját, akik segítséget nyújthatnak a digitalizálási és a javítási munkálatokban.

A digitalizált szövegek rendkívül jó minőségű szöveges formátumban elérhetők online, és ez a formátum bárki számára ingyenesen le is tölthető. Ezen túl a teljes gyűjtemény anyaga megvehető CD-n vagy pendrive-on. További tartalmak is megvásárolhatók: archivált anyagok, amelyek nem érhetőek el az ingyenes gyűjteményben, e-book-formátumok, hangoskönyvek.

A	B	C
 <p>Deutsche Digitale Bibliothek Portal für Kultur und Wissenschaft</p>	 <p>Projekt Gutenberg-DE</p>	 <p>Zeno</p>

3. ábra Német nyelvű digitális könyvtárak. A Deutsche Digitale Bibliothek debütálása 2012 első felében várható (A), a Projekt Gutenberg-DE (B) és a Zeno (C) a legnagyobb német digitális könyvtárak közé tartoznak

Összehasonlítva az eredeti Project Gutenberg gyűjteménnyel, azt tapasztalhatjuk, hogy a német Gutenberg gyűjtemény ingyenes szolgáltatásai lényegesen szűkösebbek, mint azt az eredeti gyűjtemény esetén tapasztaltuk.

### Zeno

A német digitális könyvtárak másik népszerű képviselője a Zeno. Ez a gyűjtemény is, hasonlóan a Projekt Gutenberg-DE-hez, azt állítja magáról, hogy a legnagyobb német nyelvű digitális könyvtár („der größten deutschsprachigen Volltext-bibliothek”), és különböző típusú szöveges tartalmakat szolgáltat. Érdekes azonban megjegyezni, hogy a Project Gutenberg gyűjtemények is csak teljes szövegeket digitalizálnak, tehát ezt a szempontot figyelembe véve nincs különbség a két német gyűjtemény között. Annak eldöntését, hogy melyik könyvtár a nagyobb, az olvasóra bizzuk. A Zeno gyűjtemény keletkezési körülményeiről, a gondozásáról bővebb információt a gyűjteményen belül nem sikerült megszerezni, külső hivatkozásként a Wikipédia erre vonatkozó német<sup>18</sup> és angol<sup>19</sup> nyelvű szócikkeit nyújthatnak támpontot. Mindezt figyelembe véve csak a keresési lehetőségek, a találatok minősége és mennyisége az, amely valamilyen támpontot ad a felhasználónak a gyűjteményről.





Az irodalmi művek online megtekinthető szöveges formátumban érhetőek el, amely formátum, hasonlóan a Projekt Gutenberg-DE szövegeihez, le is tölthető. A Zeno gyűjteménynek ez az egyetlen, irodalmi művekhez köthető szolgáltatása. Minden egyéb szolgáltatás a Zenon kívülre mutató linkeken keresztül érhető el.

### Többnyelvű digitális könyvtárak

Egy többnyelvű digitális könyvtár elsődleges célja, hogy egyetlen közös rendszerben szolgáltatson hozzáférést különböző nyelvű tartalmakhoz, a lehetőségek szerint minél többféle felületi nyelven. Jelen írásban négy ilyen jellegű könyvtár kerül említésre: az Europeana<sup>20</sup> (<http://www.europeana.eu>), a Google Books<sup>21</sup> (<http://books.google.com>), az IntraText<sup>22</sup> (<http://www.intratext.com>) és az International Children's Digital Library<sup>23</sup> (<http://en.childrenslibrary.org>) (4. ábra). A négy könyvtár közül három: az Europeana, az IntraText és Children's Digital Library elsődleges célja más népek, kultúrák, nyelvek megismertetése, míg a Google Books főként keresési orientáltságú, elsődleges célja könyvek lelőhelyének felkutatása.

### Europeana

Az Europeana egy internetes portál, amelynek célja, hogy Európa kulturális és tudományos öröksége a nyilvánosság számára elektronikusan vizsgálható legyen. Az Europeana tehát egy felület, egy metaadat-gyűjtemény, egyfajta virtuális könyvtár<sup>7</sup>, amely elérhetőséget biztosít az Európa-szerte digitalizált és tárolt több milliányi könyvhöz, képhez, filmhez, hanganyaghoz, illetve múzeumi tárgyak fotóihoz. Több mint 2200<sup>24</sup> (más források szerint közel 1500<sup>25</sup>) európai intézmény csatlakozott a projekthez, amelynek fő támogatója az Európai Bizottság. A felülethez huszonkilenc különböző nyelven biztosítanak elérhetőséget. Jelen pillanatban azonban ez a szolgáltatás még nem

A	B	C	D
			
Europeana	Google Books	IntraText	International Children's Digital Library

4. ábra Többnyelvű digitális könyvtárak. A bemutatott könyvtárak elsődleges felületi nyelve angol; IP cím szerinti azonosítás alapján, valamint a felhasználó kérésének megfelelően lehetőség van a felületi nyelv átállítására, érdemes azonban szem előtt tartani, hogy a kezelőnyelv módosítása – az esetek többségében – csak részlegesen történik meg



teljes körű, ami annyit takar, hogy még nem készült el a fordítás valamennyi nyelvre, nincs megoldva az oldalak folyamatos fordítása, frissítése sem. Az egyes nyelvekhez, így a magyarhoz is felkínált emberi és gépi fordítás csak részleteiben történik meg. A fordítások eredménye érdekes kombinációja az angol és a magyar nyelvnek, amelyre számtalan példát lehet találni. (Például: „További részletek of an object can be found by clicking on the item in the results list.”<sup>26</sup>, vagy „Data provider: National Széchényi Library - Országos Széchényi Könyvtár | Tartalom hozzáadása: The European Library | Hungary |.”<sup>27</sup>).

Az Europeana négy típusba sorolja a tartalmakat:

- Szöveg: könyvek, levelek, levéltári anyagok, disszertációk, versek, újságcikkek, faksimilik, kéziratok és kották.
- Kép: festmények, rajzok, nyomatok, fényképek, múzeumi tárgyról készült fotók, térképek, grafikus tervek, vázlatok és kották.
- Mozgóképek: filmek, híryananyagok és televíziós programok.
- Hanganyag: zene és hangzó szöveg hengerről, kazettáról, lemezzel, rádióközvetítésekből.

A találat elérhetőségét, formátumát, minőségét minden esetben az eredeti szolgáltatóhely határozza meg, amelyhez a 'View item at' hivatkozáson keresztül juthat el a felhasználó.

Az Europeana felhasználói felületéről, annak kezelhetőségéről megjelent egy felmérés, amelyet német kommunikáció szakos egyetemi hallgatókkal végeztek. Érdeklődő olvasók figyelmébe ajánlom az eredeti német tanulmányt és az ennek alapján készült magyar nyelvű recenziót [13] [14]. Az Europeanaról és a hozzá kapcsolódó projektekről részletes leírást találnak Kovácsné Koreny Ágnes cikkében [15], valamint Kerekes Pál már fentebb említett könyvében [16].

### Google Books

A felsorolt többnyelvű digitális gyűjtemények közül talán a Google Books a legnépszerűbb, ez viszont nem jelenti azt, hogy az itt képviselt kulturális és tudományos értékek, a találatok hasznosíthatósága egyenesen arányos a népszerűséggel. A Google Books elsősorban egy keresőfelület, amelynek elsődleges célja, hogy a felhasználót a könyvek lelőhelyének megtalálásában segítse. Mindemellett végez digitalizálási munkákat is, illetve átvesz kiadóktól és könyvtáraktól digitalizált

anyagokat, amelyeket megjelenít a felületén, ezzel is segítve a felhasználót a tájékozódásban.

Szemben a korábbi fejezetekben felsorolt gyűjteményekkel, a Google Books a felületén megjelenített könyveknél, nem a megszokott szerzői jogi szabályozást veszi alapul annak eldöntésekor, hogy egy könyvet ingyenesen elérhetővé tesz-e a felhasználó számára vagy sem. A kiadási évet veszi figyelembe, és kiadói engedély is szükséges ahhoz, hogy egy mű, amely egyébként már rég nem esik a szerzői jogi szabályozás hatálya alá, teljes egészében elérhető legyen<sup>28</sup>. Ellentétben azonban más digitális könyvtárakkal, a kiadók hozzájárulásával, esetenként a legfrissebben megjelent kiadványok egyes részleteibe is betekintést engedélyez.

A Google Books által megfogalmazott szerzői jogi politika következménye, hogy más, a könyvtárakban ingyenesen hozzáférhető, letölthető, kiváló minőségű digitális könyvek ebben a rendszerben még online módon is csak hiányosan érhetőek el, az alábbi üzenet kíséretében: „Ez egy előnézet. A rendszer korlátozza a megjelenő összes oldal számát. Szerzői jogi védelem alatt álló anyag.”

Ahhoz, hogy a sok-sok töredék könyv közül gyorsan ki tudjuk válogatni a teljes műveket, a Google Books 'Speciális Könyvkereső' vagy a Google 'Speciális keresés' szolgáltatását kell használni. Ebben a keresőben a könyvre vonatkozó adatokon túl be kell állítani két további opciót: a 'Keresés' mezőnél a 'Csak teljes nézet' kapcsolót, a 'Tartalom' mezőnél a 'Könyvek' kapcsolót. (Egy már létező találati lista esetén a lista mellett található „Teljes nézet” gombbal is lehet szűkíteni a találati halmazt.) Sajnálatos módon, a találatok száma nem teljes. A kereső rendszer ilyen jellegű következetlenségére nehéz magyarázatot találni. (Például: *Mark Twain* „The Adventure of Tom Sawyer” című művének keresése esetén a kereső mezőben az „intitle:tom intitle:sawyer inauthor:twain” jelenik meg a speciális kereső kitöltése után. Az így megtalált és online teljes egészében elérhető könyvek száma a Google szerint egy. Az előző keresést azonban tovább bővítve a kiadóval – intitle:tom intitle:sawyer inauthor:twain inpublisher:plain inpublisher:label –, újabb teljes könyvre bukkanhatunk.)

Továbbra is igaz azonban, hogy letölteni ezeket a könyveket nem lehet, mivel „Szerzői jogi védelem alatt álló anyag”, ami a Tom Sawyer esetében

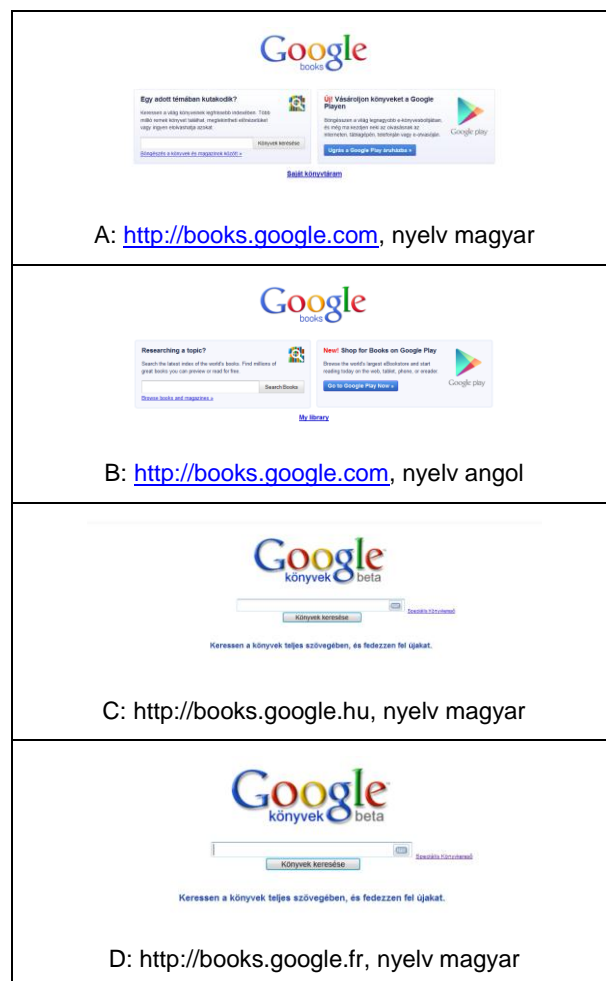
rendkívül meglepő eredmény. Sajnos *Shakespeare*-rel próbálkozva sem lényegesen jobb a helyzet. Egy „Romeo and Juliet” keresést elindítva már találunk is néhány 19. századi kiadást, amely .pdf formátumban letölthető. A gond azonban az, hogy míg a webes felületen a Google-szövegekben lehetséges a keresés, addig a letöltött .pdf formátumokban nem, mivel a lapok képként jelennek meg.

Természetesen ezeknél a találatoknál is a felhasználás módja az, amely nagyban befolyásolja a találat értékét. Ha a felhasználó azzal a szándékkal indította a keresést, hogy megtalálja azokat a könyvtárakat és könyvesboltokat, ahol fellelhetők a „Romeo and Juliet” angol nyelvű példányai, vagy hogy valamilyen digitális formátumban, ám „eredetiben” lásson, olvasson egy 19. századi kiadású Shakespeare kötetet, akkor a legjobb helyen jár. Az *E-korlátok* fejezetben elmondottaknak és a *Gallica* gyűjteményénél tapasztaltaknak megfelelően (*Gallica* fejezet) a Google Books webes felületén is kétszintű tárolás történik, a felületen az olvasó a szkennelt képet látja, míg lehetőség van a szövegen belüli keresésre is. Itt is igaz azonban, hogy nem a dokumentumon, hanem a gyűjteményen belüli keresést támogatja elsősorban a felület.

Ha a felhasználó a számítógépére, e-book-olvasójára letöltve „csak” olvasásra kívánja a szöveget használni, akkor nem szerencsés választás egy majd’ kétszáz éves kiadás kiválasztása, mivel annak minősége nagyban megnehezíti a folyamatos olvasást, és nem szabad megfedkezünk az időközben megváltozott helyesírási és tipográfiai szabályokról, amelyekhez – ha nem faksimile vagy reprint kiadásról van szó – a kiadók rendszerint hozzáigazítják a szöveget. Ezzel szemben, ha szövegen belüli keresésre, számítógépes szövegfeldolgozásra kívánjuk a könyvet felhasználni, akkor sajnos ezek a találatok nem használhatók, mivel a képként történő tárolás nem teszi lehetővé a szövegen belüli kereséseket.

A Google Books teljes körű szolgáltatása, hasonlóan az Europeanához, csak angolul érhető el, míg egyes szolgáltatások már a kiválasztott nyelven, így magyarul is. Hiányoltam azonban, hogy nincs olyan keresési opciója, amely lehetővé teszi az adott nyelven elérhető összes könyv kilistázását, és nem választhatom meg a felületi nyelvet sem. Ez utóbbi hiányosság egyik következménye, hogy csak a böngésző nyelve állítható, amellyel valamelyest szabályozható a megjelenítendő lap. Magyar

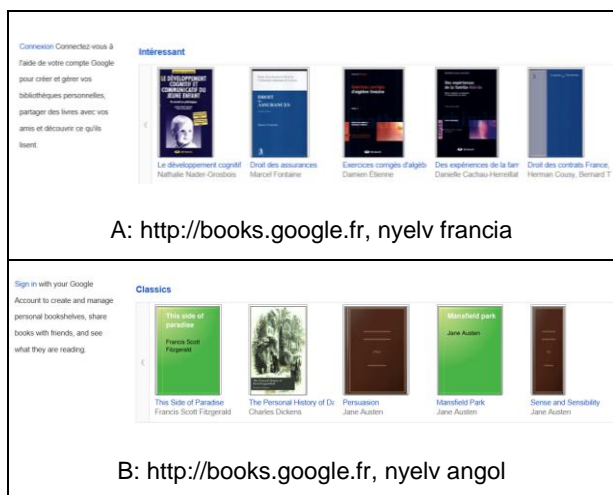
nyelvű böngészőt használva a .com végződésű doménnév esetén az angol felület magyar fordítása (5. ábra A), míg az országgóddal végződő doménnevek esetén a magyar nyitólap (5. ábra C) jelenik meg, függetlenül az országgódtól (5. ábra D). A nyelvek nem megfelelő kezelésének következménye, hogy teljesen kiszámíthatatlan, hogy mikor milyen lap jelenik meg. Példaként érdemes összehasonlítani az .fr végződésű doméint a böngésző nyelvének magyar (6. ábra D), francia (6. ábra A) és angol (6. ábra B) nyelvre állításával.



5. ábra A Google Books nyitólapja különböző végződésű doménnevek és különböző nyelvek esetén

A *Többnyelvű digitális könyvtárak* fejezetben mondtaknak megfelelően a felsorolt digitális könyvtárak közül a Google Books az egyetlen kereső orientált könyvtár. A könyvtár elsődleges célja tehát egy kereső felület biztosítása, amely a könyvek elérhetőségéről tájékoztat, a könyvpiaci, valamint a könyvtári lehetőségek bemutatásával, ki-

egészítve ezt a szövegek online elérhetőségével és esetleges letöltésével. E céljainak maximálisan eleget tesz, és mindemellett elérhetőséget nyújt, még ha csak töredékeiben is, és olykor gyenge minőségben, olyan tartalmakhoz, amelyek mind a hagyományos értelemben vett olvasóközönség számára, mind a keresgélők számára értékes találatokat eredményeznek. További információkért érdemes elolvasni *Bodó Balázs* [17] és *Kerekes Pál* írásait, amelyek részletesen foglalkoznak a „Google Books-jelenséggel” [18].



6. ábra A Google Books nyitólapja .fr végződésű doménnév esetén, a böngésző francia (fent) és angol (lent) nyelvre állításakor

## IntraText

Az IntraText digitális könyvtár igazi különlegesség, ami a szövegek egyedi feldolgozási módját illeti. A szolgáltatás sajátossága, hogy a szövegek konkordancialinkekkel jelennek meg (7. ábra). Egy ilyen hivatkozásra kattintva megjelenik a szóalak teljes, szövegen belül értelmezett konkordancialistája (7. ábra). A lista, hasonlóan a szöveghez, konkordancialinkekkel együtt megmutatja a szóalakokat. A konkordancialisták előállításán és megjelenítésén túl a rendszer statisztikákat is készít. Minden egyes szóalakhhoz megadja a szövegen belüli előfordulási gyakoriságot, valamint a szóalak – szintén szövegen belüli – ábécében és a gyakorisági listában elfoglalt helyét. A szövegről, a szövegben előforduló szóalakokról további statisztikai adatokat és listákat is megjelenít a program: a szöveg hossza, a különböző szóalakok száma, az átlagos ismétlődési ráta, a szóalakok átlagos hosszúsága, a szóalakok szókezdő és szöveg szerinti sorrendje, gyakorisági lista, gyakorisági, kezdő- és utolsó betű szerinti csoportok stb.

Az angol mellett hat másik felületi nyelven is elérhető a szolgáltatás, és ezek között szerepel a magyar nyelv is<sup>29</sup>. Igen gazdag a nyelvkészlete, mivel 40 különböző nyelven kínál a könyvtár szövegeket. A legkorábbi könyv i.e. 900 körülre tehető, míg szerzői hozzájárulással szép számban találni 21. századi műveket is.

**KÉK PILLE**

Van a méhesemben egy szalmadívány. Mikor a napóra igen fekete vonással mutatja a délutáni időt, oda szoktam leheverni és olvasgatni, vagy szenderegni. Mert olvasni is jó, meg aludni is jó. Mikor olvasok, úgy érzem, mintha egy másik embernek a lelke volna bennem, és ez gyönvörűség; mikor pedig alszom, úgy érzem, mintha az én lelkem volna egy másik emberben; és ez is gyönvörűség.

1		<u>érzem, mintha egy másik embernek a lelke volna bennem, és</u>
2		<u>halkan egy mellette ülő subás embernek -, hallottam már a múlt</u>
3		<u>zörgött. Minek fusson? Falusi embernek nem sietős az útja, legfeljebb</u>
4		<u>Mátyás - mondom az embernek -, mennyiért adná el ezt</u>
5		<u>azok a betegségei, amik az embernek. Egy test, egy vér, egyazon</u>
6		<u>teljes odaadásával. A felnőt embernek már borba kell vesztenie</u>
7		<u>beszélhet az Éghez! Én pápista embernek születtem és szeretem is</u>
8		<u>megértetni, hogy némelyik embernek az orcája miért oly szép,</u>
9		<u>Kihozat egy pohár bort az embernek.~- No - azt mondja -, hát</u>
10		<u>csöndességben eléje jön az embernek egy leány - akármiféle leány -,</u>
11		<u>szép része a mezitláb járó embernek, de festők se figyelmeznék</u>
12		<u>ebből láthatod, hogy az embernek, szabad akarata nincsen.</u>
13		<u>falusi emberek. Szegény embernek pedig nem dukál parádé a</u>

7. ábra Az IntraText digitális könyvtár a szövegek konkordancialinkekkel történő megjelenítését nyújtja (fent); a linkkel ellátott szavakra kattintva megjelenik a kiválasztott szóalak szövegen belüli konkordancialistája (lent)

A könyvtár különlegességét tehát az adja, hogy a konkordancialinkek és -listákon, valamint a szövegből nyert szóalakokra vonatkozó statisztikai adatokon keresztül lehetőséget ad a kiválasztott szöveg nyelvi elemzésére. Az IntraText vitathatatlanul bőséges forrást jelent azoknak a nyelvészeti kutatásokat végzőknek, akik szövegek számítógéppel támogatott nyelvi elemzésével foglalkoznak. Ezen túl az IntraText nyújtotta szolgáltatásokat nagyon jól fel lehet használni mind az anyanyelvi, mind az idegen nyelvű oktatásban, ha a tankönyvi mesterséges és szemiautentikus szövegeken túl igény mutatkozik irodalmi műveken keresztül történő nyelvtanulásra.

### **International Children's Digital Library**

Ha az IntraTextet igazi különlegességnek neveztük, akkor az International Children's Digital Library-t (ICDL), „A Library for the World's Children” könyvtárat nyugodtan lehet igazi gyöngyszemnek tekinteni. Mind filozófiájában, mind megjelenésében olyan értékek átadását vállalja, amelyek elsődleges célja a különböző népek és nemzetek egymás közötti kapcsolatainak erősítése, egymás elfogadása, az egymás iránti tolerancia erősítése<sup>30</sup>, mindezt egy nagyon barátságos, gyerekeknek szóló felületen. 2010-ben ez a könyvtár nyerte el az Award for International Library Innovation díjat<sup>31</sup>.

A könyvtárat elsősorban 3–13 éves gyerekeknek, valamint szülőknek, pedagógusoknak, könyvtárosoknak és más, gyerekekkel foglalkozó felnőtteknek hozták létre, azzal a céllal, hogy be tudják mutatni az egyes kultúrák, társadalmak, népek, életstílusok és életvitelük közötti hasonlóságokat és eltéréseket. A könyvtár olyan könyvek gyűjteménybe foglalását vállalta, amelyek segítik megérteni a gyerekeknek az őket körülvevő dolgokat, valamint a világot globális valójában. További célja a könyvtárnak, hogy kiszolgálja azokat a tudományos munkát végző szakembereket, akiknek a gyermekirodalom a kutatási területe<sup>32</sup>.

Az ICIDL tizenkilenc különböző nyelven érhető el és hatvanegy nyelven<sup>33</sup> kínál könyveket, amelyek online olvashatók. Olyan könyvek gyűjteménye, amelyek már nyomtatott formában megjelentek és utólag történt meg digitalizálásuk. Szemben sok más, ingyenesen elérhető digitális könyvtárral, ahol csak olyan könyveket tesznek elérhetővé, amelyek már nem esnek a szerzői jogi szabályozás hatálya alá (vagy még azokat sem), ebben a könyvtárban a szerző, illetve a tulajdonos hozzájárulásával bármilyen keletkezési dátumú könyv

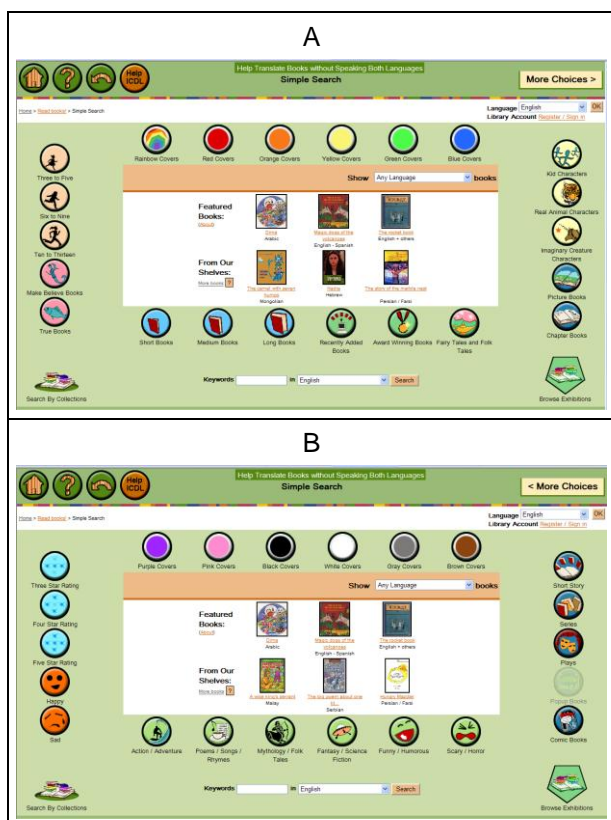
elérhető. A könyvtár egyik célja, hogy a gyerekek hozzájussanak a korábbi korszakok történelmileg jelentős olvasmányain túl a kortárs irodalomhoz is. A könyveket teljes egészében jelentetik meg, soha nem csonkított formában, ez alól kivételek azok a könyvek, amelyeknek jelent meg rövidített, vagy valamilyen más adaptációja nyomtatott formában. Töreksenek arra, hogy egy könyvnek a különböző idegen nyelvű fordításai is bekerüljenek a könyvtárba. Mindezen szempontokon túl igyekeznek olyan könyveket is megszerezni és digitalizálni, amelyek valamilyen díjat nyertek a világ bármely pontján, elfogadva az ottani értékelési rendszert.

Az ICIDL egyszerű keresőjének (Simple Search) kezelőfelülete megjelenésében és kezelésében is rendkívül gyerekbarát, elsősorban a tematikus keresést támogatja, de lehetőség van 80 különböző nyelven kulcsszavas keresésre is. A következő nagy témaköröket ajánlja az egyszerű kereső első, nyitó lapja (8. ábra A): három különböző kategória életkor szerint, két kategóriába szedve az igaz, illetve a kitalált történetek, három különböző kategória terjedelem szerint, újonnan a gyűjteménybe került könyvek, valamilyen díjat kapott könyvek, mesék és népmesék, három kategóriába csoportosítva a gyerekekről, állatokról és kitalált lényekről szóló könyvek, hét kategóriába sorolva a képes és a szöveges könyvek, s a legfontosabb, ami a felsorolásban a végére maradt, de a felületen a legelőkelőbb helyet foglalja el: hat kategóriába sorolva a borító színe. Ugyanazon tartalmak, ám más szempont szerint csoportosítva az egyszerű kereső második lapjáról is elérhetőek (8. ábra B). Nagyon ritkán látott, ugyanakkor rendkívül felhasználóbarát szolgáltatása ez a könyvtárnak. Elérhető továbbá az egyszerű kereső mindkét lapjáról a könyvtár összetett keresője a 'Search by Collections' gombon keresztül is (8. ábra A és B alsó sarok).

A 'Browse Exhibitions' (8. ábra A és B jobb alsó sarok) gombra kattintva egy olyan lapra jut az olvasó, ahol hasonló témájú könyveket gyűjtöttek kategóriákba, kiállításokba. Minden egyes kiállítás tartalmazza a téma egy összefoglalóját, hivatkozásokat a kiállításon szereplő könyvekhez, egy leírást, amely megmondja, hogy az adott könyv miért került a kiállításhoz, és végül foglalkoztató feladatokat. Jól átgondolt, nagyon értékes gyűjteményt állítottak össze<sup>34</sup>.

Az ICIDL összetett keresője is egy gyerekbarát felületen keresztül, jelölő négyzetek segítségével, a lehetőségekhez képest interaktív módon kínálja a különböző szempontok alapján történő keresést.

Az ICDL-találatok csak online érhetőek el, letöltési lehetőség nincs. Ha az adatátviteli sebesség nem megfelelő, és a betöltés nem történik meg azonnal, a várakozás idejére a képernyőn megjelenik egy igen barátságos százlábú, amely közli a felhasználóval, hogy a betöltés folyamatban van, várjunk türelmesen. A kiválasztott könyv kicsinyített lapjai jelennek meg, ezek navigációs gombként működnek. Rendkívül igényesen, laponként, képként szkennelt digitalizált könyvekhez férünk hozzá, amelyek az olvasó igényeinek megfelelően különböző nagyságokban is megnézhetőek. A lapok nagyításán túl gyakran a szöveges részek külön is kinagyíthatók. Ha a könyvtárban létezik a kiválasztott mű fordítása, az eredeti nyelv lecserélhető a fordítás nyelvére, és ily módon a könyvtár anyaga kiválóan alkalmazható nyelvtanuláshoz. A fordításokhoz azonban már nem feltétlenül kapcsolódnak a felsorolt kényelmi szolgáltatások.



8. ábra Az International Children's Digital Library egyszerű keresőjének két lapja (alapértelmezett lap, A<sup>35</sup>; második lap, B<sup>36</sup>); az egyszerű kereső elsősorban a tematikus keresést támogatja, amely jó néhány, a gyerekek által preferált kategóriát sorol fel; a különböző kategóriák beszédes, a gyerekek számára könnyen kezelhető gombokon keresztül érhetőek el

## Összegzés

Az idegen nyelvű digitális könyvtárak összeválogatásával az volt az elsődleges célunk, hogy a magyar olvasó számára bemutassuk, milyen lehetőségek nyílnak más népek, kultúrák megismerésére, hagyományos könyvtárak használatánál jelentkező fizikai határok megszűnésének következtében hogyan lehet ezeket a könyvtári tartalmakat a hagyományos értelemben vett olvasáshoz, illetve a nyelvtanulásban felhasználni, továbbá hogyan tudják a tudományos kutatásokat végző szakemberek kihasználni a könyvtárak szolgáltatásait. Mivel ennek az írásnak a kereteit meghaladja, nem volt célunk a könyvtárak kategorizálása, rangsorolása. Az elsődleges cél a kiválasztott könyvtárak könyv- és szövegtartalmának felhasználói oldalról való, gyakorlati szempontokat figyelembe vevő megközelítése volt, külön hangsúlyt fektetve az állományok különböző formátumaira és azok elérhetőségére.

A könyvtárak nagy száma miatt mindenképpen indokolt volt a válogatás, mivel egyetlen írás keretében nincs lehetőség az idegen nyelvű könyvtárak teljes körű bemutatására. A cikkben említett és/vagy bemutatott digitális könyvtárakra azért esett a választás, mert egy, vagy akár több olyan jellemzővel is rendelkeznek, amelyek miatt figyelemre érdemesek. Lehet ez az elérhetőség, a méret, a többnyelvűség, a fájlformátumok minősége és száma, a megjelenítés minősége és módja, a könyvtár ismertsége, népszerűsége, a könyvtár által képviselt értékek. Célunk volt továbbá az is, hogy rávilágítsunk a kiválasztott könyvtárak előnyeire, megmutatva, hogy miért érdemes ezeket látogatni. Esetenként, ha ez indokolt volt, az adott könyvtár hiányosságai is említésre kerültek. Érdemes azonban szem előtt tartanunk, hogy a hagyományos értelemben vett fizikai korlátok megszűnésével az olvasónak bármikor lehetősége van egy másik gyűjtemény meglátogatására, vagy a kiválasztott gyűjtemény egy későbbi időpontban történő ismételt látogatására, annak reményében, hogy a lehető legrövidebb időn belül megfelelő találatot jussunk.

A digitális könyvtárak sajátosságaiból adódik, hogy könnyű navigálni a könyvtárak között, amit még tovább egyszerűsíthet egy jól kialakított kezelőfelület, amelyre az Europeanat láttuk példának. További jellemzője ezeknek a könyvtáraknak a szolgáltatások folyamatos bővülése. A tartalmi gyarapodáson túl a szolgáltatások bővülése is nagyban befolyásolhatja a könyvtár használhatóságát. Az



egyik legfontosabb szempont lehet a hozzáférés módja és a felkínált formátumok száma. Az utóbbi évek nagy nyertese a Gallica, amely a korábbi kezdetleges, nehezen elérhető felületét, nehézkesen kereshető webhelyét lecserélte egy jól működő online felületre. A lehetséges fájlformátumok bővítésében a Project Gutenberg jár az élen, amely az elektronikus olvasó eszközök megjelenéséhez igazodva a korábbi .html és .txt fájlokat kibővítette az új formátumokkal. Hasonlóan, az új e-olvasó-eszközökhöz igazodva, az ICDL is kialakította az iPad és az iPhone eszközökön történő olvasás lehetőségét, mivel ennek a könyvtárnak a tartalma csak online érhető el. Új könyvtárak megjelenése leginkább a speciális igényeket kielégítő témákban, szolgáltatásokban valószínűsíthető, de láthatuk a német példán keresztül, hogy általános, nagy, nemzeti könyvtárak megjelenése is várható még az elkövetkező időszakokban.

A digitális könyvtárak megjelenésével egyes fizikai korlátok megszűntek, viszont keletkeztek újak, amelyeket nem lehet figyelmen kívül hagyni a szolgáltatások értékelésekor. Ilyen újkori fizikai korlát lehet a digitális írástudás hiánya, a fájlok elérhetőségének módja, egyes formátumok hiánya. Nem szabad tehát elfelejtenünk, hogy a digitális könyvtárak és könyvek értékét jelentősen befolyásolják a felhasználó lehetőségei és céljai. A felhasználó keresési szándékai, eredeti célkitűzései azok a kiindulási feltételek, amelyek eldöntik, hogy a találatok értékesnek tekinthetők-e vagy sem. A találatok azonban csak akkor tekinthetők teljes értékűnek, ha a digitális könyvtár által nyújtott szolgáltatások lehetővé teszik a szükséges hozzáférést.

## Hivatkozások

- 1 VERNE, J.: Around the World in Eighty Days. Audio Book, computer-generated. <http://www.gutenberg.org/files/8982/mp3/8982-002.mp3>. Feltöltés: 2005. szeptember 1. Letöltve: 2011. november 1.
- 2 Aesop's Fables Volume 1 The Goose That Laid The Golden Eggs. Audio Book, human-read. <http://www.gutenberg.org/files/19616/mp3/19616-02.mp3>. Feltöltés: 2006. október 26. Letöltve: 2011. november 1.
- 3 GÁRDONYI Géza: Az én falum. Kék pille. MVGYOSZ hangoskönyvek. <http://mek.oszk.hu/03000/03082/mp3/1691af05.mp3>. MEK-be került: 2005. szeptember 1. Letöltve: 2011. november 1.
- 4 Project Gutenberg. <http://www.gutenberg.org/>. Módosítva: 2011. szeptember 24. Letöltve: 2011. november 1.
- 5 Project Gutenberg Australia. <http://www.gutenberg.net.au/>. Módosítva: 2011. január 5. Letöltve: 2011. november 1.
- 6 ManyBooks. <http://www.manybooks.net/>. Letöltve: 2011. november 1.
- 7 Partners, Affiliates and Resources [http://www.gutenberg.org/wiki/Gutenberg:Partners%2C\\_Affiliates\\_and\\_Resources](http://www.gutenberg.org/wiki/Gutenberg:Partners%2C_Affiliates_and_Resources). Módosítva: 2011. július 12. Letöltve: 2011. november 1.
- 8 Classic Bookshelf. <http://www.classicbookshelf.com/>. Letöltve: 2011. november 1.
- 9 Aesop's Fables. <http://www.aesopfables.com/>. Módosítva: 2011. március 28. Letöltve: 2011. november 1.
- 10 The CD and DVD Project From Project Gutenberg, the first producer of free ebooks. [http://www.gutenberg.org/wiki/Gutenberg:The\\_CD\\_and\\_DVD\\_Project](http://www.gutenberg.org/wiki/Gutenberg:The_CD_and_DVD_Project). Letöltve: 2012. május 21.
- 11 Free eBooks by Project Gutenberg. <http://www.gutenberg.org/>. Letöltve: 2012. május 21.
- 12 Bibliothèque nationale de France, Gallica. <http://gallica.bnf.fr/>. Letöltve: 2011. november 1.
- 13 Ebooks Libres & Gratuits. <http://www.ebooksgratuits.com/ebooks.php>. Letöltve: 2011. november 1.
- 14 VERNE, J.: Le tour du monde en quatre-vingt jours (63e édition). <http://gallica.bnf.fr/ark:/12148/bpt6k64707h/f8.vocal.r=.langEN>. Letöltve: 2011. november 1.
- 15 Deutsche Digitale Bibliothek. Portal für Kultur und Wissenschaft. <http://www.deutsche-digitale-bibliothek.de/index.htm>. Letöltve: 2011. november 1.
- 16 Projekt Gutenberg-DE. <http://gutenberg.spiegel.de/>. Letöltve: 2011. november 1.
- 17 Zeno, der größten deutschsprachigen Volltextbibliothek. <http://www.zeno.org/>. Letöltve: 2011. november 1.
- 18 Zeno.org. <http://de.wikipedia.org/wiki/Zeno.org>. Letöltve: 2012. május 21.
- 19 Zeno.org. <http://en.wikipedia.org/wiki/Zeno.org>. Letöltve: 2012. május 21.
- 20 Europeana. <http://www.europeana.eu/portal/>. Letöltve: 2011. november 1.



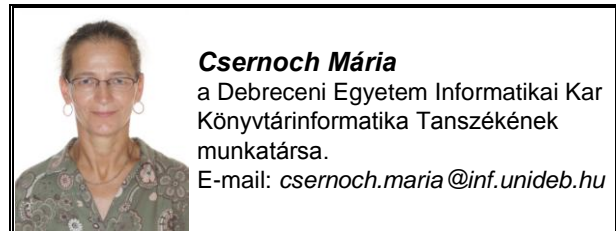
- <sup>21</sup> Google Books. <http://books.google.com/>. Letöltve: 2011. november 1.
- <sup>22</sup> IntraText. <http://www.intratext.com/>. Módosítva: 2011. november 10. Letöltve: 2011. november 11.
- <sup>23</sup> International Children's Digital Library. <http://en.childrenslibrary.org/index.shtml>. Letöltve: 2011. november 1.
- <sup>24</sup> Facts & Figures. Europeana Professional knowledge-sharing platform. <http://pro.europeana.eu/web/quest/about/facts-figures>. Letöltve: 2012. május 21.
- <sup>25</sup> Europeana: think culture. <http://europeana.eu/portal/aboutus.html>. Letöltve: 2012. május 21.
- <sup>26</sup> Találatok az Europeana-ban. [http://europeana.eu/portal/usingeuropeana\\_results.html](http://europeana.eu/portal/usingeuropeana_results.html). Letöltve: 2012. május 21.
- <sup>27</sup> Tom Sawyer kalandjai. <http://www.europeana.eu/portal/record/92006/03192429733FFC6DE8A9A5ED2584869FCD5C6882.html?start=2&query=mark+twain+tom+sawyer>. Letöltve: 2012. május 21.
- <sup>28</sup> Please note that many classic novels whose original editions are in the public domain have been republished, often with additional content, and we treat these newer editions as copyrighted works. Google respects and protects the copyrights of all publishers, and in-copyright editions of these works are not available for download: Why is the Download button only available on some books? <http://books.google.com/support/bin/answer.py?hl=en&answer=44666>. Letöltve: 2011. november 1.  
Mit lát a képernyőn, amikor a Google Könyvkeresőben keres. <http://www.google.com/googlebooks/screenshots.html>. Letöltve: 2011. november 1.
- <sup>29</sup> Catalogue by Language. Magyar. <http://www.intratext.com/HUN/>. Az IntraText formátum. <http://www.intratext.com/Aiuto/HUN/>. Módosítva: 2011. november 10. Letöltve: 2011. november 11.
- <sup>30</sup> Mission statement: The mission of the International Children's Digital Library Foundation (ICDL Foundation) is to support the world's children in becoming effective members of the global community - who exhibit tolerance and respect for diverse cultures, languages and ideas -- by making the best in children's literature available online free of charge. The Foundation pursues its vision by building a digital library of outstanding children's books from around the world and supporting communities of children and adults in exploring and using this literature through innovative technology designed in close partnership with children for children.
- <sup>31</sup> The International Children's Digital Library won the American Library Association President's 2010 Award for International Library Innovation. The ICDL was also named one of 25 Best Websites for Teaching and Learning by the American Association of School Librarians. International Children's Digital Library. Mission Statement. <http://en.childrenslibrary.org/about/mission.shtml>. Letöltve: 2011. november 11.
- <sup>32</sup> International Children's Digital Library. Collection development policy. Version 8.2. <http://en.childrenslibrary.org/about/collection.shtml>. Módosítva: 2005. március 7. Letöltve: 2011. november 11.
- <sup>33</sup> Library Fast Facts. International Children's Digital Library. <http://en.childrenslibrary.org/about/fastfacts.shtml>. Letöltve: 2012. május 21.
- <sup>34</sup> International Children's Digital Library. Exhibitions. <http://www.childrenslibrary.org/icdl/ExhibitionList?ilang=English>. Letöltve: 2011. november 11.
- <sup>35</sup> International Children's Digital Library. Simple Search 1. lap. <http://www.childrenslibrary.org/icdl/SimpleSearchCategory?ilang=English>. Letöltve: 2011. november 11.
- <sup>36</sup> International Children's Digital Library. Simple Search 2. lap. <http://www.childrenslibrary.org/icdl/SimpleSearchCategory?ids=&langid=&pnum=1&cnum=2&text=&lang=English&ilang=English>. Letöltve: 2011. november 11.

## Irodalom

- [1] KOEHLER, W.: Digital libraries, digital containers, „library patrons”, and visions for the future. Emerald Group Publishing Limited, 22. köt. 5. sz. 2004. p. 401–407. ISSN 0264-0473 DOI 10.1108/02640470410561910.
- [2] KERÉKES Pál: E-book kalauz. Az elektronikus könyv kisenciklopédiája. Budapest, Kossuth Kiadó, 2011. Az elektronikus könyv formátumai. p. 110.
- [3] KERÉKES Pál i. m. p. 37.
- [4] ESHET, Y.: Digital literacy: A new terminology framework and its application to the design of meaningful technology-based learning environments. = ED-MEDIA 2002 Work Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (14th, Denver, Colorado, June 24–29, 2002). <http://www.eric.ed.gov/PDFS/ED477005.pdf>. Letöltve: 2011. november 11.
- [5] JUKES, I. – DOSAJ, A.: Understanding Digital Children (DKs). Teaching and Learning in the New

- Digital Landscape. Prepared for the Singapore MOE Mass Lecture, 2006.  
<http://edorigami.wikispaces.com/file/view/Jukes+-+Understanding+Digital+Kids.pdf>. Letöltve: 2011. november 11.
- [6] KRESS, G.: Literacy in the new media age, 2003. p. 5–8.  
[http://seminar.net/index.php?option=com\\_content&ask=view&id=39&Itemid=1](http://seminar.net/index.php?option=com_content&ask=view&id=39&Itemid=1). Letöltve: 2011. november 11.
- [7] KERÉKES Pál i. m. p. 37.
- [8] KOLTAY Tibor – ÖRDÖGNÉ KOVÁCS Mónika: Digitális könyvtári témák a könyvtárosképzésben: A képzés Jászberényben és Szombathelyen. = Könyvtári Figyelő, 52. köt. 4. sz. 2006. p. 485–496.  
<http://ki.oszk.hu/kf/kfarchiv/2006/4/koltay.html>. Letöltve: 2011. november 1.
- [9] CHOWDHURY, G. – CHOWDHURY, S.: Bevezetés a digitális könyvtárak világába. Introduction to digital libraries. London, Facet, 2003.
- [10] KOLTAY Tibor: Bevezetés a digitális könyvtárak világába. = Könyvtári Figyelő, 49. köt. 4. sz. 2003.  
<http://ki.oszk.hu/kf/kfarchiv/2003/4/koltay3.html>. Letöltve: 2011. november 1.
- [11] KERÉKES Pál i. m.
- [12] LEBERT, M.: A Short History of eBooks., Toronto, University of Toronto, 2009.  
<http://www.gutenberg.org/ebooks/29801>. Letöltve: 2012. március 24.
- [13] SCHWEIBENZ, W.: Eine erste Evaluation de Europeana: Wie Benutzer das „Look & Feel“ des Prototypen der Europäischen Digitalen Bibliothek beurteilen. = Information Wissenschaft und Praxis, 61. köt. 5. sz. 2010. p. 277–284.
- [14] SZIVI Zsófia: Az Europeana felhasználói értékelése – az első tapasztalatok. = TMT, 58. köt. 5. sz. 2011. p. 230–231.
- [15] KOVÁCSNÉ KORENY Ágnes: Europeana Libraries – egy új európai projekt. = TMT, 58. köt. 7. sz. 2011. p. 283–299.
- [16] KERÉKES Pál i. m. p. 78.
- [17] A Google Books és a magyar könyvpiac.  
<http://prezi.com/v3yggkwst5qz/google-books-es-a-magyar-konyvpiac/>. Letöltve: 2012. március 24.
- [18] KERÉKES Pál i. m. p. 140.

Beérkezett: 2012. VI. 5-én.



## Gábor Dénes-díj, 2012. Pályázati felhívás

A korábbi évekhez hasonlóan, ez évben is meghirdetésre kerül a Gábor Dénes-díj, mely a civil szféra egyik legnevesebb műszaki alkotói elismerése ma Magyarországon. A díjjal ez évtől már nemcsak a hazai műszaki és természettudományi felsőoktatás képviselőit, a jelentősebb ágazatok, illetve iparágak (energetika, távközlés/információs technológiák, gépipar/járműipar, számítástechnika, biotechnológia/gyógyszeripar, mezőgazdaság/környezetvédelem, stb.) kutató-fejlesztő szakembereit kívánjuk elismerni és további alkotó munkára ösztönözni, hanem első ízben a határainkon túl élő magyar származású szakembereket is, akik a villamosenergia-ipar területén, illetve az ezzel szimbiózisban lévő diszciplínákban az elmúlt 5 évben, kiemelkedően eredményes mérnöki munkát folytattak, illetve kiemelkedő kutatói teljesítményt nyújtottak, vagy eredeti felfedezést tettek. Ennek értelmében 2012 decemberében, Gábor Dénes munkásságának és humanista gondolkodásmódjának szellemében, ismét ünnepélyes keretek között a megszokottnál több kategóriában kerül átadásra a pénzdíjjal is járó elismerés. A felterjesztési felhívás és az egyes kategóriákra vonatkozó kiírások, valamint háttéranyagok a

<http://www.novofer.hu/alapitvany/tartalom/menu/76>

weboldalon érhetőek el.

A felterjesztések leadási **határideje: 2012. október 10.**

Dr. Gyulai József, a kuratórium elnöke

Bárány Emese, sajtómunkatárs