

Országgyűlési dokumentumok az interneten

Az Országgyűlési Könyvtár adatbázisa

Az Országgyűlési Könyvtár 2008. június 11-én, sajtóbemutatón ismertette az Országgház Delegációs termében az elmúlt évben végzett digitalizálási tevékenységének újabb jelentős eredményét, az *Országgyűlési dokumentumok (képviselőházi naplók és irományok, főrendiházi naplók 1861–1918)* című adatbázisát.

Az adatbázis létrehozataláig szövevényes út vezetett. Néhány fontosabb mérőföldkő bemutatása bepillantást enged az Országgyűlési Könyvtár digitalizálási és ezzel összefüggő állományvédelmi tevékenységébe. Meglepő módon ugyanis a digitalizálás megkezdéséhez az állományvédelmen keresztül vezetett az út. A könyvtár Magyar parlamenti gyűjteményében féltve őrzött dokumentumok állapota folyamatosan romlott, ezért e folyamat megállítása, de legalábbis lassítása a könyvtár alapvető feladatává vált.

Az első lépések

Az úton az első lépés a Magyar parlamenti gyűjtemény állományvédelmének szisztematikussá tétele volt. A '80-as évek elején megkezdtük a bőrkötésű kötetek vegyszeres kezelését. A szükséges eszközök és módszerek kidolgozásában az Országos Széchényi Könyvtár kiváló restaurátor szakemberei voltak segítségünkre. Máig jól emlékszem arra, amikor Kastaly Beatrix az Országos Széchényi Könyvtár Hold utcai műhelyében elmagyarázta a bőrkötésű kötetek kezelésének titkait. Úgy éreztem magam itt, mint egy „boszorkánykonyhában”. Különböző anyagokból, pl. pataolajból és méhviaszból főztünk viaszos bőrpasztát és tisztító likkert. Ezekkel az anyagokkal azután éveken át kezeltük a köteteket. A bőrkötések külső védelme sikeres volt, azonban számolnunk kellett a kötetek belsejében zajló kémiai folyamatokkal, a papírlapok savasodásával is. Ennek oka az, hogy 1840 után az addigi manufakturális papírkészítést felváltotta a nagyipari papírgyártás. A tömegigény miatt a korábbi rongyalapanyagot a facsiszolat követte nyersanyagként. Az új alapanyagú papírban zajló – napjaink egyik legnagyobb állományvédelmi kihívását jelentő – kémiai folyamatok (savas lebomlás) a lapok rövid rostjait felbontják. Ennek jele, hogy a papírlap egyre sötétedő sárgás színezetet kap, és töredezni kezd. A faalapú papírból készült könyvek és iratok kémiai stabilizálás (savtalanítás) és hideg raktározás nélkül rövid idő alatt akkor is szétporladnak, ha kíméletesen kezelik azokat. A helyzetet tovább rontotta a fénymásolók használatának rohamos terjedése. Az olvasói szokások megváltozásával a kézi jegyzetelést felváltotta a gépi fénymásolatok készítése. A másolás közben fellépő hőhatás és

a másolással együtt járó fokozott fizikai igénybevétel tovább rontott a kötetek állapotán. Sajnos, Magyarországon még nem megoldott a dokumentumok tömeges savtalanítása. Ezért komolyan felmerült a lehetősége annak, hogy változatlan körülmények között néhány évtized múlva a Magyar parlamenti gyűjtemény dokumentumállománya helyrehozhatatlan károsodást szenvedhet.

A probléma kezelésében jelentős változást hozott, hogy a technika fejlődése lehetővé tette, a nemzetközi könyvtári gyakorlat pedig alkalmazni kezdte a digitalizálást mint a dokumentumok megóvásának egyik lehetséges módját.

Országgyűlési dokumentumok

A digitalizálási tevékenység bemutatása előtt röviden ismertetem azoknak az országgyűlési dokumentumoknak a körét, amelyeknek szkennelését és feldolgozását tervezzük.

Táblázat
a Magyar parlamenti gyűjtemény dokumentumairól

	Országgyűlési nyomtatványok			Egyéb dokumentumok			
Képviselőház	Naplók	Irományok	Jegyzőkönyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Főrendiház/Felsőház	Naplók	Irományok	Jegyzőkönyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Delegáció	Naplók	Irományok	Jegyzőkönyvek	Mutató	Ügyrend		

A fenti táblázat áttekintést ad az országgyűlési dokumentumok főbb típusairól. A képviselőház és a főrendiház – valamint az országgyűlés egy különleges, közös ügyek tárgyalására létrehozott bizottsága, az ún. delegáció – működését és eredményességét az alábbi sorozatok tükrözik:

- az ülésekről – a parlamenti gyorsíróiroda által magas színvonalon – készített szó szerinti *Naplók*,
- az üléseken hozott határozatok rövid és szabatos leírását rögzítő *Jegyzőkönyvek*,
- a tárgyalásokhoz benyújtott előzetes írásos dokumentumok, az *Irományok*,
- a képviselők életrajzi adatait tartalmazó az *Almanachok*,
- a képviselők elérhetőségéről tájékoztató *Lakcímjegyzékek (Lakáskönyvek)*.

A fenti dokumentumok információs értékének megértéséhez érdemes egy pillantást vetnünk kialakulásuk történetére. Az országgyűlés működésének jelentősége – az írott jog szerepe és az annak alkotásában való részvétel, valamint a politikai kulisszatitkokat és folyamatokat megismerni és megérteni vágyó közvélemény miatt – a képviseleti parlamenti demokráciákban, a polgári (jog)államokban ugrásszerűen megnőtt. Ezért már a XVIII. század végétől az ország uralkodó körei is felismerték, hogy a személyes (át- vagy félre)értelmezést megelőzhetik, ha az országgyűlésen elhangzottakat és a meghozott határozatokat hitelesített, hivatalos formában közzé teszik. Ez a szándék a politikai történések iránt egyre nagyobb

érdeklődést támaztó közvélemény akaratával egybeesett, így nyert teret a törvényhozási tevékenység (sajtó)nyilvánossága. Az események és történések hű lejegyzése érdekében fokozatosan fejlődött az élő beszéd minél pontosabb rögzítését lehetővé tevő sebes-, majd későbbi elnevezéssel gyorsírás. Éppen ezért a polgári magyar állam létrejöttének kezdete óta, az uralkodóval megkötött kiegyezést (1867) követően az egyik legfontosabb köz-, társadalom-, politikai és kultúrtörténeti forrásegyüttese az országgyűlés két házának tevékenységét dokumentáló *Országgyűlési nyomtatványok*.

A digitalizálási program I.: A házszabályok digitalizálása

A nemzetközi tapasztalatok vezettek el bennünket ahhoz, hogy kísérleti lépéseket tegyünk az országgyűlési dokumentumok digitalizálása terén, ekkor még elsősorban állományvédelmi megfontolásokból. A fokozottan veszélyeztetett dokumentumokból olyanokat jelöltünk ki, amelyeknek megőrzése és szolgáltatása egyértelműen az Országgyűlési Könyvtár alapfeladatai közé tartozik. Célunk volt az is, hogy olyan dokumentumokat válasszunk ki, amelyekről feltételezni lehetett, hogy más közgyűjtemény nem fogja digitalizálni azokat. További fontos szempont volt az is, hogy a megkezdett kísérlet viszonylag gyorsan befejezhető és eredményes legyen. Ezért a digitalizálandó dokumentumok körét szűken jelöltük ki. Így esett a választás az országgyűlési házszabályokra. A feldolgozandó dokumentummennyiség az 1848 és 1986 közötti időszak 37 kötetének, mintegy 3500 oldala volt.

Természetesen a konkrét munkafeladat megtervezésekor azonnal felmerült az állományvédelmi szemponton túl a tájékoztatás szempontja is. Nevezetesen az, hogy az eredeti dokumentumok védelme érdekében, használatukat csökkentjük, az azokból készülő fénymásolást megszüntessük, emellett azonban lehetőleg minél szélesebb körben tegyük a hozzáférhetővé a házszabályokat az olvasóközönség számára. A feldolgozás során a dokumentumok oldalairól felvételek készültek, a tartalomjegyzékek begépelésre kerültek, és a gépelés lektorálása után, a tartalomjegyzékben szereplő oldalszámokhoz a megfelelő oldalról készült képek hozzákapcsolása is megtörtént. Ennek egyik eredményeként a számítógép képernyőjén „könyvszerűen” lapozgatni lehet a dokumentumokban, másik eredményeként a tartalomjegyzék tételére kattintva közvetlenül is el lehetett jutni a tartalomjegyzék tételéhez kapcsolódó oldal képéhez. A digitalizálás lehetővé tette, hogy – az állomány védelme érdekében – megszüntessük az eredeti kötetekből történő fénymásolást. Ezzel párhuzamosan a könyvtár számítógépes hálózatán biztosítottuk a digitalizált házszabályok olvasásának lehetőségét, a kutatást és az azokból történő nyomtatást. Ez természetesen nem zárta ki indokolt esetben az eredeti kötetek megtekintését, az abban való kutatást.

A munka második fázisában került sor a digitalizált házszabályok interneten történő közzétételre. A könyvtár honlapján létrehoztunk egy „*e-Könyvtár*”-at, amelynek egyik menüpontja a „Parlamenti dokumentumok” címet viseli. Itt találhatóak a digitalizált házszabályok. Az első tapasztalatok szerint a kötetek állagromlása megállt, és az olvasóközönség is kedvezően fogadta az új technikai megoldást.

Táblázat
a digitalizált dokumentumokról (sötétebb háttér)

	Országgyűlési nyomtatványok			Egyéb dokumentumok			
Képviselőház	Naplók	Irományok	Jegyző-könyvek	Mutató	Házszabály	Almanach	Lakcímjegyzék
Főrendiház/Felsőház	Naplók	Irományok	Jegyző-könyvek	Mutató	Házszabály	Almanach	Lakcímjegyzék
Delegáció	Naplók	Irományok	Jegyző-könyvek	Mutató	Ügyrend		

**A digitalizálási program II.:
az országgyűlési almanachok (1884–1944) digitalizálása**

A kísérlet első szakaszának sikere után úgy döntöttünk, hogy újabb dokumentumkőre terjesztjük ki a digitalizálási tevékenységet. A kezdeti tapasztalatok azt mutatták, hogy a folytatáshoz hasznos lenne egy állománykímélő könyvszkennert beszerzése, amely alkalmas „ipari mértékben” kiszolgálni a könyvtár igényeit. Az eszköz beszerzéséhez szükséges pénzügyi fedezetet egy sikeresen lebonyolított Informatikai és Hírközlési Minisztérium által kiírt „24. óra – Kulturális kincseink digitalizálása” című pályázat biztosította, így 2004-ben közbeszerzéssel sikerült egy Zeutschel OS 500-as szkennert beszerezni.

Az almanachok digitalizálására való kijelölésénél alapvető szerepet játszott az a tény, hogy ez az egyik leggyakrabban használt dokumentumtípusa a Magyar parlamenti gyűjteménynek. A képviselők különböző életrajzi és egyéb adatait nagyon gyakran keresik az olvasók, különösen a kutatók.

Az almanachok digitalizálása során az 1886 és 1944 közötti időszakban megjelent 16 kötet, mintegy 8000 oldalának feldolgozására került sor. Az eljárás hasonló volt a házszabályok esetében alkalmazotéhoz. A tartalomjegyzék tételeit és a képviselők neveit betűhíven rögzítettük, és ezután a megfelelő tartalomjegyzék-adatokhoz és a nevekhez hozzákapcsoltuk a megfelelő oldalakat. Az így elkészült digitális állomány szintén az e-Könyvtár / „Parlamenti dokumentumok” menüpont alá került. Ebben az időszakban készített el a Jelenkutató Alapítvány három történelmi almanachot (Ideiglenes Nemzetgyűlés almanachja, 1944. december 21.–1945. november 29.; Nemzetgyűlés almanachja, 1945. november 29.–1947. július 25.; Országgyűlés almanachja, 1947. szeptember 16.–1949. április 12.). A gyűjtemény teljessé tétele érdekében sikerült megszerezni ezeknek a köteteknek az interneten való közzététele jogát, és a nagyközönség számára hozzáférhetővé tennünk a történelmi almanachok sorát.

Az almanachok digitalizálási projektjét egy könyvbemutató és szakmai kerekasztal-beszélgetés zárta le 2006. szeptember 15-én az Országgyűlési Könyvtárban. Ekkor került sor az 1947–49. évi országgyűlés almanachjának sajtóbemutatójára. Ez alkalmából köszöntőt mondott Szili Katalin, az Országgyűlés elnöke. A művet Soltész István, az Országgyűlés főtitkára, a kötet kiadója és Marelyn Kiss József a kötet főszerkesztője mutatta be. A bemutatót szakmai kerekasztal-beszélgetés zárta le. Az 1947–1949. évi országgyűlés tevékenységét elemző beszélgetésen részt vett Gyarmati György, az Állambiztonsági Szolgálatok Történelmi Levéltárának főigaz-

gatója, Szász Zoltán, az MTA Történettudományi Intézetének tudományos főmunkatársa, Vida István, az ELTE Politika Tudományok Tanszékének munkatársa és Domány András, a Magyar Rádió vezető szerkesztője.

Táblázat a dokumentumokról

	Országgyűlési nyomtatványok			Egyéb dokumentumok			
Képviselőház	Naplók	Irományok	Jegyző-könyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Főrendiház/Felsőház	Naplók	Irományok	Jegyző-könyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Delegáció	Naplók	Irományok	Jegyző-könyvek	Mutató	Ügyrend		

**A digitalizálási programban nem tervezett szakasz:
„minősített” minisztertanácsi határozatok**

Az országgyűlési dokumentumok digitalizálásának programját megszakította egy váratlan lehetőség. A Miniszterelnöki Hivatal a jó szakmai együttműködés keretén belül – a Magyar Országos Levéltárba történő leadás előtt – lehetővé tette, hogy másolatot készítsünk a korábban „minősített”, mára e korlátozás alól felmentett kormányhatározatok meghatározott köréről. A kormány határozatainak 1950-től létező, nagyrészt háromezres (kisebb részben más) számozású és „minősített” (nem nyilvános) része rendkívül értékes forrásanyag a korszak kutatói számára. E határozatok gyakran fontos gazdasági, politikai, diplomáciai, két- vagy többoldalú nemzetközi szerződésekkel, nem ritkán személyi kérdésekkel foglalkoztak.

Az irattárban–levéltárban szunnyadó és pusztuló 1950–1988 között keletkezett szövegek digitalizálására került sor. A 3000-res számmal induló, nem nyilvános kormányhatározatok 1955-ben kezdődtek „Titkos” vagy „Szigorúan titkos” minősítéssel. Ezeket csak egy igen szűk kör kapta meg: a párt (MDP, majd MSZMP) és a kormányzati szervek felső vezetői közül. Az elmúlt időszak tapasztalatai alapján úgy gondoltuk, hogy célszerű lenne a digitalizált dokumentumokat adatbázisba szervezni. Ezért a határozatok tárgyát, számát és dátumát rögzítettük. E metaadatokat a határozatok oldaláról készült képpel összekapcsoltuk, ami lehetővé tette a határozatok „könyvszerű” nézegetését a számítógép képernyőjén, illetve a metaadatokban (a határozatok tárgya, száma és dátuma) való keresést is. Az adatbázis a könyvtár számítógépes hálózatán áll olvasóink rendelkezésére.

**A digitalizálási program III.:
az Országgyűlés 1861 és 1918 között keletkezett dokumentumainak
adatbázisa**

A házsabályok és az almanachok digitalizálása és internetes közzététele után úgy éreztük, megérett a helyzet ahhoz, hogy nagyobb léptékben haladjunk a parlamenti dokumentumok digitalizálásával. Egy, a korábbiaknál jóval nagyobb dokumentumkört, a naplókat és irományokat jelöltünk ki. Tekintettel arra, hogy a teljes anyag több százezer digitalizálandó oldalt jelentett, ezért szűkítettük a feldolgozott

időszakot az 1861 és 1918 közötti évekre. Ezek az évek jelentik a dualizmusnak nevezett történelmi korszak (és közvetlen előzményének) időhatárait, ezáltal a feldolgozott dokumentumok egységet alkotnak. A kiválasztás során messzemenően figyelemmel voltunk az olvasói igényekre is. Ezért maradt ki a feldolgozás első szakaszából a felsőházi irományok mintegy százezer oldalnyi állományrésze (ez ugyanis kevésbé keresett, mint a felsőházi napló).

A képviselőházi naplók és irományok, valamint felsőházi naplók és irományok mennyisége több mint félmillió dokumentumoldalt jelent. Egy-egy oldal pedig kb. 5000 leütést (karaktert) tartalmaz. Az előzetes felmérés azt mutatta, hogy a több százezer oldal újragépelése és a különböző korszakok megfelelő helyesírását alapul vevő lektorálása milliárdos költséget jelentene. Erre természetesen pénzügyi források nem álltak rendelkezésre, de ekkora költség ráfordítása indokolatlan is lett volna a feladat elvégzésére.

A naplók és irományok hitelességének megőrzése a projekt alapvető feltétele volt, ezért döntés született arról, hogy ezen dokumentumok oldalainak eredeti képét meg kell őriznünk. Megkezdődött tehát 1861-től a képviselőházi naplók digitalizálása. 2007 közepére saját erőből, a korábban beszerzett Zeutschel gépen, mintegy 150 000 oldalt digitalizáltunk. A korábbi tájékoztatási tapasztalatok és felhasználói igények is felvetették annak szükségességét, hogy jó lenne olyan adatbázist építeni a digitalizált dokumentumokból, amelyben a hiteles szöveget tartalmazó képek megtekinthetősége mellett biztosított a teljes szövegű visszakeresés is. Végül az alkalmazott eljárás az ún. kétrétegű pdf.-gyártás lett. A felső – látható – réteg az eredeti dokumentumoldal jó minőségű fekete-fehér digitális képe. Az alsó – mögöttes, az olvasó számára láthatatlan – réteg pedig ugyanezen oldal optikai karakterfelismerő (OCR) programmal feldolgozott szövege. Ez a technológia egyesíti a képként vagy szöveggént digitalizált dokumentumállományok előnyeit. Az egyik feldolgozási móddal csökkenti a másik feldolgozási mód hiányosságait. A létrehozott elektronikus dokumentumok az eredetiek hű másai, tehát az olvasó számára információ nem vesz el, tökéletes példányt lát. Ugyanakkor az optikai karakterfelismerés során létrehozott szövegben keresni, a találatot megjeleníteni, sőt a kijelölt szöveget átemelni is egyszerűen és könnyen lehet. Azonban az optikai karakterfelismerés nem száz százalékos pontosságú, tehát az így kapott szövegben hibák lehetnek, ezt kompenzálja a dokumentumokról készült kép.

Már az adatbázis tervezésekor egyértelmű igényként fogalmazódott meg, hogy annak interneten elérhetőnek kell lennie.

Az irományok digitalizálásában és az adatbázis létrehozásában az Arcanum Kft., az adatbázis dizájnjának megtervezésében az RBL Marketing Kft. vett részt.

Táblázat a dokumentumokról

	Országgyűlési nyomtatványok			Egyéb dokumentumok			
Képviselőház	Naplók	Irományok	Jegyzőkönyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Főrendiház/Felsőház	Naplók	Irományok	Jegyzőkönyvek	Mutató	Házsza-bály	Almanach	Lakcím-jegyzék
Delegáció	Naplók	Irományok	Jegyzőkönyvek	Mutató	Ügyrend		

Az előző táblázatból jól látható, hogy napjainkig a digitalizálás során milyen dokumentumtípusok készültek el és váltak hozzáférhetővé az interneten.

Az Országgyűlési dokumentumok adatbázis használatának bemutatása

Az adatbázis használata átlagos számítógépes ismeretekkel és keresési gyakorlattal nem jelent problémát. Természetesen különböző keresési stratégiák építhetők fel. Az adatbázis összes funkciójának bemutatása szétfeszítené ennek a cikknek a kereteit, ezért csak a főbb funkciókat mutatom be. Az adatbázis háromféle felhasználást tesz lehetővé.

1. Lehetőség van a digitalizált naplók és irományok „könyvszerű” megtekintésére. A „Keresés a dokumentumokban” gombra kattintás után egy új keresőfelület nyílik meg. A kettéosztott lap alsó felében találhatóak a kötetekről készült képek. Itt a hagyományos használathoz hasonlóan lehet a köteteket végiglapozni (böngészni), elejétől a végéig vagy tetszés szerint válogatva az oldalak között. A kiválasztott oldal pedig kinyomtatható.
2. Lehetőség van a mutatókötetekben¹ való keresésre („Keresés a naplómutatókban”). Begépelésre kerültek a korabeli naplómutató kötetek. Ezzel lehetővé vált, hogy a név- és tárgymutató tételeihez hozzárendelődjenek a dokumentumok megfelelő oldalainak digitális képei. A mutatókötetek adataira támaszkodva történő keresés biztosítja mindazon lehetőségeket, amik korábban, az eredeti dokumentumok használata során az olvasók és kutatók rendelkezésére álltak. Ez a funkció azonban komoly többletlehetőséget is nyújt a kereséshez. A két részre osztott keresőfelület alsó felén lehet a naplókötetekben „böngészni”. Tehát kötetről kötetre haladva, a betűrend alapján tájékozódhatunk a különböző témák között.
3. A keresést további három funkció segíti:

A keresőszavak szerinti,
tárgyszavak szerinti és
ciklusok szerinti keresés.

A **keresőszavak** használata esetében a mutatókötetben előforduló bármely szóra kereshetünk. Ennek azért van különös jelentősége, mert a mutatókötetek előre megkonstruált tárgyszavakkal dolgoznak. Például ilyen a következő tárgyszó: „háznagy jelentése”. Ennél a példánál maradva, a tárgyszóban szereplő mindkét kifejezésre kereshetünk külön-külön is. Találatot fogunk kapni a „háznagy” (69 találat) és a „jelentése” (716 találat) kifejezésekre is, ha azokra külön-külön kereshetünk is. Természetesen az adatbázisokban álta-

¹ A naplómutatók által feldolgozott adattömeget jól jellemzi Endrődi Sándornak, a képviselőház naplószerkesztője előszavának részlete, melyet az 1861–1892-es első kötethez írt: „Az országgyűlés képviselőházának naplói 1861-től a most folyó országgyűlés kezdetéig 151, irományai pedig 189, együttesen tehát 340 kötetre rúgnak. Egész könyvtár, tömérdek anyag, rengeteg tárgy nagyraktára, mely azonban egy összesítő index hijával eddigelé jóformán hozzáférhetetlen volt. Ennek a munkának az a célja, hogy hozzáférhetővé tegye a nagyszámú kötetek anyagalmazát s elkalauzolja a keresőt az évtizedek során át összetorlódott tárgyak szövevényei közt.”

lánosan használatos „csonkolásokra”, illetve a keresésből egyes betűk kihagyására is mód van.

A **tárgyszavak** szerinti keresésnél elsősorban a teljes tárgyszóra való keresés biztosított. A mutatókötetek készítői által megkonstruált tárgyszavak listáját az index tartalmazza. Az előző példánál maradva, ha a háznagy jelentéseit keressük, akkor célszerű az előre megkonstruált (és az index által felkínált) tárgyszót, a „háznagy jelentése” használni a keresés során. Ebben az esetben 10 találatot fogunk kapni. Ezek a találatok pontosan az általunk keresett témával, a háznagyok jelentésével foglalkoznak.

A **ciklusok** szerinti keresés lehetősége is adott, tehát megvizsgálhatjuk, hogy különböző parlamenti ciklusokban (időszakokban) milyen tárgyszavakat, milyen témákat dolgoztak fel a mutatókötetekben. Azonban az igazi jelentősége a ciklusok szerinti keresés lehetőségének az, hogy szűkítő feltételként alkalmazhatjuk. Az előbbi példánál maradva, ha arra vagyunk kíváncsiak, hogy 1931–1936-os ciklusban volt-e háznagy jelentés, akkor a keresés során két elemre kereshetünk egyszerre. A Tárgyszó mezőben „háznagy jelentése”, a ciklus mezőben pedig „*-1931” (a* jel alkalmazása „kiváltja” a gondolatjel előtt szereplő betűket, tehát a kn² és fn³ rövidítéseket). A keresést lefuttatva két találatot kapunk. A találatok szerint a képviselőházban és a felsőházban is foglalkoztak a háznagy jelentésével az 1931–1936-os ciklusban.

4. Lehetőség van a dokumentumok szövegében való keresésre is. Ez a lehetőség nagyságrendekkel növeli a keresés hatékonyságát. Hiszen az összes dokumentum teljes szövegében szabadszavasan kereshetünk. A keresés alapjául szolgáló szöveg az optikai karakterfelismerés (OCR) eredményeként jött létre, a feldolgozás pontossága kb. 98–99 százalékos. A naplók és irományok teljes szövegében való keresés valóban forradalmasítja az ezen dokumentumokban történő kutatás hatékonyságát, hiszen sokkal differenciáltabban kereshetünk, mintha a mutatókötetekre támaszkodnánk segédeszközként. Természetesen az ilyen irdatlan adatmennyiségben történő keresés során hatalmas találati halmazokat kaphatunk. Például a „jegyző” szóra keresve 48 552 találatot, a vasút szóra keresve 22 466 találatot kapunk. Szükséges tehát, hogy különböző szűkítő feltételeket is lehessen alkalmazni. Természetesen lehetőség van szavak összekapcsolására a keresés során. Például, ha a „vasút” kifejezést a „Pécs” szóval összekapcsoljuk a keresés során, akkor 523 olyan találatot kapunk, amelyben a vasút és a Pécs kifejezés egy dokumentumoldalon szerepel. Ez természetesen még mindig nagy találati halmazzal jelent, ezért van jelentősége a további szűkítési lehetőségeknek. A következő kereső ablakok: dokumentumazonosító, az ülésnap és irományszám, illetve az oldalszám, ezek kitöltésével különböző szempontok szerint pontosíthatjuk a találati halmazokat.

Ha az adatbázis használata során további információra van szükség, akkor az Országgyűlési Könyvtár tájékoztató munkatársai készséggel nyújtanak segítséget.

2 kn = képviselőházi napló.

3 fn = felsőházi/főrendiházi napló

(Közvetlenül az *mpgy@ogyk.hu* e-mail címen is felvehető a kapcsolat.) Itt kell megjegyezni, hogy a kutatók számára kifejlesztettük az adatbázis lokális (csak a könyvtárban használható) verzióját is, ami további funkciókat is tartalmaz az internetes verzióhoz képest.

Az adatbázis fizikai jellemzői is lenyűgözők. A feldolgozott időszak 769 kötetének 312 670 oldalán több mint egymilliárd leütés (karakter) található. Ez azt jelenti, hogy az adatbázisban több mint százmillió szó vált visszakereshetővé. Ha a 769 kötetet egymásra helyeznénk és így egy oszlopot alkotnánk, akkor a Parlament kupolája csúcsának belmagasságát (27 méter) elérné.

A sajtóbemutató és a jövő

Szili Katalin házelnök az adatbázis sajtóbemutatóján az Országgyűlési Könyvtár szempontjából történelmi léptékűnek nevezte az elvégzett munkát. Kiemelte, hogy az interneten való közzététel időben és térben kitágítja a tudományos vagy egyéb célú hozzáférést a dokumentumokhoz, hiszen napi 24 órában, Magyarországon és határainkon túl – így az egykori Ausztria–Magyarország területén létrejött államokban is – könnyen elérhetővé váltak. A projekt külön érdemének tartotta, hogy olyan megoldást választott a könyvtár, ami a nemzetközi összehasonlításban is az élvonalhoz tartozik, és a feldolgozott dokumentumtömeghez képest költségkímélő, takarékos eljárás volt.

Az adatbázist szakmai szempontból Reisz T. Csaba, a Magyar Országos Levéltár főigazgató-helyettese mutatta be a sajtó képviselői számára. (Előadásának vázlatát olvasható az Országgyűlési Könyvtár honlapján.)

Szili Katalin a sajtóbemutató lezárásként fontosnak és szükségesnek tartotta a Magyar parlamenti gyűjtemény digitalizálásának folytatását, a projekt megvalósítását támogatásáról biztosította. Az adatbázis ettől az időponttól a <http://mpgy.ogyk.hu> url-címen megtekinthető.

A sajtóbemutató másnapján 566 alkalommal léptek be az adatbázisba.

Redl Károly