

## Hogyan jutnak el a virtuális látogatók a könyvtárba?

### Változó informálódási szokások

Az internet megváltoztatta az információkeresés és -szolgáltatás természetét. Azelőtt a könyvtár volt az első kézenfekvő hely, ahová az emberek információért fordultak, de ez már egyáltalán nincs így. Egy 1994-es ausztrál felmérés [1] még azt mutatta, hogy a beiratkozott olvasók durván 90%-a és a nem könyvtári tagoknak is 53%-a automatikusan a könyvtárat keresi meg, ha szeretne valamit megtudni. Ugyanakkor a 2007-ben zajlott *Pew Internet & American Life Project* kutatás [2] eredménye szerint a megkérdezettek csaknem 60%-a már az internetet nézi meg ilyenkor esetben és csak valamivel több, mint 10% fordul egy nyilvános könyvtárhoz. Láthatóan az emberek számára a könnyű használat, a kényelmes hozzáférés és a gyorsaság legalább olyan fontos, mint az információk minősége és megbízhatósága.

Az információforrások is sokkal változatosabbak lettek a digitális korban, mint amilyenek előtte voltak. Az 1990-es évek elején a tartalomipart olyan hagyományos, általában igen költségesen előállítható termékek jellemezték, mint az újságok és magazinok, a zenei felvételek, könyvek, filmek és tévéműsorok, a dobozban árult szoftverek és számítógépes játékok, előfizetéses online adatbázisok, vásárlói katalógusok... 2007-re, a webkettes technológiák elterjedésének köszönhetően, a felhasználók által generált sokféle tartalom versenyképessé vált – méretben és népszerűségben egyaránt – a hagyományos tartalomszolgáltató iparágakkal.

Ausztráliában a Wikipédia a leggyakrabban használt online referenszforrás, az USA-ban pedig egy felmérés szerint az internetezők közel fele használ valamilyen keresőgépet egy átlagos napon. Az ilyen statisztikák alapján mondják sokan, hogy a könyvtárak mint információszolgáltatók jelentősége rohamosan csökken. Érdeemes azért ezek mögé a számok mögé nézni, hogy pontosan megértsük, kik, hol és milyen fajta információk után keresgél-

nek a weben. Az alább ismertetett kutatás során azt próbálták kideríteni – a szerverek forgalmi adatait elemezve –, hogy honnan és milyen céllal érkeznek a látogatók az ausztrál *State Library of Victoria (SLV)* weboldalaira. Ez a kutatás egy nagyobb projekt része, mely az új technológiáknak a könyvtári szolgáltatásokra gyakorolt hatását vizsgálja.

### Az SLV honlapja

Victoria állam könyvtára, más könyvtárakhoz hasonlóan, a fizikai terei és gyűjteményei mellett jelentős méretű virtuális teret és változatos digitális tartalomszolgáltatásokat is kialakított a weben. Az első honlapja még a kilencvenes évek közepén született, a jelenlegi pedig 2004 óta működik. A 2006–2007-es időszakban a könyvtár épületeiben 1 milliónál valamivel több látogató fordult meg, míg ugyanezen idő alatt a honlapon több mint háromszor ennyi látogatást regisztráltak, és a kapcsolódó további szervereken pedig 22 millió felett volt ezek száma összesen. És bár a könyvtár rendszeresen végez kérdőíves felméréseket a tényleges használói között, hogy kik és miért keresik fel a könyvtárat és mennyire vannak megelégedve a szolgáltatásával, a virtuális látogatókról alig van valami információjuk.

Az SLV honlapján a szokásos ismertető (nyitva tartás, kiállítások, események stb.) mellett belépési pontok vannak a katalógushoz és egyéb – részben teljes szövegű – online adatbázisokhoz; továbbá ezen a szerveren kapott helyett a *LaTrobe* folyóirat, mely a könyvtár gyűjteményére alapozott kutatásokról közöl cikkeket; valamint jelentős mennyiségű digitális tartalom is elérhető innen. A könyvtár képgyűjteményének egyharmadát digitalizálták eddig (több mint 200 ezer képet), továbbá régi tervrajzok és térképek, újságok és hangfelvételek is találhatóak a digitális könyvtárakban.

Hatalmas mennyiségű adat van a szerver naplóállományában, amelyekből kielemezhető a felhasználók tevékenysége. Az elemzőeszközök egy része ingyenes, másokért fizetni kell. Bár ezek elsősorban piackutatási célokra lettek kialakítva, de sok, eddig nagyrészt kihasználatlan lehetőség van bennük egy könyvtár számára is, amely szeretne többet megtudni a weblapjai látogatóiról. Az elmúlt években jelentek meg már a *log*fájlok statisztikai feldolgozásáról beszámoló publikációk, például digitális könyvtárak esetében. A jelen kutatás azonban két új elemet is tartalmaz ezekhez képest: egyrészt megpróbálták „kézi” osztályozással csoportosítani azokat a webhelyeket, ahonnan belépnek, illetve ahová kilépnek a könyvtár honlapjának látogatói; másrészt elemezték azoknak a keresőkérdeéseknek a tartalmát, amelyek végül az SLV szerverére vezették az online információk után kutató felhasználókat. Mindkét vizsgálatnál nem elsősorban a leggyakoribb esetekkel foglalkoztak, hanem a *long tail*, vagyis a gyakorisági görbe elnyúló szakasza érdekelte az elemzőket, a ritkább vagy teljesen egyedi esetekből igyekeztek hasznos információkat kiszűrni a virtuális látogatók jellemzőire és motivációira vonatkozóan.

### Vizsgálati módszerek

Az itt bemutatott statisztikai adatok a *Hitwise*, a *WebTrends* és a *Google Analytics* elemzőivel készültek; mindegyik szolgáltatás másként mér, és megvannak az előnyeik és a korlátaik. Ezúttal leginkább a *Hitwise* ([www.hitwise.com](http://www.hitwise.com)) által szolgáltatott információkat használták fel. Ez a cég az ausztráliai internet-előfizetéseknek több mint az egyharmadánál regisztrálja az adatforgalmat. Egy saját fejlesztésű szisztéma segítségével a felhasználók becsült életszínvonal-profilját is meg tudja adni, ami hasznos információ. Mivel az előfizetők aktivitását figyeli, a keresőrobotok nem torzítják az adatokat. A *Hitwise*-nak annyi hátránya van, hogy a teljes felhasználói körnek csak egy – igaz, meglehetősen nagy és reprezentatív – részét méri. Elvben a *Google Analytics* lenne alkalmas a webszerver teljes forgalmának regisztrálására és kiértékelésére, de ehhez minden weblapba egy speciális kódot kellene beletenni. Mivel az SLV site több százezer oldalból áll, és csak bizonyos típusokban van benne ez a kód, így a teljes felhasználói aktivitást ezzel sem tudják elemezni.

Ahhoz, hogy a bejövő (*upstream*) és kimenő (*downstream*) irányokat osztályozni tudják, több ezer webhelyet soroltak be a *Hitwise* iparági osztá-

lyozási rendszere szerint, amit előbb némileg módosítottak, hogy könyvtári szempontból relevánsabb kategóriák is legyenek benne. A felhasználók által begépelte keresőkérdeéseket pedig a *Strauss* és *Corbin* cikkében [3] leírt „megalapozott elmélet” (*grounded theory*) nyílt kódolású (*open coding*) technikájával tették statisztikailag elemezhetővé. Minden kérdést egy vagy több címkével láttak el, és azután ezt a folyamatot többször megismételték, amíg a végső, már elég egységes kategóriákat sikerült kialakítani. Vagyis nem egy előre kitalált osztályozási rendszert alkalmaztak, hanem a kérdések végigolvasása során kristályosodtak ki a jellemző kérdéstípusok. Így persze elkerülhetetlenek bizonyos szubjektív döntések közben, de nem is az egyes kategóriák pontos mérete érdekes itt, hanem az eloszlási mintázatuk.

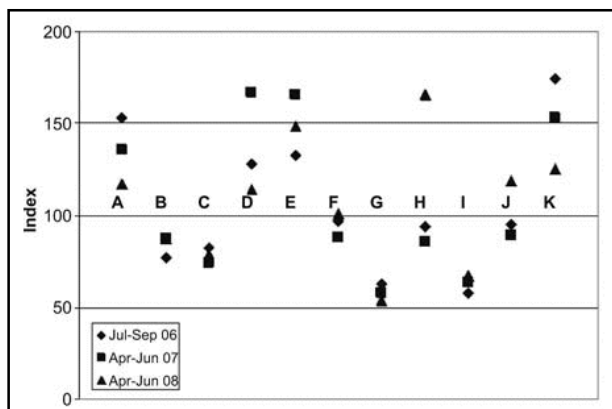
### Eredmények

Elsőként egy grafikonon ábrázolták a honlap látogatóinak számát 2006 januárja és 2008 májusa között, havi bontásban. Mivel ezekben az összesített adatokban a keresőgépek robotjai által okozott forgalom is benne volt, egy sötét sávval jelezték ezek becsült forgalmát, ami változó ugyan, de nagyjából a teljes letöltések egyharmadát teszi ki. Egy másik görbén azt is feltüntették, hogy a látogatók közül hányan érkeztek úgy, hogy egy keresés eredményeként előállt találati listából az SLV valamelyik weblapjára ugrottak. Mindkét görbe viszonylag egyenes volt a vizsgált időszakban; a legnagyobb forgalmat 2008 április–májusában regisztrálták, aminek talán az lehet a magyarázata, hogy akkor rendeztek a könyvtárban egy nagy érdeklődéssel kísért, sokfelé reklámozott kiállítást. A *Google Analytics* mérései szerint hétfőtől csütörtökig van a legnagyobb forgalom, ami azután péntekenként elkezd csökkenni. Hétféteken a látogatások száma általában mintegy kétharmadnyi a munkanapokhoz képest.

A *Hitwise* életszínvonalat mutató statisztikája látható a mellékelt diagramon (1. ábra), ami 2006 és 2008 között, három hónapos periódusok adatai alapján készült. A függőleges tengelyen a 100-as indexértéktől való eltérés mutatja, hogy az SLV honlapját látogatók között az egyes típusok mennyire vannak felül- vagy alulreprezentálva Ausztrália internetező lakosságának átlagos megoszlásához képest.

Ahogy várható volt, az ábrán a jómódú, városlakó, tanult emberek átlagon felül vannak reprezentálva,

és a társadalmilag vagy anyagilag hátrányos helyzetűek aránya pedig átlagon aluli. Azt nem lehet tudni, hogy vajon a tanultabb réteg azért jön-e nagyobb számban a könyvtár honlapjára, mert ismeri, és ezért tudatosan választja ezt az információforrást, vagy, mert olyan az internetes aktivitásának jellege, hogy nagyobb valószínűséggel bukkan rá. Érdekes viszont, hogy a **K** kategória is mennyire az átlag fölött van, ami azt jelzi, hogy az állami könyvtár virtuális jelenléte fontos a vidéken gazdálkodó emberek számára.



1. ábra Az SLP website látogatóinak profil szerinti megoszlása (Forrás: Hitwise)

Jelmagyarázat: **A** – a leggazdagabb családok a legelegánsabb városrészekből; **B** – tehető és művelt családok a középső és külső kerületekből; **C** – fiatal családok a városok növekedési sávjában; **D** – gazdag szinglik és közösen lakó egyedülállók; **E** – tanult egyedülállók és párok, közepes jövedelemmel; **F** – jó anyagi helyzetű családok elővárosokban és város-vidék peremzónában; **G** – nehéz anyagi helyzetben levő családok és szinglik városi területeken; **H** – multikulturális keveredésű közösségek; **I** – hátrányos helyzetű, városkörnyéki vagy vidéki területek; **J** – nagyvárosokban vagy vidéken élő idősek; **K** – falusi életmódot folytató gazdálkodó családok.

### Honnan jöttek a látogatók?

Erre a kérdésre kétféle válasz is van: a földrajzi helyüket a Google Analytics és a Hitwise segítségével állapították meg. Az előbbi országok szerint, az utóbbi pedig Ausztrálián belül mutatja a felhasználók eloszlását. A vizsgált időszakban minden évben 90% felett volt az ausztrál látogatók aránya, és ezek valamivel kevesebb mint kétharmada Victoria államban lakik, hatodik pedig New South Wales területén.

Ami azt a virtuális helyet illeti, ahonnan a felhasználók az SLV website-ra jutottak, az már nehezebben megállapítható. A látogató érkezhettek egy másik weblapról egy linket követve, vagy egy URL cím begépelésével, vagy egy könyvjelzőre kattintva, vagy többlapos/füles böngészéssel (*tabbed browsing*). Ha feltételezzük, hogy a legutóbb nézett weblap jellemző a felhasználó érdeklődési irányára, akkor ezeknek az *upstream* oldalaknak tulajdonosi és tartalmi elemzése alapján a látogatók információs igényeire lehet következtetni. 2008 májusában több mint egyharmaduk egy Google találati listából kattintott valamelyik SLV lapra, 7,8% pedig a *National Library of Australia*, vagy a szintén a nemzeti könyvtár által szolgáltatott *Picture Australia* képadatbázis honlapjáról érkezett. Mindkét webhelyről mutatnak hivatkozások az SLV szerverén levő digitális képekre, így nem csoda, hogy ennyire elől vannak a listában. Az első tízben ezen kívül egy-két további keresőgép, a Wikipédia, a Facebook, a Gmail és az eBay szerepel. A toplistának ez az első tíz helyezettje hozta a teljes forgalom több mint felét, összesen pedig – a Hitwise reprezentatív mintája szerint – 961 egyedi webhelyről kattintottak át SLV-s oldalra ebben a hónapban. A Hitwise ezeket ugyan automatikusan ellátta iparági kódokkal (egy kb. 150 kategóriát tartalmazó osztályozási rendszert használ ehhez), de végül át kellett kódolni őket úgy, hogy a jelen vizsgálat igényeinek jobban megfeleljenek. Ezért a kutatók minden oldalt egyetlen, leginkább jellemző kategóriába soroltak be, és a kódrendszert kiegészítették olyan kategóriákkal is, mint „könyvtárak”, „oktatási és kulturális intézmények” stb. Az újrakódolás sok nehézséggel járt, és a szubjektivitást sem tudták teljesen kizárni, mert minden esetben csak egyetlen kódot lehetett választani. Például az igen vegyes tartalmú ausztrál webarchívum a *Pandora* végül a „múzeumok, galériák és archívumok” alá került. A kategóriákat egy oszlopdiagramon ábrázolva – a 2008. májusi adatok alapján – 40% körül volt a keresőgépek aránya, 10% fölött az „egyéb” kategória, amibe egy sor, önmagában túl ritkán előforduló témát soroltak (pl. életmód, vásárlás, utazás, sport, zene, egészség, autó, étel-ital, szerencsejáték). A keresőket és az „egyebeket” figyelmen kívül hagyva a legnagyobb a könyvtári webhelyek aránya, amiben az is közrejátszik, hogy több egyetemi könyvtár honlapján vannak közvetlen keresőlinkek az SLV katalógusára. A „számítógép és internet” kategória mérete fokozatosan nőtt három év alatt; ennek a füles böngészés elterjedése és a – szintén ide sorolt – közösségi oldalak növekvő népszerűsége lehet a magyarázata. A keresőgépek aránya viszont csökkent az

évek során, talán azért, mert az SLV hátrébb csúszott a találati listákon a rangsorban, így kevésbé látható; vagy azért, mert a felhasználók inkább a gyarapodó számú linkeket követik, és így jutnak a könyvtári oldalakra.

Hasonlóképpen megvizsgálták azoknak a webhelyeknek a jellegét, ahová a látogatók továbbmentek az SLV szerveréről. Itt már a keresőgépek aránya kisebb és a második legnépszerűbb cél a nemzeti könyvtár. Az eBay itt is benne volt az első tízben 2008 májusában, aminek állítólag az a magyarázata, hogy az ott meghirdetett tárgylistákban időnként linkek vannak az SLV digitalizált gyűjteményeinek megfelelő képeire. A *downstream* webhelyek eloszlása egyenletesebb, mint az *upstream* helyeké: a tíz leggyakoribb kilépési irány csak a negyedét teszi ki az összesnek. Itt is a *long tail*, vagyis a kisebb gyakorisággal jellemzett kategóriák elemzése ad teljesebb képet, és érdekes a belépési és a kilépési site-típusok összehasonlítása is. Például az, hogy az SLV-re többen jöttek keresőoldalakról, mint ahányan továbbmentek ilyenekre, jelentheti azt, hogy megtalálták, amit kerestek. De az is okozhatja ezt a jelenséget, hogy több lapos böngészőt használnak, és miután egy találati listában rákattintottak az SLP valamelyik oldalára, visszaváltottak a listát tartalmazó földre és folytatták a többi találat nézegetését.

Hogy mi hozta a virtuális látogatókat az SLP-hez, azt a keresőkérdések elemzésével lehet kideríteni. Persze – mivel csak azokat a kérdéseket ismerik, amelyek eredményeképpen egy felhasználó eljutott az SLP valamelyik weboldalára – valójában csupán azt mutatja meg ez az elemzés, hogy hogyan találtak el ide az emberek, és mit találtak itt, nem pedig azt, hogy valójában mit kerestek. A kilenc leggyakoribb keresés a „*State Library of Victoria, Melbourne*” kifejezés szavainak valamilyen variációja volt és az összes keresés 23%-át tette ki. A 2008 tavaszán rendezett kiállítás iránti érdeklődés is látszik az akkortájt megszorodott keresőkérdéseken. A kifejezetten az SLP-re vonatkozó kereséseken kívül leggyakrabban Victoria állam egyes területei, épületei és történelmi eseményei után érdeklődők jutottak el a könyvtár szerverére, ezek 43%-ot tesznek ki a *long tail*-en belül. Valószínűsíthetően részben családfa-kutatásokhoz szükséges információkeresések lehettek ezek, de emellett a kimondottan genealógiai jellegű kérdések is elég jelentős számban fordultak elő. További 16% általános referenzkérdés volt (pl. „Hogyan tisztítsunk le egy pené-

szes falat?”), 12% pedig egy könyvre vagy egy szerzőre vonatkozott.

Hogy megtalálták-e az emberek amit kerestek, nagyon nehéz kihámozni ezekből a webszerverstatisztikákból. Egy durva becslést mindazonáltal megkíséreltek a következő módon: mivel az SLV-nek jelentős digitális képgyűjteménye van, megpróbálták kiszámolni, hogy a látogatók hány része érdeklődött a képek iránt. Ennek érdekében kigyűjtötték az *upstream* és a *downstream* site-ok közül azokat, amelyeknek az ismertetőjében vagy az URL címében a „Flickr”, „foto”, „image”, „picture” vagy „photo” szavak szerepelnek, majd összehasonlították ezek több éven keresztül arányát. Az, hogy a felhasználók közül nagyjából kétszer annyian érkeztek ilyen, képekkel foglalkozó oldalakról (és ebben a számban a keresőgépek képtalálataira kattintók még nincsenek is benne), mint ahányan továbbmentek ilyenekre, úgy értelmezhető, hogy sokan megtalálták azt a képet az SLV-nél, amelyre szükségük volt. Az pedig, hogy a három év alatt megnőtt a képek iránt érdeklődők száma, a szélessávú hálózatok elterjedésével magyarázható.

### További lehetőségek

A forgalmi adatok elemzése igazolta, hogy az SLV webhelye egyszerre jelentős digitális tartalomszolgáltatás és a fizikai könyvtár fontos bemutatkozó helye. A továbbiakban jó volna összehasonlítani a külső keresőgépekbe írt kérdéseket azokkal, amelyeket az SLV honlapjának saját keresőjébe, illetve az SLV katalógusába írnak a felhasználók. Emellett össze lehetne vetni a site gyakran látogatott részeit a teljes honlaptérképpel, hogy kiderüljön, melyek a kevésbé vagy alig használt területei a webhelynek. Elemezni lehetne a látogatók mozgását is a site-on belül. Arra viszont már nem alkalmasak ezek az adatok, hogy különbséget tegyünk a „kereső” és a „böngésző” típusú felhasználók között. Előbbiek azért jönnek az SLV honlapjára, mert valamilyen információra van szükségük; utóbbiak viszont csak kíváncsiságból kattintanak rá egy érdekes bélyegképre valamelyik kedvenc oldalukon, ami azután átviszi őket az SLV szerverén tárolt eredeti digitalizált dokumentumhoz.

A logfájlok nagy előnye, hogy olcsón és a könyvtár használóit nem zavarva juthatunk hatalmas mennyiségű adathoz. Nem lehet viszont így elég információt szerezni a felhasználókról és szándékaikról, valamint arról, hogy megtalálták-e, amit keres-

tek. Ilyen irányú tájékozódáshoz inkább kérdőíves felmérést célszerű végezni.

### Irodalom

- [1] MERCER, C.: Navigating the economy of knowledge: a national survey of users and non-users of state and public libraries. = Libraries Working Group of the Cultural Ministers Council, 1995.
- [2] ESTRABOOK, L. – WITT, E. – RAINIE, L.: Information searches that solve problems. = Pew Internet Project, 2007.  
*[http://pewinternet.org/pdfs/Pew\\_UI\\_LibrariesReport.pdf](http://pewinternet.org/pdfs/Pew_UI_LibrariesReport.pdf)*

- [3] STRAUSS, A. – CORBION, J.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA., Sage Publications, 1998.

**WALLER, Vivien: How do virtual visitors get to the library? = The Electronic Library, 27. köt. 5. sz. 2009. p. 815–830./**  
**<http://www.emeraldinsight.com/10.1108/02640470910998533>**

(Drótos László)