

Ferenczy Gábor

AZ INTERNET MINT A FELDERÍTÉS ÚJ ADATFORRÁSA

Ma már senki sem lepődik meg azon a kijelentésen, hogy az információs társadalomban élünk. Az információs társadalomban az információ különleges szerepet tölt be, ezért törekedni kell arra, hogy minél gyorsabban, minél többet megszerezzünk belőle. Aki előbb kerül egy adott információ birtokába az fölénybe kerül a versenytársaival szemben. Az információs hadviselés során is az a célunk, hogy az információ gyors megszerzése által vezetési fölénybe kerüljünk.

Információt sokféle eljárással és sokféle forrásból lehet szerezni. A világban végbement változások hatására a nyílt források szerepe óriási mértékben megnőtt. Sok esetben szinte kizárólag csak ezek állnak a rendelkezésünkre az adatszerzésre.

Az Internet is egy nyílt forrás, amely az utóbbi néhány évben robbanásszerű fejlődésen ment keresztül. Az Internetet már szinte mindenki el tudja érni, ha még nem is otthonról, de a munkahelyéről, vagy a folyamatosan szaporodó Internet kávézók segítségével mindenképpen. Az Interneten óriási mennyiségben állnak rendelkezésre az információk. Ebből az irdatlan méretű adathalmazban feltétlenül lehet olyant találni, ami a felderítés számára értékes.

A NYÍLT FORRÁSOK

Mit tekintünk nyílt forrásnak? Nyílt forrásnak hívjuk a nyilvánosan elérhető információkat (például a törvényesen hozzáférhető információ megszerzése lekérdés, vagy megfigyelés útján), valamint azokat a nem feldolgozott információkat, amelyeknek korlátozott a terjesztése és elérhetősége. Más szavakkal nyílt forrásnak minősül minden olyan eszköz, hely, kiadvány, műsor, amihez hozzá lehet férni a piacon, vagyis azok az anyagok, amelyekhez nem tiltják a hozzáférést.

Nyílt források típusai

A fenti meghatározás alapján nyílt forrásnak tekintjük az alábbiakat:

- könyvek;
- napilapok, folyóiratok;
- térképek;
- fényképek;

- adatfájlok;
- digitális képi anyagok;
- rádió- és televízióadások;
- szabadalmi bejegyzések;
- konferenciák előadásai és írásos anyagai;
- egyetemi-, főiskolai jegyzeteket és szakdolgozatok;
- nemzetközi szerződések, bizalomerősítő egyezmények;
- internet.

A számos nyílt forrás közül csupán csak az Internettel kívánok a továbbiakban foglalkozni. A nyílt források között az Internet talán a legújabb. Az utóbbi tíz évben fejlődött át egy szűk körben használt rendszerből egy olyan a világot átfo-gó hálózattá, ami mindenkinek a rendelkezésére áll a nap huszonnégy órájában.

Az Internet méretei

Az Internet a nyílt információk hatalmas mennyiségű és változatos gyűjtemé-nyének tárháza, a hálózatokba kapcsolt számítógépek hálózata, amelynek a fej-lődése soha nem látott méretű volt az elmúlt néhány évben. Nagyon nehéz feladat lenne megbecsülni a hálóra csatlakozó személyek és szervezetek számát, de becslé-sek szerint naponta 400 és 1000 között van az újonnan megjelenő weboldalak szá-ma. A hálózatra csatlakozó egyének, szervezetek száma jelenleg 800 milliós nagy-ságrendű lehet, az összekapcsolt hálózatok száma meghaladhatja a 40 ezret.

Az elérhető adatok mennyisége roppant nagy. Egyes vélemények szerint az összes információ mennyisége elérheti a 2—10 terabájtot (10^{12}). Egy terabájt az egy millió megabájt, vagy egymilliószor egymillió karakter. Összehasonlításkép-pen egy átlagos könyvtárban, ahol 300 ezer könyv van kb. 3 terabájt adat van.

INFORMÁCIÓSZERZÉS AZ INTERNETRŐL

Az Interneten való szörfölést általában valamilyen konkrét céllal kezdjük el (pl. elolvassuk a legfrissebb híreket), de sok esetben csak taláalomra keresünk vala-mit. A gyakran látogatott, vagy a frissen felfedezett oldalaink címét felírhatjuk, de praktikusán, a számítógépen tárolhatjuk. De mit is rejt egy weboldal címe? Ezt vizsgáljuk meg a következőkben.

Címzés az Interneten

Az Interneten lévő számítógépeket az IP címük azonosítja. Ez egy négybájtos szám, amit az általános szokásnak megfelelően pontokkal elválasztva szoktak ábrázolni. Például: 172.180.20.2. Az IP cím két részből áll, egy hálózat azonosí-tó részből és egy gépazonosító részből.

A számokból álló IP címek hosszúak, nehezen megjegyezhetők. Ennek a kiküszöbölésére az egyes gépek rendelkeznek egy ún. domain címmel is, amely általában rövid, értelemmel bíró betűszavak, rövidítések ugyancsak pontokkal elválasztott sorozata.

A domain név rendszerben¹ a címek allokálása hierarchikus struktúrában történik. A domain nevek utolsó tagja általában az országnév kétbetűs — az ISO 3166 szabványának megfelelő — rövidítése (pl. „.hu” Magyarország betűjele). Előfordulnak ettől eltérő legfelső szintű domain nevek² is. A legáltalánosabban használtak a következők:

- .com — ipari, üzleti (commercial) felhasználók számára;
- .edu — egyetemek, oktatási (educational) intézmények számára;
- .net — a hálózat (network) adminisztrációs szervezetei számára;
- .org — az olyan egyesületek, szervezetek (organization) részére, amelyek a fenti kategóriákba nem sorolhatóak;
- .int — nemzetközi (international) szövetségi szervezetek, és nemzetközi adatbázisok számára;

Csak az Egyesült Államokban használt domainek:

- .gov — szövetségi kormány (government) hivatalok számára;
- .mil — a hadsereg számára.

Ugyancsak előfordulhat, hogy a legfelső domain szintjén egy hálózat neve szerepel, amennyiben a címzett nem az Interneten van. Ilyen lehet például a BIT-NET és az UUCP végződés.

A hálózat nagysága miatt az egyes információk, szolgáltatások megtalálása nem mindig könnyű. A különböző jellegű erőforrások egységes leírására jött létre az ún. URL-jelölés³. Ennek általános formája:

<protocol>://<Internet_cím>:<port_szám>/<könyvtár>/<fájl_név>?<paraméter>

A leírás általában nem ilyen bonyolult, és a legtöbb esetben a komponensek nagy része hiányzik. A Web-klinseknél az URL megadásával lehet elérni az egyes forrásokat az Interneten, de többségüknel a http:// előtagot nem is kell beírni, mert a kliens automatikusan kiegészíti vele a címet.

A leggyakrabban előforduló protokollrövidítések a következők:

- http — hypertext átviteli protokoll⁴: a Web-szerverek és -kliensek közötti kommunikáció szabályait leíró protokoll;
- ftp — fájl átviteli protokoll⁵ a TCP/IP hálózati protokoll-csomag azon része, mely az állományok gépek közötti (hibamentes) átmásolását szabályozza;

¹ DNS — Domain Name System

² TDL — Top-level Domain

³ URL — Unified Resource Locator

⁴ http — hypertext transfer protocol

⁵ ftp — file transfer protocol

- gopher — kliens/szerver elven működő rendszer az Interneten levő (elsősorban szöveges) információforrások egységes kezelésére, hierarchikus menükön keresztül;
- news — hírcsoport, kommunikációs fórumok összessége, az elektronikus faliújságok és a levelezőcsoportok keveréke a témakörök szerint;
- telnet — folyamatos, interaktív terminálkapcsolatot lehetővé tevő protokoll két gép között;
- mailto — levél küldése egy címre.

Egy példa az URL címre:

— ftp://ftp.microsoft.com/peropsys/windows/readme.txt

Ennek jelentése: egy ftp protokollal elérhető, az ftp.microsoft.com gépen a peropsys/windows könyvtárban található readme nevű fájl.

Az internetes információszerezés előnyei

Az Interneten megtalálható anyagok tartalma széles skálán mozog, a napi friss hírektől az archivált újságcikkekig, a tudományos írásoktól a rock rajongók üzenetváltásáig. Sok egyén, társaság, kormányhivatal és egyéb szervezet közread különböző információkat az Interneten. A hírcsoportok, vagy beszélgetési csoportok (newsgroup) ezrei témakörökbe szervezett üzeneteket és vitafórumokat tartalmaz az emberi élet minden területéről, a befőzési receptektől a szenvedélyes vitáig a kulturális és politikai élet problémáiról.

Mivel az Interneten nagyon olcsón lehet közétetni az információkat — összehasonlítva a hagyományos kiadási eljárásokkal — szervezetek, katonai egységek és kormányzati hivatalok növekvő mennyiségben alkalmazzák az Internetet arra, hogy rajta keresztül tájékoztassák az ügyfeleiket, tagságukat, vagy az állampolgárokat.

A napilapok és folyóiratok egyre nagyobb számban jelennek meg az utcai árusok mellett az Interneten is, sőt léteznek olyan újságok, amelyek nyomtatott formában meg sem jelennek csak elektronikus kiadásuk létezik.

Az Interneten keresztül már régóta lehet rádiót hallgatni, de az utóbbi időben már egyre több televízió műsort is lehet a hálózaton keresztül nézni. A televízió adások széleskörű elterjedésének még egyelőre határt szab az átviteli utak keskeny keresztmetszete és lassúsága. Ezen a gondon segíthet a kábeltelevíziós társaságok által nyújtott Internet szolgáltatás, amely már elegendő sávszélességet fog biztosítani.

Számos könyvet, lexikont már nem csak nyomtatott formában olvashatunk, hanem az Interneten keresztül is. Óriási méretű adatbázisok állnak rendelkezésre, ahol kedvünkre kereshetünk.

Az egyetemek és főiskolák oktatói között elterjedőben van az, hogy hozzáférhetővé teszik az Interneten keresztül is az általuk oktatott tantárgyak tematikáját, felkészülési kérdéseit, előadásai anyagait, jegyzeteit. Emellett számos hallgató is közzéteszi a kész szakdolgozatát, ezzel biztosítva széles vitafórumot a benne leírtaknak.

Az Interneten keresztül országok, vidékek, városok részletes térképeit, vagy ezek műholdról készült légi fényképeit tölthetjük le. Megnézhetjük fotóművészek, fotóriporterek és amatőr fotósok fényképeit. Ellátogathatunk egy múzeumba, vagy egy kiállításra úgy, hogy a fotelünkben ülünk otthon, de ugyanígy sétát is tehetünk egy-egy nagyváros utcáin is.

Az Interneten keresztül szinte minden nyílt forráshoz hozzá lehet férni részben, vagy teljes egészében. A jövőben nagy valószínűséggel folyamatosan fog növekedni az Interneten hozzáférhető ilyen információk és lehetőségek száma. A nem is olyan távoli jövőben szinte minden elérhető lesz majd számítógépes hálózaton keresztül is.

Az Internet tehát egy nagyon jól használható, könnyen elérhető, olcsó információforrás, ami képes lehet annyi adatot szolgáltatni, ami jelentős részét teheti ki egy adott területről szerezhető összes információnak. Az Internet az információ Nirvánája, óriási, sokkultúrájú könyvtár, ami nyitva van éjjel, nappal, a nap minden órájában az egyszerű számítógép felhasználónak.

Az internetes információszerezés hátrányai

Ne felejtsük el, hogy az Internet csak egy a nyílt források között. Több ezer adatbázis csak a kereskedelmi forgalomban hozzáférhető. Ezek a kereskedelmi információforrások gyakran jobban szervezettek és indexeltek, mint az Interneten keresztül elérhetők, igaz sokszor jóval drágábbak is azoknál. Nagyon fontos megjegyezni, hogy a gyors növekedés ellenére, a nyílt forrásoknak csak 10 százaléka érhető el jelenleg ebben a formában, a többi csak nyomtatásban áll a rendelkezésünkre.

Az Interneten lévő összes anyag mennyisége feltételezések szerint nagyjából kettő, vagy három átlagos (300 000 kötetes) könyvtárnak felel meg. A legnagyobb nyílt források tárházának ma az Egyesült Államok Kongresszusi Könyvtára tekinthető, ahol több mint 110 millió könyv, újság, folyóirat, mikrofilm, és egyéb speciális formátumú anyag található meg 470 nyelven, valamint közel 200 000 aktuális periódikum olvasható, amelyek közül 80 000 nem angol nyelvű.

Szemben egy könyvtárban található anyagokkal, az Interneten az egyes oldalak tartalma folyton változik. Oldalak megjelennek és váratlanul, minden figyelmeztetés nélkül eltűnnek. Egyes szakértők szerint a hálózaton lévő dokumentumok átlagos élettartama csak kb. 30 nap.

Az Interneten lévő adatbázisokat nem gondolja senki, szemben a könyvtárakkal és a kereskedelmi adatbázisokkal. Nem létezik olyan katalógus rendszer, aminek a segítségével az anyagok könnyen gyorsan megtalálhatók lennének. Az internetes keresés jártasságot, kitartást és nagyon jó kereső eszközöket igényel.

Az Internetes keresés hasonló ahhoz, amikor az ebédszünetünkben elugrunk egy nagy áruházba, hogy például megvegyünk egy cipőfűzőt. Tudjuk, hogy mit akarunk venni, de számos dolog vonja el a figyelmünket az eredeti tervtől. Először megérezzük a sülő pizza szagát, így veszünk belőle egy szeletet. Azután

néhány érdekes könyvet veszünk észre a könyvesbolt kirakatában, így eltöltünk egy kisidőt azok lapozgatásával. Ezután veszünk még valamilyen nyalánkságot, majd egy csinos pólót. Másfél óra elteltével sietünk vissza dolgozni teletömött hassal és egy kisebb csomaggal, de cipőfűző nélkül.

Az Internetre szinte akárcsak feltehető adatokat. Vannak olyan szolgáltatók, ahol ingyen és bérmentve lehet tárhelyhez jutni, ahova el lehet helyezni akármilyen fájlt, adatot, képet, stb. A kiszolgáltatók nem felelnek az így közreadott adatok tartalmáért, tényyszerűségéért. Kizárólag csak akkor avatkoznak be, és szüntetik meg a weboldalt, ha azzal valamilyen törvényt, vagy előírást sértenek meg (pl. uszít valamelyik kisebbség ellen). Sokan ki is használják ezt a lehetőséget, mert ez nagyon olcsó módja az információ terjesztésének. Az információt feltöltő egyéneket, szervezeteket nagyon nehéz arra kötelezni, hogy csak az igazat, a színtiszta igazat adják közre. Sokszor gyakorlatilag utolérhetetlen, hogy ki helyezett el adatokat egy adott helyen, mert a szolgáltatók az általuk kért adatokat nem ellenőrzik vissza, így bármilyen fals adatot megadhatnak a felhasználók (pl. nem létező email címet) az azonosításukra. Ezért szinte minden következmény nélkül közölhetnek bárkiről, vagy bármiről bármit. Ez — valljuk meg őszintén — nem minden esetben rossz, hiszen ezért tarják sokan az Internetet a legdemokratikusabb információforrásnak, mert itt nem lehet — vagy csak igen körülményesen — megakadályozni a vélemények szabad áramlását.

Az Internet tehát egy olyan információforrás, aminek a tartalma gyorsan változik, nagyon rosszul szervezett, valamint kétes hitelességű. Felhasználásához nagyon pontosan meg meghatározott keresési terv és egyre kifinomultabb számítógépes eszközök szükségesek.

ÖSSZEGZÉS

A nyílt forrásokból, így az Internetről szerzett adatok, nagyon fontos kiegészítői az olyan titkos eljárásokkal szerzett adatoknak, mint a humán- (HUMINT), a képi- (IMINT) és rádióelektronikai- (SIGINT) felderítés. Ezekből a forrásokból első sorban a tendenciákra, fejlődési irányokra lehet következtetni. A megszerzett adatokat a felhasználás előtt fokozottan szükséges ellenőrizni, mert azok nem minden esetben felelnek meg a teljes valóságnak, hiszen számos esetben szándékosan elferdített híreket tesznek közzé megtévesztésül. Erre nagyon jó példát láthattunk a koszovói válság idején, mikor is a szerbek sok olyan hírt, adatot, képet adtak közre az Interneten, ami rossz színben tüntette fel a NATO csapatok tevékenységét.

Sok esetben kizárólag csak a nyílt források állnak rendelkezésünkre ahhoz, hogy egyáltalán valamilyen információt tudjunk szerezni egy esetleges krízishelyzetben. Az így szerzett információk arra feltétlenül alkalmasok, hogy megérthessük és figyelemmel kísérhessük egy ország társadalmának politikai, gazda-

sági, katonai terveit, esetleges szándékait. Segítségükkel előre lehet jelezni a várhatóan bekövetkező krízishelyzeteket, forradalmakat, válságokat.

A felderítésben az Internetről összegyűjtött adatok referencia anyagként szolgálhatnak a más felderítési forrásból származó adatokkal való összevetéshez. Amennyiben egy más módon — pl. rádiófelderítéssel — szerzett információra ráerősít az Internetről származó adat, akkor nagy valószínűséggel megbízhatunk annak hitelességében. Ilyen módon, hosszú távon kialakíthatunk egy képet arról, hogy mely internetes forrás ad megbízható és melyik ad megbízhatatlan információkat.

Az Interneten óriási mennyiségű adat áll a rendelkezésünkre. Ebből természetesen csak nagyon kevés az, ami a részünkre hasznos tartalommal bír. Fontos megjegyezni, hogy sok esetben csak nagyon rövid ideig állnak rendelkezésre, mert lehet, hogy csak véletlenül kerültek fel a hálóra, valakinek a tévedéséből, vagy nem ritkán rosszakaratából. Ezért szükség van arra, hogy a nap 24 órájában folyamatosan figyeljük az Internetet. Folytonosan keressünk új információkat, adatokat, fájlokat, stb., illetve ellenőrizzük az addig már megismerteket. Ezt nyilván csak számítógéppel és erre alkalmas szoftverrel lehet realizálni. Abban az esetben, ha valamilyen új forrást fedezünk fel, vagy valamilyen változást észlelünk, azonnal le kell menteni az adatokat, mert azok bármikor eltűnhetnek. Az adatok tárolásához megfelelő méretű tárolókapacitásra van szükség.

A megszerzett nagymennyiségű adatot ezek után fel kell dolgozni. A feldolgozást egy erre alkalmas számítógépes programmal lehet kezdeni, amely témakörönként szétválogatja az anyagot. Ezután már emberi elemzéssel lehet véglegesen elönteni azt, hogy mit lehet felhasználni a megszerzett anyagból és mit nem.

Az Internet olyan információforrás, amivel számolnunk kell most, és a jövőben egyre inkább. A felderítéssel foglalkozó szervezetek nem engedhetik meg maguknak azt a luxust, hogy nem folytatnak adatszerzést az Internetről. Az Internetről való adatszerzés hasonló az aranymosáshoz, ahol sok köbméter földet mozgatunk meg, hogy kinyerjük azt a néhány grammnyi aranyat.

FELHASZNÁLT IRODALOM

- [1] BJORE, Mats: Six Years of Open Source Information
<http://www.oss.net/Proceedings/ossaab/aab1/aab1an.htm> 1998.05.24.
- [2] Dr NYE, Joseph: the Chairman of the National Intelligence Council, speaking to members of the Security Affairs Support Association at Fort Meade, Maryland, on 24 April 1993.
- [3] STUEDEMAN, William: Teaching the Giant to Dance: Contradictions and Opportunities in Open Source within the Intelligence <http://www.oss.net/Proceedings/ossaaa/aaa1/aaa1ae.htm> 2000.09.20.
- [4] TOMPKINS, Alan D.: The Internet: Resource or Quagmire??
<http://www.fas.org/irp/agency/army/tradoc/usaic/mipb/1997-1/tompkins.htm> 1999.05.19.
- [5] Dr. VÁRHEGYI István - SALLAI József: Információs hadviselés, vezetési hadviselés alapjai (IW, C2W), BJKMF főiskolai jegyzet, Budapest, 1998.