



ZRÍNYI MIKLÓS NEMZETVÉDELMI EGYETEM
Hadtudományi Doktori Iskola
Alapítva: 2002-ben

Ferenczy Gábor Zoltán okleveles mérnök alezredes

**INTERNET ALAPÚ NYÍLT
INFORMÁCIÓSZERZÉS
ELVI RENDSZERTECHNIKAI
MEGVALÓSÍTÁSA**

Doktori (PhD) értekezés

Tudományos témavezető:

Dr. Várhegyi István nyá. ezredes
címzetes egyetemi docens, a hadtudományok PhD doktora

BUDAPEST
2007

TARTALOMJEGYZÉK

Bevezetés.....	4
A tudományos probléma megfogalmazása.....	5
Kutatói hipotézisek.....	6
Kutatási célkitűzések.....	8
Kutatási módszerek.....	10
1. A nyílt forrású felderítés alapfogalmai, forrásai és azok képességei.....	11
1.1. A nyílt forrású felderítés helye a felderítő tevékenységen belül.....	11
1.2. A nyílt forrású információszerezés alapfogalmai.....	13
1.3. A nyílt források jelentőségének változása.....	14
1.3.1. A költségvetés változásának a hatása.....	14
1.3.2. A nyílt információk értéke.....	15
1.3.3. Változások az információszerezés terén.....	17
1.3.4. A nyílt forrású információszerezés előnyei.....	18
1.3.5. A nyílt források korlátai.....	21
1.4. A nyílt információk felhasználásának szintjei.....	22
1.5. A nyílt forrású információszerezés csatornái.....	24
1.5.1. Hagyományos média.....	24
1.5.2. Szürke irodalom.....	25
1.5.3. Szakértők és megfigyelők személyes tapasztalatai.....	26
1.5.4. Kereskedelmi műholdas felvételek.....	27
1.5.5. Kereskedelmi online információszolgáltatók.....	27
1.5.6. Internet.....	28
1.5.7. Egyéb források.....	29
1.6. A nyílt elektronikus információk jelentőségének és szerepének változása.....	31
1.7. Az internet alapú információszerezés kapcsolata a többi nyílt forrású csatornával.....	31
1.7.1. Az internetes információszerezést nehezítő tényezők.....	36
1.8. Összegzés és következtetések.....	37
2. Az internet és a keresési technikák fejlesztésének irányai.....	40
2.1. Az internet technikai fejlesztése.....	40
2.1.1. Az Újgenerációs Internet (NGI).....	41
2.1.2. Az Internet2.....	42
2.2. Az internet protokoll fejlesztése.....	44
2.2.1. Az IPv4 problémái.....	44
2.2.2. Az IPv6 jellemzői.....	45
2.3. A világháló keresési technikái.....	49
2.3.1. Az internet mérete.....	49
2.3.2. Kereső szolgáltatások.....	51
2.3.3. A keresőeszközök csoportosítása.....	52
2.3.4. A keresőmotorok.....	53
2.3.5. Témakatalógusok.....	57
2.3.6. Blogkeresők.....	58
2.3.7. A mély web.....	59
2.4. Az internetes keresés fejlesztése.....	61
2.4.1. A szemantikus web.....	62
2.4.2. A szemantikus web modellje.....	63
2.5. Összegzés és következtetések.....	64
3. Az internet alapú információszerezés rendszertechnikai modellje.....	67
3.1. Az internetes információszerezés jelentősége, szerepe.....	67
3.2. Az internet csatlakozás.....	69
3.2.1. A csatlakozás módjának megválasztása.....	70
3.3. Az INFF modell rendszertechnikai felépítése.....	70
3.4. A rendszer biztonsági kérdései.....	73
3.4.1. Tűzfal.....	73
3.4.2. Behatolás elleni védelem.....	74
3.4.3. Vírusvédelem.....	75

3.4.4.	Identitás elrejtése (anonimizálás).....	75
3.5.	Az adatszerző alhálózat	77
3.5.1.	Forráskutató munkahely.....	77
3.5.2.	A forrásértékelő munkahely.....	79
3.5.3.	Forrás-adatbázis	83
3.5.4.	Letöltés-vezérlő.....	84
3.5.5.	Kiértékelő munkahely.....	85
3.5.6.	Adatszerző-adatbázis	86
3.5.7.	Adatszerző-webszerver	88
3.6.	Adatfeldolgozó alhálózat	88
3.6.1.	Elemző munkahelyek.....	89
3.6.2.	Feldolgozó-adatbázis	89
3.6.3.	Feldolgozó-webszerver	90
3.7.	Végfelhasználói alhálózat	91
3.8.	Útvonalválasztók	91
3.9.	A rendszer működése.....	91
3.10.	A szükséges géppark.....	93
3.10.1.	Számítógépek konfigurációja.....	93
3.10.2.	Számítógépek száma	94
3.10.3.	Operációs rendszer	95
3.11.	Humán erőforrás	96
3.12.	A szolgáltatott információk mennyisége.....	97
3.13.	A megszerezhető információk típusai.....	98
3.14.	A megfigyelésre érdemes oldalak.....	98
3.15.	A rendszer alkalmazásának szintjei	99
3.16.	A rendszer kialakításának várható költségei	99
3.17.	Összegzés és következtetések.....	100
4.	Az értekezés tézisei és végkövetkeztetései, új tudományos eredmények	102
4.1.	Új tudományos eredményeim	106
4.2.	További kutatási és felhasználási lehetőségek	107
	Felhasznált irodalom	109
	Az értekezéssel kapcsolatos publikációim.....	115
	1. számú melléklet.....	118
	2. számú melléklet.....	121
	3. számú melléklet.....	123
	4. számú melléklet.....	125
	5. számú melléklet.....	128
	6. számú melléklet.....	129

BEVEZETÉS

Mindenkinek szüksége van valamilyen információra, mindig, mindenhol és mindenkor. És ez így volt az emberiség egész történelme során, de az információk típusai, azok tárolási módja és módszerei, illetve az azokhoz való hozzájutás képességei és lehetőségei alapvetően megváltoztak. A ma emberének az információ tárolására számtalan eszköz áll rendelkezésére, a papírtól a filmen keresztül, a mágneses- és optikai tároló eszközökig. Az információhoz való hozzájutás is sokkal könnyebb, mert ehhez számos olyan modern elektronikus eszköz áll rendelkezésre, mint a telefon, a rádió, a televízió és az utóbbi időben a számítógépes világháló, az *internet*.

Az információszerzésre az életben- illetve a versenyben maradás érdekében van szükség, mert akinek nem áll rendelkezésére elegendő, időszerű, megbízható és fontos (releváns) információ az hátrányba kerül a többiekkel szemben. Ez az oka annak, hogy minden lehetséges forrást fel kell használni az információ megszerzésére.

Az információ szerepe folyamatos változáson ment keresztül. Ma már nem egyszerűen csak a hírt jelenti, hanem annál sokkal többet: olyan felhasználható és kiaknázható erőforrás, ún. értéknövelő (a gazdaságban hatékonyságnövelő) tényező, amit a természetben, társadalomban, gazdaságban, technikában stb. keresni kell. A változások egyik fő mozgatója az *információs forradalom*, ami a számítástechnika és a távközlés területén végbement mennyiségi és minőségi növekedést, valamint e két terület egymáshoz való rohamos közeledését, az alkalmazásokban való együttes megjelenését és integrációját jelenti.

Az emberi civilizációkban és kultúrákban olyan mélyreható változások következtek be, amelyek gyökeresen megváltoztatták a társadalom anyagi, szellemi és szolgáltatási termelési módját. *Beléptünk az információs ipari termelési világkorszakába, amely felváltotta a hagyományos gépi ipari termelési világkorszakot. Ennek következtében kialakul a fejlett tudományra alapozott, intenzív tudásgazdaság és tudástársadalom, közismert megnevezéssel az információs társadalom*¹.

Az információs társadalom a nyíltságon, az információk szabad áramlásán alapul, ezért a nyílt információk válnak dominánssá. A nyílt információk szerepének növekedését az információterjesztés eszközeiben és technikájában végbement ugrásszerű fejlődése tette lehetővé. A nem is olyan távoli múltban még csak az újságokból, vagy más nyom-

¹ Várhegyi I. – Makkay I.: Információs korszak, információs háború, biztonságkultúra, OMIKK, Budapest, 2000, 9. old.

tatott anyagokból lehetett friss hírekhez jutni. Az elektromosságnak a hírközlésben való felhasználása eredményeként egyre újabb találmányok (távíró, telefon) könnyítették meg az emberek közötti kommunikációt. A nyílt információk mennyiségének megnövekedését a rádiózás, de főként a televíziózás elterjedése okozta. Az információ terjesztésének igazi forradalmát azonban az internet széleskörű elterjedése hozta el. A 90-es évekre már világméretűvé vált számítógépes világhálózat segítségével az információhoz való hozzájutás már olyan egyszerű lett, hogy ma már nem az információ hiánya, hanem annak túlzott mennyisége okoz problémát.

A TUDOMÁNYOS PROBLÉMA MEGFOGALMAZÁSA

Az információs társadalomban a hadsereg is mélyreható változáson megy keresztül. Ez a változás alapvető hatással van a hadtudományra és a hadügy gyakorlatára egyaránt. *Az információ társadalmi, műszaki értékének és jelentőségének forradalmi megváltozásával összefüggésben, valamint egy természetes és gyors evolúciós fejlődés eredményeként kialakultak az információs műveletek², ami az információs fölény, az időfölény és a vezetési fölény³ kivívására irányuló, egységes elgondolás szerint kialakított, feszített szellemi és fizikai küzdelem, amelyet a szemben álló katonai felek magasan képzett vezetőállománya, valamint e sajátos küzdelembe bevont, a különböző műveleteket kivitelező információs harcosok folytatnak⁴.*

Az információs műveletek elképzelhetetlenek időbeni, fontos és hiteles felderítési adatok nélkül. Az időszerű és megbízható információ biztosítása a parancsnok számára a korszerű műveletek során minden katonai szervezet egyik legalapvetőbb feladata. Az információt számos forrásból lehet megszerezni, amelyek nyíltak és titkosak lehetnek.

A 80-as évek végétől elindult változások nyomán megszűnt a bipoláris világrendszer, véget ért a két katonai tömb szembenállása, az állandó háborús fenyegetettség. Ennek hatására az addig a „vasfüggöny” mögé zárkózott országok megnyitották a határaikat, és nyílt társadalmat kezdtek kialakítani. Ez az újfajta nyitottság nagyszámú új, nyílt információforrás megjelenését eredményezte. Számos új napilap, hírlap, magazin, periodikum, könyv jelent meg, valamint új rádió- és TV adó kezdte el a műsorsugárzást.

² Információs műveletek – Information Operation (IO).

³ Várhegyi I. – Makkay I.: Információs korszak, információs háború, biztonskultúra, OMIKK, Budapest, 2000, 125. old.

⁴ Információs harcos – information warrior.

A XXI. század elejére végbement technikai fejlődés jellemzője, hogy az információs technológiák integrációjának következtében az információáramlás egyre inkább elektronikai adathordozókon keresztül valósul meg. Ez a megállapítás fokozottan igaz a nyílt információforrásokra. Ez az oka annak, hogy az információs műveleteknek az elektronikus úton végzett nyílt információszerzés nagyon fontos alkotó elemévé válik.

A nyílt információk gyűjtése mindig is a hírszerző munka részét képezte, de az elektronikus úton, azon belül is a számítógépes hálózatok révén megszerezhetőké még nem tekint vissza hosszú múltra. Így csak az utóbbi néhány évben kezdtek megjelenni az internet figyelésével kapcsolatos szakirodalmak. Az Egyesült Államokban lefolytatott kutatások alapján egy kézikönyvben foglalták össze ennek elvi végrehajtását⁵. Ebben az anyagban azonban nem adják meg azt, hogy az internet alapú nyílt információszerzést milyen rendszertechnikai megoldásokkal, eszközökkel, szervezettel lehet megvalósítani. Ennek kidolgozását mindenhol a saját nemzeti szervezetekre bízzák. Ezzel foglalkozó cikkek, tanulmányok vagy szabályzatok nincsenek. Követelményként fogalmazódott meg, hogy az internet alapú információszerzés eredményes és gyors végzéséhez elkerülhetetlen egy rendszertechnikai modell kidolgozása, és ez alapján egy nagymértékben automatizált, modern eszközökből, korszerű elven megvalósított internet figyelő rendszer kialakítása. Ez felválthatná a jelenlegi kevésbé automatizált, nem egységes szervezetben végzett, egyénileg folytatott internetes információszerzést.

KUTATÓI HIPOTÉZISEK

1. A Magyar Honvédség számára mindennapi realitássá vált – a NATO csatlakozás következményeként –, hogy részt vállaljon a világban kipattanó *válságok kezelésében, béketámogató műveletekben*. Ezeket a feladatokat sok esetben olyan térségben kell végrehajtani, amely korábban nem volt a felderítés fókuszában (pl. Afganisztán, Irak). Ilyen esetben a katonai műveletek (a missziós tevékenység) megtervezéséhez, valamint azok kezdeti szakaszában, jelentős mértékben lehet *a nyílt információkra támaszkodni*, mert azokat azonnal be lehet szerezni, a nap bármelyik pillanatában *könnyen, gyorsan és olcsón hozzáférhetők*. A nyílt források szerepének felértékelődését az információszerzésben végbement gyökeres változások okozták. Amíg a hidegháborús szembenállás időszakában a nyílt források inkább kiegészítő információkat szolgáltatottak, addig napjainkban előfordulnak olyan esetek (pl. vál-

⁵ Intelligence Exploitation of The Internet, Saclant, Norfolk, 2002.

ságkezelés), amikor már alapinformációk is megszerezhetők ezekből a forrásokból. Az információs társadalomban tapasztalható információrobbanás eredményeként a nyílt információk mennyiségileg lényeges többségbe kerültek a minősítettekkel szemben, ezért egyre több információigényt lehet ilyen forrásokból kielégíteni. Megfelelő munkamódszerek alkalmazásával a felkutatott nyílt információk legalább olyan *hitelesek, pontosak* és is *időszerűek* lehetnek, mint a minősített információk, mindemellett *használatuk egyszerű és kockázatmentes*.

2. Az *információt* nem elég megszerezni, *fel is kell tudni azt dolgozni*. Az elektronikában, az elektrotechnikában és az informatikában az utóbbi évtizedekben végbement fejlesztések eredményeként az információk többsége már digitális elektronikus formában keletkezik, vagy ilyenné alakítják át. Ennek a változásnak a kiváltó oka, hogy azok az információk kezelhetők a *leghatékonyabban*, amelyek ilyen formában állnak rendelkezésre. Az internetről szerezhető információk eleve digitális elektronikus formában érhetőek el, ami lehetővé teszi a feldolgozásuk azonnali megkezdését. A feldolgozó munka a rendelkezésre álló *számítógépen futó célprogramok* segítségével nagymértékben gépesíthető és automatizálható.
3. Az *információt*, annak megszerzése és feldolgozása után *el kell juttatni a felhasználókhoz*. Az információt *elektronikus úton* lehet a *leggyorsabban és leggazdaságosabban a legtöbb felhasználóhoz* eljuttatni. Mivel a nyílt információk terjesztésének a legolcsóbb és legalkalmasabb módja a digitális elektronikus jelek *számítógépes hálózaton való továbbítása*, ezért *az internetre tevődnek át a nyílt információs csatornák*, amire ma már számos példa van: adatbázisok, újságok, könyvek, múzeumok, oktatási és kulturális anyagok, stb. Ez a tendencia a jövőben is tovább folytatódik. Emellett az internetbe integrálódik *a kommunikáció és a műsorsugárzás*, így a távközlés (telefon), a rádió és a televízió is. Ennek következtében *a világháló válik a legnagyobb nyílt információforrássá*, amit egyre nagyobb mértékben lehet és kell felhasználni információszerzésre, ezért azt folyamatosan – a nap huszonnégy órájában – figyelni kell.
4. Az internet figyeléséhez olyan *állandó szervezetre van szükség, amelynek rendszer-technikailag megfelelően kialakított – magas technikai szintet képviselő – eszközpark áll a rendelkezésére*. Ennek létrehozása napjaink sürgető és elodázhatatlan feladata, mivel az internettel, mint a nyílt, elektronikus hálózati információkhoz való hozzáférés eszközével, szerzett információkkal – a más forrásokból származó információkkal összehasonlítva – jelentős mértékben *felgyorsítható a döntéshozatal*,

ami hozzájárul a döntési időnyeréshez. Az internetet figyelő rendszer minél előbbi kialakítása lehetővé teszi a szakembereknek a minél gyorsabb és *hatékonyabb* internetes *információszerző tevékenység* szakmai fogásainak elsajátítását, az adatbázisok megfelelő mennyiségű anyagokkal való feltöltését, valamint a megbízható információforrások megtalálását. Emellett a megfelelően kialakított információszerző és feldolgozó rendszer nem csak az információ megszerzésére és feldolgozására, hanem a feldolgozott anyagoknak a *felhasználókhöz való eljuttatására is alkalmas* a meglévő hálózati kapcsolatok felhasználásával.

5. Az interneten az információk mennyisége rendkívül gyorsan növekszik. Ezek nincsenek semmilyen könyvtárba, vagy valamilyen más logikai rendszerbe szervezve, ezért a szükséges *adatok megtalálása komoly felkészülést igényel* az internetes keresési technikák használatában. Megbízható, hiteles és fontos információt csak megfelelő forrásból lehet szerezni. Az információszerzés szempontjából értékes *források kiválasztása időigényes, megfelelő ismereteket és jártasságot kíván meg* a forrásértékelést végzőtől. Felhasználás előtt az információkat értékelni kell, hogy kiszűrjük közülük a felderítési értékkel nem rendelkezőket. *Az információk értékelését gyorsan és nagy szakértelemmel kell végezni*, ami tapasztalatot és magas szintű nyelvismeretet igényel. A források különböző adatait és a megszerzett információkat *adatbázisokban* kell tárolni. Ezek *kezelését*, karbantartását hozzáértő számítógépes *szakembernek tudják csak szakszerűen végezni*. Az internet alapú nyílt információszerzés tehát egy olyan komplex feladat, amelyet csak jól felkészült, nagy tapasztalatokkal rendelkező, számítástechnikában jártas, idegen nyelveket ismerő szakemberek végezhetnek, ezért célszerű lenne ezt a tevékenységet *külön szakterületként kezelni a felderítésen belül*.

KUTATÁSI CÉLKITŰZÉSEK

A témámmal kapcsolatban a következő *fő kutatási célok*at tűztem ki magam elé:

1. Megismerni, majd rendszerbe foglalni az internet alapú nyílt információszerzéssel kapcsolatos ismereteket.
2. Bizonyítani, hogy az internet alapú információszerzés eszköztára adaptálható a nyílt forrású információszerzésre.

3. Bizonyítani, hogy az internet alapú nyílt információszerzés olyan speciális felkészültséget és ismereteket követel az azt végrehajtó állománytól, ami indokolja annak új szakterületként való kezelését.
4. Kidolgozni egy, az internet alapú információszerzés megvalósítására alkalmas, rendszermodellt, ami egyaránt képes a megfelelő információforrások felkutatására, az információk automatikus gyűjtésére, azok feldolgozásra és a felhasználókhoz való eljuttatására.

A fő kutatási céljaim elérése érdekében az alábbi *részfeladatok* elvégzését tartotam szükségesnek:

Összefoglalni a nyílt forrású információszerzéssel kapcsolatos ismereteket és alapfogalmakat. Megismerni annak helyét és szerepét az információszerző tevékenységben. Meghatározni a nyílt forrású információszerzés csatornáit, azok képességeit és kapcsolatát az internettel. Bizonyítani, hogy a napjainkban is zajló fejlesztések hatására az internet válik a legnagyobb nyílt információforrássá.

Összefoglalni az internetnek azokat a fejlesztéseit, amelyek lehetővé teszik a gyors és hatékony információszerzést, valamint az internetes keresési technikák működési elvét, tulajdonságait, korlátait, felhasználási lehetőségeit és fejlesztéseit.

Részletesen kidolgozni a rendszermodell elemeinek a rendszerben betöltött feladatát, működésük automatizálhatóságát és szerepét. Meghatározni a rendszer kialakításához és működéséhez szükséges eszközöket, humán erőforrást, valamint megbecsülni kialakításának várható nagybani költségeit.

A kutatási céljaim eléréséhez a szerteágazó szakterületek közül *nem folytattam* az alábbiak kutatását:

- ☞ az illegális módszerekkel való információszerzést;
- ☞ az internet fejlesztésének azon területeit, amelyek nem kapcsolatosak a sáv-szélesség, az átviteli sebesség és az újgenerációs internet protokoll kidolgozásával és bevezetésével;
- ☞ a számítógépes biztonságok a tűzfalak, a behatolás elleni védelem, vírusvédelem és a névtelen internetes böngészés (anonimizáció) alapismeretein kívül eső kérdéseket;
- ☞ a letöltésre kerülő anyagok feldolgozásával és rögzítésével kapcsolatos szakmai kérdéseket;
- ☞ az adatbázisok típusaival, program nyelvükkel, adattárolási megoldásaival és általában minden működésükkel kapcsolatos kérdéseket;

- ☞ a rendszer egyes elemei működéséhez szükséges konkrét operációs rendszerek és programok megválasztása kérdéseit;
- ☞ az összegyűjtött anyagokból való jelentés elkészítésének munkamódszereit, eljárásait és követelményeit.

KUTATÁSI MÓDSZEREK

Kutatási céljaim elérése, és az értekezés elkészítése érdekében az alábbi módszereket alkalmaztam:

- ☞ felkutattam és feldolgoztam a témához kapcsolódó naprakész szakirodalmat, tudományos dolgozatokat, doktori értekezéseket;
- ☞ célirányos kutatásokat folytattam könyvtárakban és az interneten;
- ☞ konzultáltam a számítástechnikában, a hálózat építésében és használatában, valamint a nyílt forrású információszerezésben jártas szakemberekkel;
- ☞ folyamatosan figyelemmel kísértem az internet fejlesztésével kapcsolatban publikált kutatási anyagokat;
- ☞ kutatásokat végeztem az interneten fellelhető információk milyenségéről, mennyiségéről és azok változásáról;
- ☞ tanulmányoztam az internetes keresési technikákat, összehasonlításokat végeztem az egyes rendszerek működésével, keresési eredményeivel és hatékonyságával kapcsolatban;
- ☞ alkalmazás közben tanulmányoztam olyan programokat, amelyek szükségesek lehetnek az általam kialakítani tervezett rendszer működéséhez;
- ☞ figyelemmel kísértem az internet fejlődését, illetve fejlődésének tendenciáit;
- ☞ rendszereztem a megszerzett ismereteket;
- ☞ a célkitűzéseimhez kapcsolódó tevékenységem és kutatásaim részeredményeit rendszeresen publikáltam magyar és angol nyelvű szakmai kiadványokban és adtam elő konferenciákon;
- ☞ az elvégzett munkáról a témában jártas szakértőktől és kollégáimtól kértem és kaptam reflektálásokat, észrevételeket, segítséget, amelyeket beépítettem az értekezésembe.

A témával kapcsolatos dokumentációk, adatok, információk gyűjtését és rendszerezését 2007. február végén fejeztem be.

1. A NYÍLT FORRÁSÚ FELDERÍTÉS ALAPFOGALMAI, FORRÁSAI ÉS AZOK KÉPESSÉGEI

A nyílt forrású felderítés/információszerzés (*Open Source Intelligence – OSINT*) előretörése a XX. század utolsó évtizedének közepén kezdődött. Ekkor kezdték el az Egyesült Államokban azt a kutatómunkát, amelynek az volt a célja, hogy megvizsgálják a nyílt forrású információszerzés lehetőségeit, korlátait és kidolgozzák annak munkamódszereit. A kutatások eredményeként az információszerzésnek ez a módja már túllépett a katonai felderítés keretein, így a nemzetbiztonsági szolgálatok, a rendvédelmi szervezetek, sőt magánvállalkozások is alkalmazzák. Mára elmondhatjuk, hogy a felderítés által megszerzett információk jelentős része nyílt forrásokból származik.

1.1. A NYÍLT FORRÁSÚ FELDERÍTÉS HELYE A FELDERÍTŐ TEVÉKENYSÉGEN BELÜL

A meghatározott feladatok sikeres végrehajtása érdekében minden lehetséges adatforrásból információt kell gyűjteni. A felderítés általános elfogadott adatforrásai a következők⁶:

- ☞ **Rádióelektronikai felderítés (REF)** (*Signal Intelligence – SIGINT*), ami a rádiófelderítést (*RÁF*) (*Communication Intelligence – COMINT*) és a rádiótechnikai felderítést (*RTF*) (*Electronic Intelligence – ELINT*) foglalja magában. Feladata a híradó rendszerek és elektronikai eszközök, valamint a légi-(űr-), föld- (víz-) felszíni és föld- (víz-) alatti kísérleti és rendszeresített rádióelektronikai eszközök által kibocsátott elektromágneses jelek mérése és elemzése.
- ☞ **Humán felderítés** (*Human Intelligence – HUMINT*). Az emberi kapcsolatok felhasználásával folytatott titkos információszerzés.
- ☞ **Képi felderítés** (*Imagery Intelligence – IMINT*). A fotografikus, radar, elektrooptikai, infravörös, hő, illetve multispektrális érzékelők által vett jelekből képzett képanyagok megszerzése és elemzése.
- ☞ **Technikai jelfelderítés** (*Measurements and Signatures Intelligence – MASINT*)⁷ olyan információk megszerzését és értékelését végzi, amelyek nem tartoznak az ellenség híradó (kommunikációs) rendszereinek felderítését vég-

⁶ A meghatározások alapját az AJP 2.0. 1-2-5 oldalán szereplő definíciók képezték.

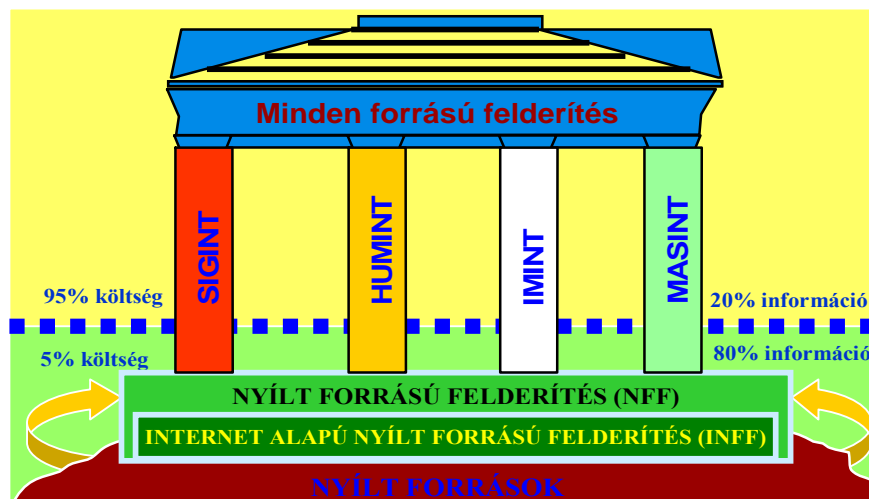
⁷ Az IMINT és a MASINT tevékenységet első sorban csak a gazdag és fejlett hadiiparral rendelkező országok alkalmazzák.

ző SIGINT feladatköréhez. Ilyenek lehetnek többek között a tömegpusztító fegyverekkel és a legújabb technológiai eredmények segítségével kifejlesztett fegyverekkel végrehajtott kísérletekből származó jelek és információk, amelyek az ellenség megnövelt harci képességeire utalnak.

☞ **Nyílt forrású felderítés (NFF)** (Open Source Intelligence – OSINT).

- **INTERNET ALAPÚ NYÍLT FORRÁSÚ FELDERÍTÉS (INFF)** (angolul Internet OSINT – I-OSINT). Ez a felderítési mód a NFF szerves része, annak egy csatornája, ami a számítógépes hálózatok szolgáltatásainak és eszközrendszerének felhasználásával szerez nyílt információt. Végzéséhez speciális ismeretekre van szükség, ezért külön szakterületként célszerű kezelni az NFF belül.

A minden oldalú információszerzésnek, amelynek részét képezi a rádióelektronikai, humán, képi és technikai felderítés (SIGINT, HUMINT, IMINT, MASINT), fontos összetevője – azok alapja – a nyílt forrású felderítés (NFF) és benne az internet alapú nyílt forrású felderítés (INFF), amely nélkül nem lehet teljes képet kapni egy adott kérdéskörben (1.1. ábra). Jóllehet a nyílt forrásokból származó információk nem helyettesíthetik a minősített információkat, de számos téren képesek kiegészíteni azokat⁸. Segítségükkel kijelölhető a költségesebb információszerzési eljárások erőfelfejtésének fő iránya, emellett gyors, naprakész információkkal tudja ellátni a felhasználókat. Távoli térségekben vezetendő katonai műveletek esetén pedig gyakran a tervezés, illetve a műveletek kezdeti időszakának fő információszerző lehetősége.



1.1. ábra Az internet alapú nyílt forrású felderítés helye a minden forrású felderítésen belül⁹

⁸ NATO OSINT Reader, Saclant, Norfolk, 2002, 10.old.

⁹ Az ábra a NATO OSINT Handbook, Saclant, Norfolk, 2001. 36. oldalán lévő ábra felhasználásával készült.

1.2. A NYÍLT FORRÁSÚ INFORMÁCIÓSZERZÉS ALAPFOGALMAI

Mindenek előtt fontosnak tartom, hogy összefoglaljam a nyílt forrású információszerzéssel kapcsolatos legfontosabb alapfogalmakat¹⁰.

Nyílt forrás (*Open Source – OS*): mindazok a nyilvánosan elérhető források, amelyekhez törvényes eszközökkel (lekérdezés, megfigyelés, előfizetés, megvásárlás, stb. útján) lehet hozzáférni. Más szavakkal nyílt forrásnak minősül minden olyan eszköz, hely, kiadvány, műsor, amihez hozzá lehet férni a kereskedelmi-kiadási terjesztésben, illetve az ezeken kívül eső nem minősített anyagok, amelyekhez nem tiltják másoknak a hozzáférést.

Nyílt forrású adat (*Open Source Data*): olyan nyers (még nem feldolgozott), nyomtatott, kisugárzott, szóban közölt vagy más formájú dokumentum, tény, ismeret, stb., amit az elsődleges források szolgáltatnak. Ez lehet például fénykép, magnófelvétel, műhold által készített kép vagy valakinek a személyes levele. Másképpen megfogalmazva: a nyomtatott, képi vagy jel formátumban megjelenő, ismétlődő, megerősítő vagy új ismereteket tartalmazó, de még nem értékelt anyag.

Nyílt forrású információ (*Open Source Information – OSIF*): olyan nyílt forrású adatok, amelyeket adott szemlélet alapján gyűjtenek, értékelnek és feldolgoznak. Általában olyan széles körben terjesztett anyagok, mint a könyvek, újságok, rádió- és televízió-műsorok, stb.

Nyílt forrású információszerzés (*Open Source Intelligence – OSINT*): olyan információgyűjtő eljárás, amely során a nyilvánosan (a publikum számára) elérhető forrásokból az információkat felkutatják, elemzik, értékelik és felhasználják egy adott cél elérése érdekében, általában a parancsnok és annak közvetlen törzse, által feltett kérdés megválaszolására. Más szavakkal az információszerzés kipróbált eljárásainak alkalmazása a széles körben hozzáférhető nyílt adatforrásokra. A nyílt forrású információszerzés nem kizárólag katonai felderítési vagy információszerzési kategória, mivel ezt a tevékenységet a civil szférában is folytatják.

Nyílt forrású felderítési adat: a nyílt forrásokból, a felderítés aspektusából értékelt információ, amely a felhasználó által meghatározott információigény kielégítésére alkalmas. A nyílt forrású felderítési adat tartalmazza mindazokat a tényeket, amelyeket a nyílt forrásokat felderítő szervek (adatforrások) összegyűjtöttek és az azok felhasználásával született következtetéseket és értékeléseket.

¹⁰ A meghatározások alapját a NATO Osint Handbook, Saclant, Norfolk, 2001. 2-3 oldalán lévő meghatározások képezték.

1.3. A NYÍLT FORRÁSOK JELENTŐSÉGÉNEK VÁLTOZÁSA

A hidegháború megszűntével a világ teljesen átformálódott. A kétpólusú világ egypólusúvá vált, már csak egy szuper hatalom van, az Egyesült Államok. A csökkenő feszültség hatására világszerte elindult a hadseregek létszámának csökkentése. Az országok egyre nyitottabbak lettek. Az addig szigorú titokként őrzött katonai információk nagy részét nyilvánossá tették, minek következtében a korábban nagy költségeket felemésztő titkos felderítő eljárások veszítettek jelentőségükből, mert az információk többsége már nem minősített, így valamilyen nyílt forráson keresztül is elérhető. Emiatt felértékelődtek a nyílt forrású anyagok, amelyek mindennap óriási mennyiségben keletkeznek szerte a világon.

A nyílt források mennyisége eddig példa nélkül álló módon megnőtt. Világszerte *a napi-, heti-, havilapok és a különböző periodikák száma az 1972-es 70 ezerről több százezerre növekedett*¹¹. A robbanásszerű növekedés a volt Szovjetunió államaiban figyelhető meg a legjobban, ahol több ezer olyan újságot adnak ki, amelyek a rendszerváltozás előtt egyáltalán nem léteztek.

A nyomtatott sajtó mellett az elektronikus média (rádió, televízió) is széles körben elterjedt. Ma már elképzelhetetlen a mindennapi élet rádiózás, de még inkább televíziózás nélkül. A kezdeti egy csatorna helyett ma már minden országban számos földi és műholdas csatorna adása fogható, akár a világ másik végéről is.

A hagyományos média mellett az internet megjelenése és térhódítása is hozzájárult a nyílt információk mennyiségének növekedéséhez. Az interneten keresztül az élet szinten minden területéről lehet valamilyen információt szerezni. Ma már több ezer online kereskedelmi adatbázist kínálnak a felhasználóknak, amelyek nagy többségének komoly felderítési értéke van.

1.3.1. A költségvetés változásának a hatása

A nyílt forrású információk felértékelődését a felderítő szervezetek anyagi helyzetében végbement változások is meghatározzák. A rendszerváltozás óta a katonai, és ezen belül természetesen az információszerzésre fordítható *költségvetés* is, nagymértékben *lecsökkent*. A „csináljunk többet kevesebből” elv jellemzi a helyzetet a felderítő szervezetekben. Ha ugyanazt az információt tudjuk megszerezni egy külföldi újságból vagy televí-

¹¹ NATO OSINT Reader, Saclant, Norfolk, 2002, 9.old.

zióból, illetve egy drága műholdon telepített érzékelőből, akkor nem lehet kérdés, hogy melyik a költséghatékonyabb forrás.

A költségvetés csökkenésével párhuzamosan a hadsereg létszámában, és ezen belül a felderítő szervezetekben is, mélyreható változások mentek végbe. Az egykori tömeghadseregből egy jóval kisebb profi hadsereg létrehozása lett a cél. Így a felderítésnek a feladatát *kevesebb emberrel, de továbbra is hatékonyan* kell ellátnia. Ez csak *a nyílt források fokozott felhasználásával és munka nagy részének gépesítésével oldható meg.*

A nyílt információk növekvő részben elektronikus, pontosabban digitális elektronikus formában keletkeznek, ezért az interneten keresztül is elérhetőek. Az így megszerzett anyagok tárolásához, feldolgozásához, elemzéséhez és kiértékeléséhez rendelkezésre állnak a korszerű, gyors számítógépek. A számítástechnika alkalmazása hatékonyabbá, gyorsabbá teszi az információszerzést és –feldolgozást, kiválthatóvá teszi az olyan feladatok végrehajtását, amelyek nem igényelnek humán intelligenciát (pl. keresés, figyelés, letöltés). Az információszerzésnek ez a formája *kis költséggel tud nagy mennyiségű hiteles információt szolgáltatni*, mert a kiépítése után már nem merülnek fel nagyobb kiadások. Nem kell minden évben lecserélni a számítógépeket és a bevált szoftvereket, mert azok nem avulnak el ilyen gyorsan. A működtetés során felderített és szükségesnek tartott forrásokra való előfizetés jelenthet újabb költséget.

1.3.2. A nyílt információk értéke

Még napjainkban is vannak olyan információszerzéssel foglalkozó szakemberek, akik nem ismerik fel a nyílt információk jelentőségét. Véleményük szerint a felderítésnek nem a nyílt információk gyűjtése a feladata, hanem a minősített információk megszerzése. Ezzel kijelentéssel az a probléma, hogy figyelmen kívül hagyja azt, hogy a nyílt információkat mindig is gyűjteni kellett. Emellett az információrobbanás következtében a nyílt információk mennyiségileg lényeges többségbe kerültek, ami azzal jár, hogy egyre több információigényt lehet nyílt információkból kielégíteni.

A *minősített információval* kapcsolatban sokáig az volt a vélekedés, hogy az a maga hozzáférési és kezelési korlátozása miatt megbízható minőségű és a legmegfelelőbb a valós helyzet megjelenítésére. Azonban ez nem minden esetben van így. *Tartal-mazhat apró ferdítéseket vagy elhanyagolásokat* éppen úgy, mint más források. Például a történelem bebizonyította, hogy a nyugati felderítő szervezetek sok mindent rosszul feltételeztek a Szovjetunió és a Varsói Szerződés államainak katonai és gazdasági erejéről.

A minősített anyagok előtérbe helyezése a hidegháborúra vezethető vissza. Az a titokzatosság, ami ezt az időszakot körülvette, néhány elemzővel azt hitette el, hogy csalhatatlan, s büszke volt a különlegességére. Ennek oka az, hogy a rejtjelző- és egyéb különleges berendezések, valamint a biztonsági eljárások használata egyfajta fontosság és hozzáértés érzetét adják az azt alkalmazóknak. Azonban a mai konfliktusokra nem lehet a régi eljárásokat alkalmazni. Mialatt reggelizünk és hallgatjuk a híreket arról, hogy forradalom tört ki valamelyik országban, nem várhatunk arra, hogy majd a titkos hírszerző rendszeren keresztül megtudjuk, hogy mi is történt ott „valójában”.

A hidegháború idején nem volt természetes és elfogadható, hogy a *minősítetteknél jobb, vagy legalább olyan jó információk szerezhetők a nyílt forrásokból*. Ma már természetes és magától értetődő, hogy egy napilapban az adott országban lezajlott legutóbbi választásokról megjelent áttekintő elemzés, vagy a térségben régóta fennálló etnikai feszültségekről szóló cikk sokkal *tisztább képet ad a valós helyzetről*, mint egy magas minőségű olyan jelentés, ami részletesen megadja a kézifegyverekbe való lőszer pontos számát, amit valahol egy titkos raktárban őriznek, vagy az a tény, hogy csak 57 helikopter működik a lehetséges 64 közül.

A nyugati országokban, elsősorban az Egyesült Államokban, erős törekvés van a nyitottság irányába. Ez fokozottan igaz a kormányzatokra, ahol sok titkot és titkos technológiát felszabadítottak a minősítés alól. Ugyanígy az iparban (kivéve a hadiipart) is egyre növekszik a felismerése annak, hogy az információk elzárása jelentősen megnöveli a belső költségeket, emellett nagyon nehezen lehet biztosítani. Speciális kémiai receptek vagy más „védhető” titkok esetében ez természetesen nincs így, de az általános mérnöki gyakorlatban, kereskedelemben is elfogadottan a nyíltságra összpontosítanak, és inkább időben megelőzik az ellenfeleket, mint a titkokat védik. A versenyelőny fokozatosan eltolódik a titkosságról a nyíltságra, a „titokban” végzett kutatásokról a „gyorsan” kiaknázott lehetőségekre. Ez lehetőséget nyújt a „tájékozódást” lehetővé tevő és a nyíltan rendelkezésre álló információk azonnali összegyűjtésére.

Fontos megjegyezni, hogy a nyílt forrású információszerzés nem képes helyettesíteni a minősített információkat. *Dr. Joseph S. Nye* az Egyesült Államok *Nemzeti Felderítő Tanácsának*¹² egykori vezetője sok esetben használta a kirakós játék (puzzle) hasonlatot a nyílt forrású felderítés és a többi hagyományos felderítő eljárás közötti összefüggés leírására:

¹² A National Intelligence Council feladat, hogy koordinálja az elnök részére készülő tömör hírszerzői jelentéseket.

„A nyílt információforrások felderítéséből kapjuk meg a kirakó játék külső darabjait, amelyek nélkül sem elkezdni, sem befejezni nem lehet azt. Jellemzően ezek önmagukban nem bírnak nagy jelentőséggel. A kirakó játék legfontosabb belső darabjai, amelyekhez gyakran a legnehezebben és a legdrágábban lehet hozzájutni, a hagyományos felderítő eljárásokból szerezhetők meg. A nyílt források figyelése a szükséges alapja a felderítő eljárásnak, de nem pótolhatja az egész, minden oldalú felderítő erőki-fejtést.”¹³

Véleménye szerint a nyílt információforrások a 70-90%-át elégíthetik ki a teljes felderítő igényeknek. Olyan összefüggéseket tárhatnak fel, amelyek rámutathatnak egy adott téma lényegére, illetve segíthetik azt, hogy a minősített anyagokban milyen irányban kell összpontosítani a keresést.

1.3.3. Változások az információszerzés terén

A hidegháború megszűntével a világpolitikai helyzet a várakozások ellenére nem egyszerűsödött, hanem a várakozások ellenére, inkább bonyolultabb lett. Most olyan új, merőben eltérő *kihívások*nak kell megfelelni, mint a nemzetközi terrorizmus, a tiltott fegyver- és drogkereskedelem, az etnikai feszültségek és a migráció.

Fontos kérdés, hogy hogyan lehet a felderítő tevékenység annyira rugalmas, hogy mindig a kellő időben megválaszolja az adott helyzetben felmerülő legfontosabb politikai és katonai kérdéseket? Ebben a nyílt forrású anyagok szerepe jelentős mértékben megnövekedett, mivel azok bármikor, egyszerűen összegyűjthetők, reális időben tájékoztatnak a világ tényeiről, történéseiről, folyamatosan és azonnal rendelkezésre állnak.

Az információszerzés eljárásainak és képességeinek megváltozása mellett új jelenség, hogy *az információszerzés privatizálódik*. Ma már nem csak állami szervezetek végeznek felderítő munkát, hanem civilek is folytatnak adat- és információgyűjtést, elemzést, valamint értékelést. Ez sok esetben meglepően termelékenyebb és költséghatékonyabb, sőt sokkal diszkrétebb is lehet.

A nyílt forrásból származó anyagokat *meg lehet osztani* közvetlenül az iparral, a sajtóval, a törvényhozással és a szövetségesekkel mindenféle politikai kockázat nélkül. Az egyes szervezetekkel összeköttetést lehet teremteni számítógépes hálózaton keresztül, ami lehetővé teszi a folytonos kapcsolatot az egyes munkatársak, a felső vezetés,

¹³ Dr. Joseph Nye, the Chairman of the National Intelligence Council, speaking to members of the Security Affairs Support Association at Fort Meade, Maryland, on 24 April 1993.

valamint az elemzők között. *A megszerzett anyagokat online módon is elérhetővé lehet tenni.*

A nyílt információszerzésben szakítani kell azzal a korábbi filozófiával, miszerint az egyes országokat valamilyen szisztéma szerint besoroltuk első-, másod vagy sokadrendűnek, és az elsőrendűről minden elérhető, fontos vagy kevésbé fontos információt összegyűjtöttünk, ha volt rá szükség, ha nem. Ezzel a módszerrel óriási méretű adathalmaz gyűlt össze az elsődleges prioritású országokról, de szinte semmi az alacsonyabb rendűnek besoroltakról. Így abban az esetben, amikor ott valamilyen konfliktus tör ki nem áll rendelkezésre semmilyen információ. A mai megváltozott világban azt az információt kell elsődlegesen megszerezni, amit a felhasználók kérnek. Ezzel elkerülhető, hogy főlegesen gyűjtsünk össze olyan információkat, amelyekre senkinek sincs szüksége, de azok, amelyek igazán fontosak lennének nem állnak rendelkezésre.

Az információs korszak technológiája kitágítja a lehetőségeket arra, hogy miért és mikor szerezzünk meg egy információt. Átállhatunk az eddigi „*minden esetre*” („just-in-case”) való *információszerzésről*, vagyis arról, hogy előre összegyűjtsünk minden adatot, az „*éppen időben*” („just-in-time”) elvű *megközelítésre*, ami azt jelenti, hogy nem az információkat gyűjtjük össze, hanem azokat a forrásokat, amelyekből a megfelelő információ adott esetben kinyerhető. A nyomtatott anyagok és a mikrofilmek voltak valamikor a domináns tároló médiumok a régi „minden esetre” paradigma alatt. A globális, közvetlen elérésű, digitális elektronikus információ lesz domináns az „éppen időben” új paradigma alatt. *Ehhez a legjobb információforrás az internet.*

1.3.4. A nyílt forrású információszerzés előnyei

A nyílt források nem képesek önmagukban teljes mértékben kielégíteni a felhasználók igényeit, de számos előnyük van a többi információszerzési eljárással szemben.

1. Csökkenti az információ megszerzéséhez szükséges költségeket. A nyílt források felderítésének hosszabb idejű tapasztalata alapján, *Robert D. Steele*¹⁴ szerint: „A titkos adatforrásokból (HUMINT, SIGINT, IMINT, MASINT) a felhasználható információ 20%-a szerezhető meg, amire a költségek 95%-át kell rááldozni, a nyílt forrásokból (NFF) a felhasználható információ 80%-a származik, amihez a költségek 5%-át kell felhasználni”.¹⁵ Ezek szerint a minősített információk megszerzése igényli a költségek

¹⁴ 1992-ben megalapította az OSS Inc.-t (Open Sources Solution – Nyílt Forrású Megoldások Egyesülete).

¹⁵ Robert David STEELE: *Multinational Intelligence Can CENTCOM Lead the Way? Reflections on OSINT & the Coalition*, Vivas Presentation to the Coalition Coordination Center 27 January 2006, 5. old.

nagyobb részét, de ehhez képest csak kevés információt lehet ezekből a forrásokból szerezni, igaz hogy sokszor ezek a leglényegesebbek és a legfontosabbak.

2. Időszerű. A nyílt információszolgáltatóknak – különösen a médiának – érdeke, hogy az általa nyújtott információk minél frissebbek legyenek, mert ezzel növelhető az eladott példányszám, illetve a nézettség. Emiatt a nyílt forrásokból lehetséges olyan *aktuális információkat (current informations)* szerezni, amelyek az adott időszakban zajló napi eseményekről tájékoztatnak a történésekkel szinte egy időben. Ez lehetővé teszi a válságok kialakulásának előrejelzését, lefolyásának figyelemmel kísérését, valamint a stabilitást veszélyeztető események azonnali felismerését, vagyis a riasztást.

3. Kiterjed minden területre. Számos esetben fordul elő az, hogy bizonyos események hatására, olyan ország vagy térség kerül a figyelem középpontjába, amely addig nem volt a felderítés számára fontos. Ilyenkor nagy segítséget tudnak nyújtani a nyílt információk a tájékozódásban, az ismeretek megszerzésében.

Ezt jól illusztrálja egy példa az első Öböl-háborúból¹⁶: „A Sivatagi Vihar¹⁷ során, egy manőver előkészítéséhez, szükséges volt információt szerezni arról, hogy hogyan lehet harcjárművekkel átkelni Dél-Irak sivatagos területein. Meg kellett tudni, hogy a homok hol túl laza ahhoz, hogy a harckocsikat ne tartsa meg, és hol vannak olyan szakadékok, amelyeknél meg kell állni a járműoszlopnak és hidászati eszközöket kell alkalmazni. Nagy adatvadászat indult el. Az *Egyesült Államok Kongresszusi Könyvtáráról*¹⁸ kiderült, hogy nagyon hasznos információforrás. Felderítő tisztek egy csoportja három napot töltött azzal, hogy az egykori régészeti kéziratokat áttanulmányozza és megtalálja a közlekedési lehetőségekről a feljegyzéseket. Régészek korábban, a XX. század elején, aprólékosan feljegyeztek mindent a naplójukban erről a vidékről, amint teveháton lassan keresztül haladtak a sivatagi homokon.”¹⁹

A felderítő szervezetek nem tudhatják előre összegyűjteni az összes olyan információt, amire a jövőben szükség lehet, de az nagyon is lehetséges, hogy olyan értékes információkat szerezzenek, amiket valaki más gyűjtött és állított össze. A nyílt információs rendszer teszi lehetővé ezeknek az információknak a megtalálását.

¹⁶ 1990. 08. 02-én kezdődött Kuvait Irak általi lerohanásával. A konkrét katonai művelet 1991. január 17-én kezdődtek és február 28-án fejeződtek be.

¹⁷ angolul: Desert Storm.

¹⁸ Library of Congress, honlapja: <http://www.loc.gov/>

¹⁹ Open Source Intelligence Resources for the Intelligence Professional, The 434th Military Intelligence Detachment, New Haven, 1994. 7. old.

4.) Kockázatmentes. Az ipari kémkedés, de még a rejtett- és technikai információszerezés is komoly politikai kockázatokkal jár, hiszen mindig kínosak az esetleges lebukások, és bizalmi válsághoz vezethetnek.

A nyílt információkhoz a hozzáférést – annak lényegéből adódóan – természetesen nem korlátozzák, sőt a terjesztőinek ahhoz fűződik érdeke, hogy minél szélesebb körhöz jusson el. A nyílt információk gyűjtése ezért sehol sem ütközik törvénybe, így nem jár semmilyen politikai vagy egyéb kockázattal.

A titkos anyagok megszerzésének egyik speciális és kockázatmentes módja lehet, ha azokat még *előpublikációként*²⁰, vagyis még minősítés előtti állapotában szerezzük meg. A legtöbb újonnan fejlesztett technológia esetén a kutatás időszakában a részeredményeket nyilvános publikációkban ismertetik. *Az adott terület szakértői képesek lehetnek ezekből a részletekből összerakni az egész projektet.* Ezen kívül minden kettős hasznosítású technológia – ami katonai és polgári felhasználásra is kerülhet, ezért az exportját ellenőrzik – egy bizonyos ideig *béta változatban*²¹ hozzáférhető. Azok a hivatalok, amelyek a kettős használatú technológia titkosítását irányítják, sokszor akár évekkel is le vannak maradva az aktuális fejlesztések mögött, ez különösen igaz a szoftver iparra. Így a legjobb idő megszerezni a „titkos” vagy zárolt technológiát azalatt az idő alatt, amíg béta változatban van, és még nem lett minősített.

5. Megosztható külső szakértőkkel. A nyílt forrásokból származó anyagok esetén lehetőség van arra, hogy számos szakértő és kommentátor elemzésének és megjegyzésének vessük alá azokat. Vannak az életnek, a tudománynak vagy szakmának olyan speciális területei, amelyekhez csak az abban jártas szakemberek értenek. Vannak olyan témák, amiknek a megértéséhez speciális felkészültségre van szükség. A nyílt forrású anyagok esetében nyugodtan *bevonhatók az elemző munkába külső szakemberek is.*

6. Együttműködési lehetőség a polgári információszerző szervezetekkel. Nyílt forrású információszerzést nem csak a katonai felderítés folytat, hanem magán szervezetek, hírügynökségek és akadémiai intézmények. Számos országban több polgári információszerző szervezet létezik, mint katonai vagy nemzetbiztonsági. A felderítő szervezetek és a magán társaságok tevékenységében sok a közös vonás. Mindketten, lehetőség szerint minél olcsóbban, megbízható információhoz akarnak jutni. Vannak olyan esetek, amikor a katonai felderítésnek körülményes információhoz jutni, de a magán

²⁰ Open Source Intelligence Resources for the Intelligence Professional, The 434th Military Intelligence Detachment, New Haven, 1994. 18. old.

²¹ Olyan teszt változat, amikor a szokásos belsők mellett már külső személyeket is bevonnak a programok tesztelésbe.

szervezetnek erre van lehetősége. Jó példa erre az Irak ellen folytatott első háború, ahol a CNN²² stábjának lehetősége volt forgatni az országban, és ennek segítségével meg lehetett győződni a bombázások eredményességéről minden katonai erőfeszítés nélkül. A felderítésnek meg kell találnia az együttműködés lehetőségeit a magán hírszerző szervezetekkel.

1.3.5. A nyílt források korlátai

A nyílt forrásokból származó anyagoknak a fent említett előnyei mellett számos olyan tulajdonságai vannak, amelyek megnehezítik a hatékony információszerzést.

1. Mennyiség. A napjainkban végbement információrobbanás miatt már nem az a probléma, hogy nincs elég információ, hanem éppen fordítva, inkább túl sok van belőle. Minden nap több ezer napilap, hírlap, kutatási jelentés, könyv, magazin, rádióadás, TV műsor, dokumentum és egyéb termék keletkezik. Ezek nyilvánvalóan értékes információt tartalmaznak, de *feldolgozásuk, tárolásuk, indexelésük időigényes* és csak valamilyen nagyteljesítményű számítástechnikai eszközzel lehetséges. Óriási erőfeszítést, szakértelmet és jártasságot igényel az adatfeldolgozóktól, hogy használható információt lehessen kinyerni ebből a hatalmas mennyiségű tényből és véleményből.

2. Sokféleség. A naponta keletkező információk *típusa, alakja, megjelenési formája* széles skálán belül változik, az egyszerű kutatási ötletet tartalmazó írógéppel írt papírlapoktól a több száz magnó-, vagy videó kazettányi idegen nyelvű konferencia anyagokig.

A sokféleség nem csak a megjelenési formákban jelentkezik, hanem az *anyagok nyelvében* is. Az idegen nyelvű anyagokat le kell fordítani, s ehhez az adott nyelvet nagyon jól ismerő *fordítókra van szükség*. A nyílt információk többsége csak eredeti nyelven található meg. Például a hírügynökségek nagy része is csak a saját nyelvén adja közre a saját híreit. Abban az esetben, ha más nyelven (pl. angolul) is kiadják azokat, akkor nem biztos, hogy tartalmaznak minden információt, mert esetleg valamilyen szempont szerint szelektálják a híreket.

3. Különböző nézőpontok. Ha valahol kialakul egy konfliktus két fél között, akkor elolvashatjuk a két szembenálló oldal szereplőinek tollából származó cikkeket, jelentéseket, kommentárokat. Az *ellentétes érdekű csoportok eltérő módon értelmezik ugyanazt az eseményt*, történést, vagy helyzetet, hogy ezzel is támogatást nyerjenek a saját

²² CNN: Cabel News Network – Kábel hírhálózat, 24 órás híreket sugárzó TV csatorna, amit 1980-ban alapított Ted Turner.

politikai vagy társadalmi céljaikhoz. A felderítő elemzők csak komoly gyakorlat birtokában képesek kiszűrni a különböző forrásból származó, azonos témáról vagy esetről szóló jelentésekből a valóságos értéket. Az ugyanazon események értékelése közötti különbségek sokszor olyan értékes információt jelentenek, mint maguk az eredetiek.

4. Pontatlanságokat és szándékos ferdítéseket tartalmazhat. Mivel nagy mennyiségű és változatos tartalmú információ áll rendelkezésre a nyílt forrásokból, ezért komoly *jártasságot igényel az elemzőtől, hogy megbecsülje egy adott információ minőségét és pontosságát.* A kormányok általában felnagyítják a számukra hízelgő számadatokat, annak érdekében, hogy a politikai programjuk sikerét bizonyítsák. Előfordul, hogy állambiztonsági, politikai, üzleti vagy más érdekből a valóságnak nem teljesen megfelelő információkat hoznak nyilvánosságra.

Az információk pontatlanságát nem csak szándékos ferdítés okozhatja, hanem a tudósítói szakma gyorsabbá és elnagyoltabbá válása. Ez főként az internetes újságírással igaz. A spekulációk, félkész információk is könnyen nyilvánosságra kerülnek, hiszen mindenki tisztában van azzal, hogy utólag minden korrigálható. A különböző történetek valós időben, az olvasók szeme előtt épülnek fel, mivel a szerzők közvetlenül az internetre publikálnak, gyakran szerkesztői ellenőrzés nélkül. A hagyományos újságírással szemben az egyik legfőbb értéke az információk többszörös ellenőrzése, az internetes újságírásban értelmét veszíti.

Végül következtetésként elfogadhatjuk Robert Steele véleményét, mely szerint: „Ha az információ csak 85%-os pontosságú, de a megfelelő időben rendelkezésre áll és azt meg tudom osztani másokkal is, akkor az sokkal használhatóbb, mint az, ami minősített forrásból származik, de a túl hosszadalmas és különleges biztonsági kezelési igény miatt csak késve hozzáférhető”.²³ Az időben érkező információ a jelenre, illetve a jövőre, míg az elkésett elsősorban a múltat érinti.

1.4. A NYÍLT INFORMÁCIÓK FELHASZNÁLÁSÁNAK SZINTJEI

A nyílt információkkal kapcsolatban ma már nem kell bizonyítani, hogy azok fontosak lehetnek a politika, a nemzetbiztonság és a felderítő elemzés minden szintjén.

1.) Politikai stratégiai, nemzetbiztonsági szinten. A nyílt források körültekintő átvizsgálásával – még akkor is, ha azok cenzúrázottak, vagyis tartalmilag hiányosak –

²³ ON INTELLIGENCE: Spies and Secrecy in an Open World, OSS International Press, Fairfax, 2001. 112. old.

olyan ismereteket lehet szerezni, amelyek alapján lehetségessé válnak az élet különböző területet átfogó, elemző, értékelő, összefoglaló, áttekintő jelentések összeállítása. Az ilyen jelentések alapján a kormányzati vezetők, döntéshozók megérthetik egy adott ország politikáját mozgó érdekeket és szándékokat, valamint segíthetik a társadalmi elégedetlenségek, forradalmak, felkelések és zavargások előrejelzését. Hadászati szinten a nyílt források nagyon jól alkalmazhatók az olyan lexikális ismeretek és anyagok felhalmozásához, amelyek ahhoz kellenek, hogy fel lehessen ismerni a veszélyek kialakulásának lehetőségét egy adott területen, valamint kiválóak a csapatok és a közvélemény tájékoztatására.

2.) Védelmi stratégiai szinten. A nyílt források lehetővé teszik az erő kifejtés megtervezését, valamint az egyesített és koalíciós műveletek koordinálását. Például abban az esetben is nagyon jól használhatóak, amikor olyan területen kell műveleteket végrehajtani, ahol az időjárási körülmények nagyban eltérnek a hazaitól. Ebben az esetben a megszerzett ismeretek alapján, fel lehet készülni arra, hogy a műveletek során alkalmazni kívánt eszközök hogyan fognak viselkedni ezek között a szokatlan körülmények között, illetve megbecsülhetők azok élettartamának, határfokának esetleges csökkenése, vagy növekedése. A nyílt forrásokból a tanácsadók, döntés-előkészítők, szakértő elemzők részére olyan háttérismeretek gyűjthetők össze, amelyek apró részletekbe menően mutatnak be egy adott helyzetet, eseménysort, ezzel segítve azok látókörének, ismereteinek kibővítését, naprakészen tartását, valamint az analógiák, az összefüggések és a kapcsolatok felismerését.

3.) Védelmi műveleti és technikai szinten²⁴. A nyílt források közül nagyon értékesek lehetnek az olyan kereskedelmi forgalomban kapható térképanyagok, mint a 1:50.000-hez léptékű, szintvonalakkal ellátott térképek olyan területekről, amelyekről nem áll még rendelkezésre digitális térkép. A harmadik világban végrehajtott missziók tapasztalati alapján, amerikai szakértők arra a meggyőződésre jutottak, hogy egy parancsnoknak nem szabad félnie attól, hogy kereskedelmi képanyagokat és térképeket használjon fel a parancsnoki munkája támogatásához. A kereskedelmi képi anyagok között olyan multispektrális érzékelőkkel készült képek is vannak, amelyek felhasználásával el lehet készíteni egy terület háromdimenziós térképét, például megtervezni egy repülőtér megközelítését és az arra való leszállást.

²⁴ Megfelel a Magyar Köztársaság összhaderőnemi műveleti tevékenységének.

A műveleti területen lévő polgári kommunikációs és számítástechnikai képességekről megszerzett ismeretek nagyon fontosak lehetnek a parancsnoknak. Ez alapján meg tudja becsülni a szemben álló fél elektronikai hadviselés folytatásához rendelkezésre álló képességeit, továbbá megértheti, hogy hol és hogyan kell gyengíteni az ellenfél ilyen irányú képességeit, illetve hogyan használja ki azokat a saját kommunikációs igényeinek kielégítésére.

Segítségükkel információhoz juthatunk az adott terület olyan ismérveiről, amelyek alapján megtervezhetők a végrehajtandó műveletek, mivel képet lehet alkotni a mozgási lehetőségekről, a fegyverek alkalmazhatóságáról, a terep átláthatóságáról, a logisztikai képességekről és egyéb feltételekről.

1.5. A NYÍLT FORRÁSÚ INFORMÁCIÓSZERZÉS CSATORNÁI

A kapcsolódó szakirodalom alapján²⁵ az alábbi csatornák merülnek fel, mint a nyílt forrású információszerzés adatforrásai:

- ☞ hagyományos (nyomtatott és elektronikus) média;
- ☞ szürke irodalom;
- ☞ szakértők és megfigyelők személyes tapasztalatai;
- ☞ kereskedelmi műholdas felvételek;
- ☞ kereskedelmi online információszolgáltatók;
- ☞ **INTERNET;**
- ☞ egyéb (nyelvviskolák, egyetemek, üzleti élet, információbrókerek, újságírók, nemzetközi és nem kormányzati szervezetek).

1.5.1. Hagyományos média

A hagyományos média (újságok, rádió- és televízió műsorok) segítségével figyelemmel lehet kísérni a napi eseményeket, vagy a válságokat, azok kirobbanásától a rendeződéséig. Ezt az teszi lehetővé, hogy a világon szinte mindenütt vannak tudósítók, vagy ha kell, ki tudnak küldeni az események helyszínére, vagy ha az nem lehetséges, akkor annak közelébe. Emellett az adott ország saját tömegtájékoztatása is igénybe vehető a tájékozódáshoz. Minden újság működtet valamilyen adatbázist, amiben a megjelent számokat és egyéb hasznos információt tárolnak. Ezeknek a segítségével elemzéseket

²⁵ NATO Open Source Intelligence Handbook, Saclant, Norfolk, 2001. 1. fejezet 1. old.

lehet készíteni egy kialakult válság okairól, történelmi, gazdasági, vallási vagy egyéb háttéréről, a szemben álló felekről, stb.

Az elmúlt tizenöt évben – az országok nagy többségében – óriási számban indultak új *televízióadók*, amelyek egymással versengenek azért, hogy mindenki előtt megszerezzenek egy-egy fontos információt. Nagyszámban vannak olyan *műsorok*, amelyekből hasznos információhoz lehet jutni. A dokumentum-, katonai-, katonapolitikai-műsorok például olyan sokrétű háttér-információt adhatnak egy felkészült elemző számára egy országról vagy egy adott eseményről, amelyet más médiából nem lehet megszerezni. Sok elemző véleménye szerint egy ország megismeréséhez sokkal többet ér, ha az aktuális eseményeket *vizuálisan* tekintjük meg, mint megbízni a hírlapokban leírtakban. Itt érvényesül az „egy kép többet mond ezer szónál”²⁶ effektus.

A televíziós hírműsorok jelentősen közelebb hoznak minket a világban történő eseményekhez és válsághelyzetekhez, valamint lehetővé teszik az olyan zárt társadalmakról, mint például a katonai vezetés alatt álló országokról, az ismereteink kiszélesítését. A külföldi televízióadások elemzése után a felderítő szervezeték képesek a hírműsorból olyan *videó anyagot összeállítani*, amely segítheti a politikai vezetőket az egyes események jobb megértésében.

1.5.2. Szürke irodalom

A *szürke irodalomba* (gray literature) tartozó anyagok közös tulajdonsága, hogy *nehezen hozzáférhetőek, kívül maradnak a kereskedelmi-kiadási terjesztési hálózaton* és sokszor nem is tudunk a létezéséről. A szürke irodalom, a médiára való tekintet nélkül, magába foglalja: a munka- és vitaokmányokat; kefélynyomatokat; a kutatási, piackutatási, gazdasági és egyéb jelentéseket; a szabványokat; az úti beszámolókat; a nem hivatalos kormányzati dokumentumokat; a jegyzőkönyveket; a tanulmányokat, a disszertációkat és téziseket; a kereskedelmi irodalmat; a szabadalmi bejegyzéseket; a konferenciák előadásait és azok képi-, video- és írásos anyagait; az egyetemi-, főiskolai jegyzeteket és szakdolgozatokat; a nemzetközi szerződéseket.

Ezek az anyagok átfogják a tudományt, a társadalmi-gazdasági, a politikai és a katonai ismereteket. Szürke irodalmat leginkább nonprofit szervezetek, kutató szervezetek (laborok és intézetek), nemzeti kormányhivatalok, magán kiadók (sajtó csoportok/politikai pártok), részvénytársaságok, kereskedelmi társaságok/szövetségek, kutatóhelyek, egyetemek és főiskolák produkálnak.

²⁶ Kínai közmondás.

A nemzeti és nemzetközi szintű szabályozásoknak köszönhetően napjainkban már nem reménytelen e dokumentumok számbavétele. Jól szervezett dokumentumszolgáltató rendszerek és dokumentumszolgáltató központok jöttek létre (pl. *GLWG*²⁷, *BLDSC*²⁸, *DSC*²⁹, *INIST*³⁰, *NTIS*³¹ stb.). A probléma nem az, hogyan érhető el, hanem az, honnan szerezhetünk tudomást a létezésükről. A szürke irodalom adatbázisai egy-egy tudományterületen vagy nagyobb témakörökön belül szerveződnek. Egyes becslések szerint a szürke irodalom mennyisége 3-4-szerese a hagyományosénak.

1.5.3. Szakértők és megfigyelők személyes tapasztalatai

A legkülönlegesebb nyílt forrás maga a szakértő, megfigyeléseket végző nagy tapasztalatokkal rendelkező ember. Nincs olyan hírszerző szolgálat, amely valaha is képes lesz mindig minden szükséges adatot időben összegyűjteni és feldolgozni. Akkor lenne tökéletes a hírszerző szolgálat, ha képes lenne arra, hogy az információ felhasználóját közvetlenül összekapcsolja a számára fontos információforrással. A tapasztalt ember segítségével valós idejű hozzáférést lehet biztosítani más forrásokhoz, az emberi elme feladatérzékeny, képes különböző forrásokból gyorsan összegyűjteni az ismereteket, azokat feldolgozni majd a megértett, kiértékelt ismereteket másoknak igény szerint átadni.

Vannak a világnak olyan területei (pl. Afrika), amely adottságairól és kondícióiról csak nagyon kevés anyagot publikálnak. Néhány témában még akkor sem lehet információt találni, ha az anyagok átnézésére és megszerzésére kellő idő és eszköz áll rendelkezésre. Ilyenkor lehetnek hasznosak az adott területet jól ismerő, a helyet megjárt olyan emberek, mint a nemzetközi méretű nem kormányzati szervezetek (Pl. Vörös Kereszt, Orvosok Határ Nélkül) tagjai, szakemberei. Ezek a személyek olyan információkkal is szolgálhatnak, amelyek a helyszínről származnak, mert vannak a helyi lakosság köréből kikerülő munkatársaik is, akik a helyszínen vagy annak közelében tartózkodnak és természetesen beszélnek a helyi nyelvet.

²⁷ GLWG: Gray Literature Work Group: az Egyesült Államokban a szürke irodalom elérésének biztosítására alakított munkacsoport.

²⁸ BLDSC: British Library Document Supply Centre – Brit Nemzeti Könyvtár Dokumentum Ellátó Központ (<http://www.bl.uk/>). Az Egyesült Királyság fő szürke irodalom dokumentumszolgáltatója.

²⁹ DSC: The Directory of Social Change – Társadalmi Változások Könyvtára (<http://www.dsc.org.uk/>)

³⁰ INIST: L'Institut de l'Information Scientifique et Technique – Tudományos és Technikai Információk Intézete (<http://www.inist.fr/>). Franciaország fő szürke irodalom dokumentumszolgáltatója.

³¹ NTIS: National Technical Information Service – Nemzeti Technikai Információ Szolgálat (<http://www.ntis.gov/>). Az Egyesült Államok fő szürke irodalom dokumentumszolgáltatója.

1.5.4. Kereskedelmi műholdas felvételek

Sokáig csak tudni lehetett arról, hogy a műholdak által szolgáltatott képi anyagok nagmértékben segítik a felderítő munkát, de ilyen képekhez nem juthatott hozzá hétköznapi ember. A műholdtechnika széleskörűvé válásával már elérhetőek lettek a nagy felbontású (pl. *Landsat*³², *SPOT*³³, *Google Earth*³⁴) műholdas képek, amelyek akár katonai célra is felhasználhatók. A méteres felbontású electro-optikai képek ma már teljesen természetesen a magánszektorban is.

A műholdról készült képek mellett a távérzékelés felhasználása is egyre szélesebb körben terjed el. A *meteorológiai előrejelzésre* specializálódtak mellett, *kutató-* és a *földfelszint letapogató*, a civil szférában alkalmazható, kereskedelmi célokat szolgáló műholdak alkalmazása is kifejlődött. Ezek eszközei lehetnek aktív vagy passzív működésűek, optikai, infravörös, ibolyántúli vagy mikrohullámú frekvenciatartományt alkalmazók, és kifelbontású (néhány 100 m/pixel) vagy nagyfelbontású (néhány m/pixel) felvételeket készítenek. Segítségükkel akár megfelelő méretarányú térkép (harcászati szinten 1:50.000-es) is készíthető egy adott területről. A kereskedelmi műholdfelvételeket számtalan, műholddal nem rendelkező ország, közöttük NATO-tagországok is használják.

1.5.5. Kereskedelmi online információszolgáltatók

Létezik sok olyan kereskedelmi online szolgáltató, amely *előfizetési díj, vagy használati díj ellenében biztosít hozzáférést az általa összeállított információkhoz*. Ezek az adatbázisokba gyűjtött információk az élet egy-egy területét (gazdasági, piaci, politikai, katonai) fogják át. Az adatbázisok szakemberek által feldolgozott adatokat tartalmaznak, egyszerű keresési módszerük viszonylag gyorsan elsajátítható. Tudni kell, hogy az adott cél érdekében melyik adatbázisban található meg nagy valószínűséggel a keresett válasz. Ez nem is mindig olyan könnyű, mert például a *Dialog Corporation*³⁵ több, mint 900 különféle témájú adatbázist kínál az ügyfeleinek, ezek mellett 7000 szakfolyóirat elektronikus kiadását is elérhetővé teszi, döntően teljes szöveges formátumban.

Az online elérés lehetővé teszi, hogy az *adatbázisok készítésük ütemében* – a hír-adatbázisok akár a nyomtatott hírek megjelenése előtt – *válnak elérhetővé*. Az adatbázis-gyűjtemények szolgáltatói azt is kínálják, hogy egyszerre több adatbázisban is lehes-

³² Honlapja: <http://www.landsat.org/>.

³³ Honlapja: <http://www.spot.com/>.

³⁴ Honlapja: <http://earth.google.com/>.

³⁵ Honlapja: <http://www.dialog.com/>

sen futtatni ugyanazt a keresőprofilot. Így mérsékelhető a keresési idő, és az információ szóródásából eredő keresési veszteség.

Az adatbázisok tartalomra és az adatok részletezettsége szerint is tagozódnak. A vállalatokról szóló információt több szintre lehet bontani: alap, jogi, minősített, marketing és egyéb. Ezek többé-kevésbé árkatagóriákat is jelentenek: míg az alapadatok ingyenesek, a cégstruktúrát, hitelminősítést, esetleg a vezetők hobbiját tartalmazók jóval drágábbak. Emellett mód van a *témafigyelésre* is: az előre megadott profilban a szolgáltató akár az információ *megjelenésének pillanatában küldi elektronikusan a megfelelő tételeket* a megrendelő postafiókjába. Az online keresés esetén nagyobb lehetőség nyílik a keresési stratégia rugalmas változtatására is. A vaskos (és méregdrága) piaci jelentések egészének megrendelése helyett mód van arra, hogy csak a szükséges egy-két táblázatot vagy összefoglalót töltsük le, fizessük ki, akár a ma közreadott felmérésből.

Nagy *vállalatok, társaságok és kormányhivatalok* is működtetnek online adatbázisokat. Például 1971-ben az IBM³⁶ elkezdte működtetni az ITIRC³⁷-nek hívott rendszerét, amelyet bármely IBM telephelyről el lehet érni. Ezen megtalálható minden IBM készülék kezelési leírása, az összes számítógéppel kapcsolatos publikáció, belső dokumentum, jelentés, újság- és szakfolyóirati cikk. Hasonló rendszereket használnak a kormányzati hivatalokban és az Egyesült Államok felderítő közösségében is jó néhány éve.

1.5.6. Internet

Az internet nem más, mint a hálózatok hálózata, vagyis az információ továbbítására egymással valamilyen módon (rézérpár, koaxiális kábel, sodrott érpár³⁸, optikai kábel, rádió-, infravörös-, lézerhullám) összekötött, a TCP/IP³⁹ protokollkészletet alkalmazó, hálózatba (LAN⁴⁰, WAN⁴¹, WLAN⁴²) kapcsolt számítógépek együttese. Létrejöttét az az igény indokolta, hogy az Egyesült Államok vezetési-, irányítási rendszerét decentralizálják, és ezzel növeljék túlélőképességét egy atomtámadás esetére. A kutatómunka 1969-ben kezdődött az ARPANET⁴³ keretein belül. Ekkor még csak négy csomópontból állt a hálózat. A fejlődése kezdetben lassú volt, mert elsősorban csak katonai kutatási

³⁶ IBM: International Business Machine – Nemzetközi Üzleti Gépek.

³⁷ ITIRC: IBM Technical Information Retrieval Center – IBM Technikai Információ-visszakereső Központ.

³⁸ Két fajtája van: árnyékolatlan sodrott érpár - Unshielded Twisted Pair – UTP; árnyékolt sodrott érpár Shielded Twisted Pair – STP.

³⁹ Transmission Control Protocol/Internet Protocol – átvitel vezérlési protokoll/Internet protokoll

⁴⁰ LAN: Local Area Network – helyi hálózat.

⁴¹ WAN: Wide Area Network – nagytávolságú hálózat.

⁴² WLAN: Wireless Local Area Network – Vezeték nélküli helyi hálózat.

⁴³ Advanced Research Project Agency Network – Korszerű Kutatásokat Tervező Intézet Hálózata

projekt volt, így 1979-ben a kapcsolódó hosztok száma 188, tíz évvel később, 1989-ben is csak kb. 80 ezer volt (lásd 6. számú melléklet).

A fejlődés felgyorsulását két dolog tette lehetővé: a *WWW*⁴⁴ bevezetése 1991-ben, ami megkönnyítette az egyszerű felhasználónak is az internet használatát, és az *üzleti élet bekapcsolódása* a fejlesztésbe, ami új pénzforrásokat jelentett. Ezután a fejlődés meredeken ívelt felfelé, s mára a teljes világot átfogó, milliárdnál több felhasználót kiszolgáló rendszerré változott.

Az interneten megtalálható anyagok tartalma széles skálán mozog, a napi friss hírektől az archivált újságcikkekig, a tudományos írásoktól a zenerajongók üzenetváltásáig. Sok egyén, társaság, kormányhivatal, intézmény és egyéb szervezet közread különböző információkat az interneten. A hírcsoportok, blogok vagy fórumok ezrei témakörökbe szervezett üzeneteket és vitafórumokat tartalmaznak az emberi élet minden területéről. *Az internetről szinte minden olyan adat beszerezhető, ami a hadműveleti tervezés megkezdéséhez és a csapatok felkészítéséhez szükséges.*

1.5.7. Egyéb források

Az élet számos területén keletkeznek olyan nyílt adatok, információk, amelyek értékek lehetnek, illetve sok olyan szervezet, intézmény és egyén van, ami/aki szakértelmével, ismereteivel, vagy különleges tudásával segítheti az információszerző munkát.

Nyelviskolák: a magyarul tudó, de idegen anyanyelvű diákok jelenthetnek komoly segítséget egy adott ország anyagainak lefordításában. A nyelviskolák gyakran alkalmazzák is őket rész munkaidőben fordítási feladatok végrehajtására.

Egyetemek: mivel az oktatás mellett ma már minden egyetem kutató munkát is végez, ezért az általuk összeállított adatbázisok szakmailag megbízható forráslehetőséget jelentenek

Könyvtárak: az információs társadalomban is az információgyűjtés és -feltárás alapintézményei maradnak. Nagymennyiségű újságot, különböző témájú könyvet gyűjtenek össze, amelyekből hasznos információ szerezhető országokról, gazdasági, politikai és egyéb kérdésekről.

Üzleti élet: általános vagy specifikus információt nem csak a katonai felderítés és a nemzetbiztonsági szolgálatok gyűjtenek, hanem számos polgári szervezet is. Az üzleti élet résztvevői sok esetben a nyílt forrású anyagok minden fajtáját összegyűjtik, és utána feldolgozzák, katalogizálják. Tevékenységük során folyamatosan figyelemmel kísé-

⁴⁴ WWW: World Wide Web – világháló.

rik a bírósági döntéseket, piackutatási eredményeket, napilapok és folyóiratok cikkeit, az egyes cégekről és azok tevékenységéről rendelkezésre álló anyagokat, valamint minden anyagi szolgáltatás ellenében hozzáférhető dokumentumot. Összegyűjtik egy ország esetében a politikai korrupcióra, a klimatikus viszonyokra és egyéb specifikus adatokra vonatkozó ismereteket. Egy adott országban beruházó szakkégek információkat adhatnak akár a kommunikációs és informatikai infrastruktúra elemeiről, valamint azok telepítési helyeiről, amennyiben részt vettek a kiépítésében. Építőipari vállalatok ismerhetik katonai vagy stratégiai jelentőségű épületek, katonai célra is használható repülőterek, kikötők, stb. tervrajzait is.

Újságírók: általában csak egy szakterületre specializálódnak (pl. technika, gazdaság, sport, politika, katonai kérdések). Különleges felkészültségüknel fogva alkalmasak szakmai riportok készítésére, háttér információk vagy háttéranyagok összeállítására. Segítségükkel sokkal egyszerűbb a különböző országok vagy térségek szakújságíróval felvenni a kapcsolatot.

Információbrókerek: az információkeresésre szakosodott szakemberek, akik megrendelőik problémájának megoldása érdekében – térítésért – testreszabott információs csomagot állítanak össze. Kézzelfogható a hasonlóság az újságíróval, mivel mindkettőn adott témában gyűjtenek össze adatokat, információkat, hogy azokat, megfelelő formába rendezve, ügyfeleik vagy olvasóik saját tudásuk építésére használhassák. Különbség kettőjük között, hogy az információbróker célcsoportja egy vagy néhány ember, esetleg munkacsoport, az újságíróé a széles nyilvánosság.

A civil társadalom szervezetei: ezek a nem kormányzati szervezetek (NGO⁴⁵), a szakmai, a karitatív és a bázisszervezetek, az állampolgárokat a helyi életbe bekapcsoló szervezetek, illetve, specifikus hozzájárulással, az egyházak és a vallási közösségek. Ezek a szervezetek világra kiterjedő információs hálózattal rendelkeznek, sokszor még olyan térségekben is, amelyek kívül esnek a NATO érdekövezetén. Mivel ezek a szervezetek a helyi lakosságból is választanak munkatársakat, ezért különleges nyelvek ismeretével is rendelkeznek.

⁴⁵ A nem-kormányzati szervezet angol elnevezése: non- governmental organisation, rövidítése: NGO.

1.6. A NYÍLT ELEKTRONIKUS INFORMÁCIÓK JELENTŐSÉGÉNEK ÉS SZEREPÉNEK VÁLTOZÁSA

Az információs társadalomban az *információhoz való hozzáférés* elemi érdeke mindenkinek. Azok, akiket elzárnak az információtól, vagy valamilyen oknál fogva nem képesek ahhoz hozzájutni, hátrányba kerülnek azokhoz mérten, akik képesek arra. Az információhoz való hozzáférést biztosíthatják *ingyenesen*, vagy *térítés ellenében*; lehet az információ *nyílt* vagy *minősített*; elérhetjük valamilyen *nem elektronikus módon* (nyomtatásban, filmen, stb.), vagy *elektronikusan*.

Azzal, hogy az információt valamilyen térítési díj ellenében szolgáltatják, már szűkítik a hozzáférők körét. Ha az információt titkosítják, magától értetődően ezzel is sokakat fosztanak meg attól, hogy azt megismerjék. A nem elektronikus formában rendelkezésre álló információhoz csak korlátozott mértékben lehet hozzáférni, mivel azt csak bizonyos példányszámban állítják elő. Természetesen az elektronikus információknak is van korlátozottságuk, például a műsorszórók adáskörzetének nagysága, vagy a vételhez szükséges eszköz hiánya.

Az elektronika fejlődése lehetővé tette, az információ előállítás és terjesztés terén, analóg technika felváltását a digitálissal. A digitális elektronikus jel feldolgozása, szerkesztése, másolása, tömörítése, terjesztése sokkal olcsóbb az analógénál, valamint sokkal ellenállóbb a torzulás, veszteségek és a zajok ellen. Ez az oka annak, hogy a technikai fejlődés abba az irányba halad, hogy az információkat eleve digitális alakban állítsák elő, és az analóg módban meglévőket is átalakítsák erre a formátumra. *A digitális elektronikus jelek alkalmasak az elektronikus úton, így számítógépes hálózaton (interneten) keresztül való terjesztésre is.* Az internetes adatátvitel fontos előnye, hogy az információk már eleve digitális elektronikus jelként állnak rendelkezésre, így megtakarítható az esetleges átalakítás ideje és költsége, ezáltal azonnal megkezdhető azok gépi feldolgozása.

1.7. AZ INTERNET ALAPÚ INFORMÁCIÓSZERZÉS KAPCSOLATA A TÖBBI NYÍLT FORRÁSÚ CSATORNÁVAL

Az internet fejlődése során egymás után jelentek meg újabb és újabb szolgáltatások. Kezdetben a hálót csak arra használták, hogy a tudósok megosszák egymás között az ismereteiket. Az elektronikus levelezés 1971-es megjelenése volt az első lépés a más jellegű, nem csak a hálózat fejlesztésével kapcsolatos használat felé. Az átviteli sebes-

ség és a sávszélesség fokozatos növekedése egyre több szolgáltatás megjelenését tette lehetővé. A technikai fejlesztés mellett az olyan – a használatot és az információ megtalálását megkönnyítő – programok megjelenése, mint a böngészők (*Mosaic*⁴⁶), illetve a kereső programok (*Wandex*⁴⁷, *Aliweb*⁴⁸) segítettek az internet használatának felfutásában. A szolgáltatások fejlesztése mára már azt eredményezte, hogy *az internet vált az információszerzésnek és a kommunikációnak a legnagyobb médiumává.*

Mára már minden *hírlapnak, napilapnak, hetilapnak, folyóiratnak, magazinnak* van internetes honlapja, ahol részben vagy egészében, ingyen, vagy előfizetési díj ellenében elolvasható az aktuális számuk. Emellett a régebbi számokat is le lehet kérni, így egy adott témakörnek, eseménysorozatnak utólag is meg lehet ismerni az előzményeit, kiváltó okait és hátterét. A nyomtatott sajtó mellett már jó ideje létezik az *elektronikus sajtó* is, amely a hírszolgáltatás új dimenzióját nyitja meg. Itt már nem létezik lapzárta, az információk folyamatosan – a történésekkel közel egyidőben – érkeznek. A hírek nem csak a megszokott álló, hanem mozgó képekkel is tarkítottak. Emellett a felhasználók gyors informáltságát elősegítendő, ún. *RSS*⁴⁹ *hírcsatornákra* is fel lehet iratkozni, amelyek az új hírek megjelenésről azonnali tájékoztatást adnak. A nyomtatott sajtó szerepe csökken, helyettük a *hírportálok* (pl. *Origo*⁵⁰, *Index*⁵¹) a médiapaletta megkerülhetetlen részévé, meghatározó *tömegmédiummá válnak*. Saját stábjuknak köszönhetően a hírportálok egyre gyakrabban jelennek meg *elsődleges hírforrásként*.

A szélessávú internetezés lehetővé tette az internetes *rádióközvetítést* (IP rádió). Hallgathatók a földi műsorszórók adásai, illetve olyanok is, amelyek csak hálózati terjesztésűek. Ennek köszönhetően a műsortovábbításban megszűnik az a korlát, amit az elektromágneses hullámterjedési tulajdonságok állítanak. Nem csak az adók sugárzási körzetében, hanem a világ bármely pontján lehet az internetes rádióadásokat hallgatni, csak megfelelő sávszélességű összeköttetés kell hozzá (legalább 32 kbit/s).

A minőségi *internet alapú televízióadások* (*IPTV*⁵²) széleskörű elterjedésének még egyelőre határt szab az átviteli utak keskeny keresztmetszete és lassúsága, de nap-

⁴⁶ A Mosaic volt az első böngésző program, 1992-ben kezdték a fejlesztését, 1993-ban jelent meg az első változata.

⁴⁷ Wandex: World Wide Web Wanderer volt az első kereső program, 1993-ban fejlesztették ki a MIT-en, ma már nem létezik.

⁴⁸ Aliweb: Archie Like Indexing for the Web még ma is elérhető kereső program, honlapja: <http://www.aliweb.com/>.

⁴⁹ Weboldalak változásait tükröző, rövid XML fájl. Az RSS-nek több meghatározása is van: Really Simple Syndication, Rich Site Summary vagy RDF Site Summary.

⁵⁰ Honlapja: <http://www.origo.hu/>.

⁵¹ Honlapja: <http://www.index.hu/>.

⁵² IPTV: Internet Protocol Television: internet protokoll alapú televízió.

jainkban terjedőben vannak a szélessávú kapcsolatok (*ADSL*⁵³, TV kábel). Jelenleg, jellemzően, csak kis felbontásban (általában 320x240) lehet tévézni, de így elegendő a keskenyebb sáv szélességű kapcsolat is (min. 32 kbit/s), ami még csak nagyon gyenge kép és hangminőséget biztosít. A megfelelő minőséghez legalább 1024 kbit/s-os összeköttetés szükséges, ami már több csatorna vételét is lehetővé teszi⁵⁴.

A mai *könyvtárba* már nem kell feltétlenül elmenni ahhoz, hogy igénybe vehessük a szolgáltatásait, mert az *online módon rendelkezésre álló adatbázisok* segítségével, tájékozódni tudunk a mindenkori állományról, és azokat előfizetés ellenében le is tölthetjük. Ily módon a katalógus fogalma kitágult, hiszen rég átlépték már a könyvtár falait, és nem csupán egy adott könyvtárban fellelhető dokumentumokról adnak tájékoztatást. A kitűnő technológiai lehetőségek és a közös felhasználói felület miatt nincs különbség a fizikailag egy helyen tárolt, illetve a virtuálisan közös (ám fizikailag elkülönülő) katalógusok között.

A hagyományos könyvtárak mellett viszonylag hamar kialakultak az *elektronikus könyvtárak* is. Létrehozásuk alapötletét az adta, hogy az interneten egyre növekvő számban jelentek meg teljesszövegeű dokumentumok. Könyvtárosok, könyvtáros lelkületű emberek pedig úgy gondolták, hogy ezeket az elektronikus dokumentumokat valamilyen formában össze kellene gyűjteni és feldolgozni, különben elvesznek a hálózat egyre jobban növekvő elektronikus dzsungelében. Erre a kihívásra válaszul jöttek létre az *elektronikus könyvtárak*, amik digitális formában tárolt *dokumentumok rendszerezett, feldolgozott, tartósan megőrzött és egy közösség számára hozzáférhetővé tett gyűjteménye*. Fontos tulajdonsága, hogy ide nem alkalomszerűen elhelyezett dokumentumok kerülnek, hanem minden dokumentum keresztülmegy a feldolgozó szakember kezén, aki ugyanúgy, mint a hagyományos értelemben vett könyvtárakban, egyfajta bibliográfiai leírást készít hozzá, s a könyvtár rendszere szerint sorolja be annak állományába. Az elektronikus könyvtár arra törekszik, hogy a könyvtár dokumentumai hosszú távon és megbízhatóan megőrződjenek és megtalálhatóak maradjanak.

Számos *könyvet, lexikont* már nem csak nyomtatott formában olvashatunk, hanem az interneten keresztül is, valamint vannak eleve csak elektronikus formában létező könyvek, ún. *ebook-ok*⁵⁵. Például *Michael S. Hart* (1947, Tacoma) amerikai számítógé-

⁵³ ADSL: Asymmetrical Digital Subscriber's Line – aszimmetrikus digitális előfizetői vonal.

⁵⁴ Magyarországon 2006 őszén indult internet alapú televízióadás.

⁵⁵ Magyarországon például a Magyar Elektronikus Könyvtárban (MEK) lehet ilyen anyagokat elérni, honlapja <http://www.mek.iif.hu/>.

pes szakember, 1971-ben megalapította – *Project Gutenberg*⁵⁶ néven – önkéntesekből álló mozgalomát, amelyben elektronikus könyveket készítenek a már nem jogvédett kötetekből, és azokat az interneten elérhetővé teszik.

Az olyan nagy *adatbázissal rendelkező szolgáltatók*, mint a vállalatok, speciális szolgáltatók, kutató intézetek, és egyetemek, egytől egyig hozzáférhetővé tették adatbázisaikat, amelyekhez ugyan többnyire térítés ellenében lehet hozzáférni, de azok olyan felbecsülhetetlenül értékes, rendszerezett, szakértelemmel megszerkesztett információkat tartalmaznak, amilyenekhez máshonnan nem lehet hozzájutni.

Napjainkra alapvető trend lett a *digitális technológia általános alkalmazása*, ami magával hozta, hogy a számítástechnika és a telefónia technológiája konvergál egymáshoz. A megindult protokollfejlesztések lehetővé teszik a *beszédátvitel* (telefonálás) *internetes alapon* való megvalósítását (*VoIP*⁵⁷).

Mivel az interneten nagyon *olcsón és gyorsan lehet közzétenni az információkat*, – összehasonlítva a hagyományos kiadási eljárásokkal – szervezetek, katonai egységek és kormányzati hivatalok, kulturális intézmények növekvő mértékben alkalmazzák arra, hogy rajta keresztül tájékoztassák az ügyfeleiket, tagságukat, vagy az állampolgárokat. Kormányzati szinten komoly fejlesztések folynak az ún. *e-kormányzat* kialakítására, ami lehetővé teszi az állampolgároknak az interneten keresztül történő tájékoztatását és az elektronikus ügyintézését. Emellett az internetes *kereskedelem* is egyre nagyobb teret hódít.

Az *egyetemek és főiskolák oktatóinak* ma már kötelező, hogy hozzáférhetővé tegyék az interneten keresztül is az általuk oktatott tantárgyak tematikáját és felkészülési kérdéseit. Sok oktató előadásainak anyagait és a tantárgyainak jegyzeteit elektronikus formában is elérhetővé teszi az interneten keresztül. A bolognai folyamat előírásának megfelelően az új tananyagoknak már a 75%-a hozzáférhető elektronikus formában is a *Scopus*⁵⁸ rendszeren keresztül. Emellett számos hallgató is közzéteszi a kész szakdolgozatát vagy diplomatervét, ezzel segítve társainak az adott témában való ismeretszerzést.

Az interneten keresztül *országok, vidékek, városok részletes térképeit*, vagy ezek műholdról készült, vagy *légi fényképeit* tölthetjük le. Ezek olyan jó minőségűek is lehetnek, hogy segítségükkel térkép is készíthető olyan területekről, amelyekről nem áll

⁵⁶ A programról bővebben a <http://www.promo.net/pg/> honlapon lehet tájékozódni.

⁵⁷ VOIP: Voice Over Internet Protocol – Hangátvitel internet protokoll felett, a témáról bővebb információt lehet szerezni a <http://voip.lap.hu/> honlapon.

⁵⁸ Egy multidiszciplináris, absztraktokat és bibliográfiai adatokat tartalmazó, URL címek megadásával linkeléssel komplex szolgáltatások nyújtására is alkalmas adatbázis.

rendelkezésre digitális térkép. Megnézhetjük fotóművészek, fotóriporterek és amatőr fotósok fényképeit, de feltehetjük akár az általunk készített digitális formában meglévő, képeinket is. Ellátogathatunk valamelyik múzeumba, vagy egy kiállításra úgy, hogy a fotelünkben ülünk otthon, de ugyanígy akár virtuális sétát is tehetünk egy-egy nagyváros utcáin.

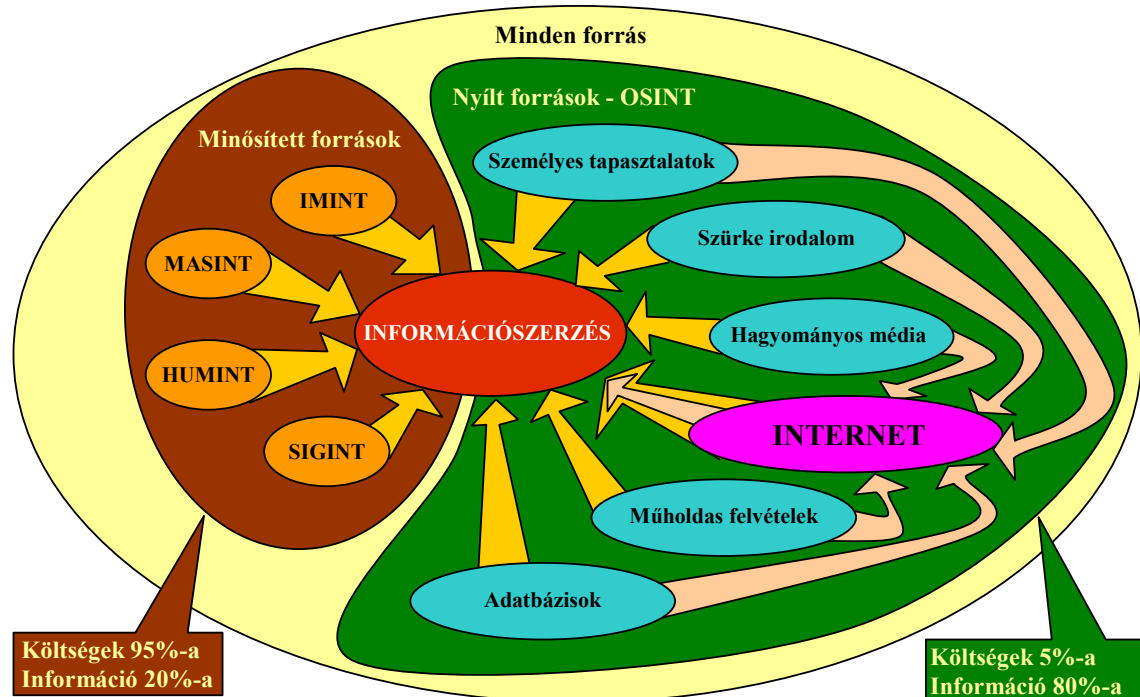
Az internet állandó figyelése esetén lehetőség van akár *minősített anyagok megszerzésére* is. Erre az ad lehetőséget, hogy valaki véletlenül vagy tévedésből teszi hozzáférhetővé a számítógépén tárolt anyagait, amire nagyon sok a példa. Előfordul az is, hogy szándékosan, rosszakaratból, ártó szándékkal teszik hozzáférhetővé mások értékes, nem a nyilvánosságnak szánt, esetleg kompromittáló, bizalmas adatait. Az ilyen információk többnyire csak nagyon rövid ideig érhetők el, de folyamatos figyélssel és kereséssel meg van az esély a megtalálásukra. Az ilyen információkat más forrásból is meg kell erősíteni, mert azok hitelessége nem mindig ellenőrizhető.

A *Web 2.0*⁵⁹-nak nevezett kezdeményezésnek köszönhetően az *emberek közti kapcsolatok* teremtésének, építésének és ápolásának fontos terepévé és eszközévé válik a világháló. Segítségével különböző *közösségek jöhetnek létre*, amit a közös érdeklődés, vagy más közös tulajdonság megléte alakíthat ki és tarthat fenn. A Web 2.0 egy új irányzat, szemlélet a webes szolgáltatások készítése kapcsán. A szolgáltatások az internetet használók tartalomkészítő, önszervező erejét próbálják meg összefogni, és ebből újat alkotni, modern eszközöket, felhasználóbarát szemléletet és nyílt megoldásokat használva az elkészítéskor. Így válik az internet a felhasználók számára egy barátságos, sok lehetőséget, *nagymennyiségű információ elérését biztosító új médiává*.

Összegezve, az interneten keresztül jószerivel már minden nyílt forráshoz hozzá lehet férni részben, vagy teljes egészében (1.2. ábra). A jövőben tovább növekszik az interneten hozzáférhető szolgáltatások száma és ezzel arányosan a rajta lévő információk mennyisége. A tendencia az, hogy minden – TV, rádióadások, újságok, telefon, könyvek, múzeumi anyagok, képek, de még az otthoni háztartási berendezések is – elérhetővé válik a világhálón keresztül, vagyis minden szolgáltatás egy helyre, az internetbe integrálódik. Az internet ezért 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 egy óriási, sokkultúrájú elektronikus könyvtár, ami nyitva van éjjel, nappal az információt kereső fel-

⁵⁹ A 2004-es év közepén bukkant fel először ez a fogalom, Tim O'Reillynek és Dale Doughertynek köszönhetően.

használóknak, és ontja magából az időszerű, az élet szinte minden területére kiterjedő információt annak, aki rendelkezik megfelelő képességekkel és eszközökkel az információ ezen különleges bányájának kiaknázására.



1.2. ábra A mindenoldalú információszerzés adatforrásai

1.7.1. Az internetes információszerzést nehezítő tényezők

Az internetes információszerzés könnyű, gyors és olcsó, de vannak olyan tényezők, amelyek megnehezítik ezt a munkát.

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, ezért nem elég a böngészőben könyvjelzővel megjelölni az oldalakat a későbbi megtalálás érdekében, hanem azonnal le kell tölteni és archiválni kell a fellelt információt.

A világhálón lévő *adatokat senki sem rendezi* valamilyen logika szerint, nem gondozzák, *nincs* az internetnek egy általános *tartalomjegyzéke*, ami alapján megtalálhatnánk a rajta levő, számunkra fontos információkat. Pontosabban ennek a munkának még csak az elején tartanak, de már vannak biztató jelek (pl. szemantikus web). Az internetes keresés jártasságot, kitartást és minőségi kereső eszközöket igényel.

A megtalálható információk *hitelessége nem mindig biztosított*, mivel a közzétett információkért a tárhely szolgáltatók nem felelnek, kizárólag csak akkor avatkoznak be,

és szüntetik meg az adott oldalt, vagy kötelezik a felhasználót annak tartalmának törlésére, ha azzal valamilyen törvényt, vagy előírást sért meg (pl. uszít valamelyik kisebbség ellen).

Az információt *feltöltő személyek kilétét, sok esetben, gyakorlatilag lehetetlen kideríteni*, mert a szolgáltatók az általuk kért adatokat nem mindig ellenőrzik vissza, így bármilyen hamis adatot is megadhatnak a felhasználók (pl. nem létező, ún. „eldobható” email címet) az azonosításukra. Ezért szinte minden következmény nélkül közölhetnek bárkiről, vagy bármiről bármit (lásd egykori besúgók és titkos ügynökök neveinek nyilvánosságra hozása).

Az internet segítségével nagyon sok információhoz lehet hozzáférni, de fontos megjegyezni, hogy azok mennyiségének gyors növekedése ellenére, a nyílt forrásoknak még *kevesebb, mint a fele érhető el* jelenleg a hálózaton keresztül, a többi nyomtatásban, vagy más formában áll csak rendelkezésre.

Az internet tehát egy olyan információforrás, aminek a tartalma gyorsan változik, nagyon rosszul szervezett, valamint kétes hitelességű, a rajta lévő információk megtalálásához nagyon pontosan meghatározott keresési terv és egyre kifinomultabb számítógépes eszközök szükségesek.

1.8. ÖSSZEGZÉS ÉS KÖVETKEZTETÉSEK

1. A kétpólusú világrendszer megszűnése utáni demokratikus átalakulásoknak, valamint a kommunikációs és informatikai technológia gyors, nagyléptékű fejlődésének, elterjedésének köszönhetően a nyílt információk mennyisége nagymértékben megnőtt. Az információs társadalomban a nyíltság és az átláthatóság (transzparencia) igénye kerül előtérbe, mert az információk védelme megnöveli a költségeket és nehézkessé teszi a munkavégzést, ezért a gyorsaság válik fontossá. Aki meg tudja előzni a versenytársait az információ megszerzésben és feldolgozásában, az előnybe kerül a többiekkel szemben.

A nyílt információk mennyiségének megnövekedése és dominánsná válása lehetővé teszi, egyben ki is kényszeríti a nyílt forrású felderítés végzését. Ez a felderítési mód szerves része a minden oldalú felderítésnek. Bizonyított tény, hogy a leggyorsabb információszerzést a nyílt források teszik lehetővé, mivel azok a nap minden pillanatában keletkeznek és rendelkezésre állnak. . Váratlan krízis helyzetekben sokszor az egyetlen rendelkezésre álló információforrások.

2. Az olyan nem elektronikus formában elérhető nyílt információk, mint a nyomtatott anyagok (könyv, újság), a filmen tárolt anyagok, papíralapú térképek és képek, stb. szerepe fokozatosan csökken, mert a felhasználók áttérnek a digitális formátumra.

Az elektronikus formában az információk tárolása, kezelése és keresése egyszerűbb, valamint a feldolgozása olcsóbb, gyorsabb, mert nagy részben gépesíthető. Ezek a jó tulajdonságai az okai annak, hogy a hagyományos, nem elektronikus információkat is fokozatosan átalakítják ebbe a formába. Elektronikus formában az információkat gyorsan, egyidőben sok felhasználóhoz lehet eljuttatni, amelynek eszközei az elektronikus média és az internet.

3. Az elektronikában napjainkban végbemenő gyors változások eredményeként az elektronikus információtovábbítás technikájában és eszközeiben (rádió, TV, telefon, internet) egyfajta konvergencia figyelhető meg.

Az egyes szolgáltatások között elmosódnak az őket elválasztó határok. Ennek a konvergenciának – a mai fejlődési irányokat figyelembe véve – az lesz a végeredménye, hogy az internetbe integrálódik a rádiós és televíziós műsorszórás (elektronikus média), valamint a beszédátvitel (telefon).

4. A nyílt anyagok hálózatos elérése válik jellemzővé (e-book, elektronikus könyvtár, hírportálok, e-kormányzat, elektronikus oktatóanyagok, e-kereskedelem, elektronikus adatbázisok, digitális képek és térképek, stb.).

A változások alapján megállapítható, hogy az internet válik a legfontosabb nyílt információforrássá, amely folyamatosan, a nap huszonnégy órájában rendelkezésre áll, gyors, olcsó, sokrétű tájékoztatás nyújt (szöveges, képi, hang), az új információkat a történésekkel közel egyidőben juttatja el a felhasználóhoz, a megszerzett anyagok könnyen tárolhatók, kereshetők és feldolgozhatók.

A nyílt információszerzést nehezítő tényező, hogy a megszerzhető információ mennyisége hatalmas. Ennek a nagymennyiségű anyagnak a feldolgozása kézi módszerekkel lehetetlen, csak számítógépek alkalmazásával oldható meg, mert azok lehetővé teszik számos feladat automatikus elvégzését.

A gyors információszerzés, -feldolgozás és -elosztás követelménye miatt szükség van az internet folyamatos figyelésére, amit egy erre a feladatra kialakított, nagymértékben automatizált számítógépes rendszer végez. Egy ilyen rendszer képes lehet arra, hogy megtalálja a felhasználók által feltett kérdések megválaszolására alkalmas forrá-

sokat, emellett figyelemmel kísérje az aktuális eseményeket (*current information*). Az automatikusan begyűjtött információkból értékelés, archiválás, szelektálás, csoportosítás, elemzés és feldolgozás után megadott határidőre, vagy megadott időközönként jelentések, tájékoztatók, háttéranyagok készíthetők, amik a hálózati kapcsolatok révén a felhasználókhoz el is juttathatók. Egy ilyen rendszer felválthatja a mai egy-két emberrel, nem szervezetszerűen végzett – manuálisnak nevezhető – kereső munkát egy szervezett keretek között, meghatározott létszámmal, rendszertechnikailag optimálisan kialakított számítógépes eszközökkel folytatott információszerzéssel.

2. AZ INTERNET ÉS A KERESÉSI TECHNIKÁK FEJLESZTÉSÉNEK IRÁNYAI

Az internet kezdeti átviteli sebessége csak néhány kbit/s⁶⁰ volt. Ez elegendő volt az addig kifejlesztett programok futtatására, de már akkor is gondot jelentett a nagyobb méretű fájlok le- és feltöltése. Az internet megadott átviteli sebessége csak elméleti jellegű, mert az egyidejű felhasználók számával arányosan csökken annak pillanatnyi értéke. Ez az oka annak, hogy az átviteli sebesség növelése mindig a fejlesztések homlokterében volt.

Az internet népszerűvé válásával együtt újabb és újabb szolgáltatásokat fejlesztettek ki a hálózat lehetőségeinek kihasználására. A 90-es évek elejére az alkalmazott protokollokkal kapcsolatban már előre lehetett látni, hogy azok képességei nem teszik lehetővé a továbbfejlődést, ezért elkezdtek azok fejlesztését.

2.1. AZ INTERNET TECHNIKAI FEJLESZTÉSE

Az internet kifejlesztése az Egyesült Államokban kezdődött és a mai napig itt folynak a legfontosabb fejlesztői projektek. Az átviteli sebesség és a protokollok fejlesztésére 1996-ban elindították az *Újgenerációs Internet Kezdeményezésnek (NGII)*⁶¹ nevezett programot⁶², ami két összetevőből állt: az *NGI*⁶³ és az *Internet2*. Ezek közös célja, hogy kialakítsanak egy nagy sebességű (Gbit/s) hálózatot.

Az internetet a használt technológia és a támogatott alkalmazások alapján három generációba lehet osztani:

☞ az *első generációs* internet jellemzői:

- megbízhatatlan;
- Ipv4 protokollt használ;
- nincs minőségi szolgáltatás (QoS⁶⁴);
- nincs garantált sávszélesség;
- alacsony átviteli sebesség.

⁶⁰ 1986-ban az NFSNet hálózat átviteli sebessége 56kbit/s volt.

⁶¹ NGII: Next Generation Internet Initiative – Újgenerációs Internet kezdeményezés.

⁶² A programot részleteiben tárgyalja a Jagannathan, Prasanna Kumar: Global Next Generation Internet Initiatives, CIS, Ohio, 1999.

⁶³ NGI: Next Generation Internet – Újgenerációs Internet.

⁶⁴ QoS: Quality of Service – Minőségi szolgáltatás. Hálózatok és hálózati eszközök képessége az erőforrások meghatározott rend szerinti felosztására, és garantált sávszélesség biztosítására.

☞ a *második generációs* internet jellemzői:

- IPv6 protokollt használ;
- multicast címzés;
- van minőségi szolgáltatás (QoS)
- van garantált sávszélesség.

☞ a *harmadik generációs* internet (még csak kutatási stádiumban van *SUPERNet* néven) jellemzői:

- igen nagy átviteli sebesség mellett minden fejlett szolgáltatást képes biztosítani;
- DWDM használata⁶⁵;
- optikai rétegek használata routolásra és a kapcsolásra.

A továbbiakban csak a második generációs internet összetevőit ismertetem.

2.1.1. Az Újgenerációs Internet (NGI)

Az USA következő kormányzati ügynökségei alkotják az NGI-t: *DARPA*⁶⁶, *DOE*⁶⁷, *NASA*⁶⁸, *NIH*⁶⁹, *NIST*⁷⁰ és az NSF. A projekt indításának célja két hálózat kifejlesztése volt. Az egyik százszor gyorsabb, mint a jelenlegi internet és 100 intézet csatlakozik rá: *Net1*, amely a következő hálózatokból áll: *vBNS*⁷¹, *NREN*⁷², *DREN*⁷³, *ESNET*⁷⁴, *NISN*⁷⁵. A másik ezerszer gyorsabb a mostani internetnél és 10 intézet csatlakozik rá: *Net2*, ami csak a *SUPERNet*-ből áll.

A *vBNS* hálózatot 1995-ben indították. Nagy sebességű, szélessávú, üvegszálak kapcsolatát létesít a nagy kutatóintézetek, az egyetemek és a szuper-számítóközpontok kiválasztott csoportjának tagjai között, másodpercenként 622 Mbit/s információátviteli sebességgel. A hálózatot csak a szélessávú alkalmazáskutató projekteken résztvevők használhatják, ami azt jelenti, hogy nem használható az általános internet forgalom le-

⁶⁵ DWDM: Dense Wavelength Division Multiplexing – sűrű hullámhossz osztásos multiplexálás. A DWDM esetén általában több mint 8 fényablakot alkalmaznak adó és vevő oldalon. 16, 40 és 80 ablakos rendszerek az általánosan elterjedtek. Matematikailag 111 ablak lehetséges egyetlen optikaiszál-páron a ma használt hullámhosszakkal.

⁶⁶ DARPA: Defense Advanced Research Projects Agency – Fejlett Kutatások Védelmi Ügynöksége.

⁶⁷ DOE: Department of Energy – Energia Minisztérium.

⁶⁸ NASA: National Aeronautics and Space Administration – Nemzeti Légügyi és Űrhajózási Hivatala.

⁶⁹ NIH: National Institutes of Health – Nemzeti Egészségügyi Intézet.

⁷⁰ NIST: National Institute of Standards and Technology – Nemzeti Szabványügyi Intézet.

⁷¹ vBNS: Very high speed Backbone Network Service – szupergyors gerinchálózati szolgáltatás.

⁷² NREN: NASA Research & Education Network – NASA Kutatói és Oktatói Hálózatok.

⁷³ DREN: Defense Research & Engineering Network – Védelmi Kutató és Tervező Hálózat.

⁷⁴ ESNET: Energy Sciences Network – Energia Tudósok Hálózata.

⁷⁵ NISN: NASA Integrated Service Network – NASA integrált szolgáltatású hálózata.

bonyolítására. A vBNS továbbfejlesztett változata a vBNS+, amit 1999-ben indított el az NFS⁷⁶ és az MCI WorldCom⁷⁷.

Az *NREN* projekt a NASA fő hozzájárulása az NGI-hez. Ebben szövetségi ügynökségek, akadémiai szervezetek és iparvállalatok dolgoznak együtt. Céljuk a nagysebességű hálózatokhoz olyan új technológiák és szolgáltatások kifejlesztése, amelyek forradalmi alkalmazások használatát teszik lehetővé, elsősorban a NASA számára. Az egyik fő alkalmazási terület a *Digital Earth* nevű program, amelyben a tudósok a Föld dinamikáját és atmoszféráját vizsgálják nagyfelbontású képek és nagymennyiségű adatok felhasználásával, amelyek méretei már TB nagyságrendűek.

A *DREN* hálózatban az amerikai *Védelmi Minisztérium* (DoD⁷⁸) *Nagyteljesítményű Számítástechnika Központja* (HPC⁷⁹) és a *WIDE*⁸⁰ központja együtt került összekötésre a fent említett hálózatokkal. Az *ATM*⁸¹ és IP szolgáltatások közötti együttműködési lehetőséget vizsgálják. Kezdetben a DREN több mint 60 helyéből körülbelül 25 volt ATM, a többi IP hálózat volt. A hálózat feladata elsősorban a védelemmel kapcsolatos kutatások és fejlesztések végzése.

Az *ESNET* az Energia Tudósok nemzetközi, ATM gerincű (backbone) hálózata. Egyike azoknak a fő hálózatoknak, amelyeken együttműködnek egymással az amerikai, ázsiai és európai hálózatok. Jelenleg 10 Gbit/s-os (OC-192⁸²) sebességen működik. A kutatók az IPv6-ot (a 6Bone keretein belül), az osztály alapú sorbaállást (CBQ⁸³), a minőségi szolgáltatást (QoS) és a virtuális hálózatokat vizsgálják.

A *NISN* a NASA integrált szolgáltatású hálózata, amelynek a célja, hogy egyetlen hálózattal helyettesítsék a NASA sok különálló hálózatait. Elsősorban a NASA feladatainak a végrehajtására hozták létre. A legnagyobb sebessége 155 Mbit/s (OC-3).

2.1.2. Az Internet2

Az Internet2 egy konzorcium, amiben több mint 200 amerikai egyetem, az informatikai ipar és kormányzat munkacsoportjai vesznek részt. A fejlesztő munka az NGI-vel párhuzamosan fut. Ez a nagysebességű (gigabites nagyságrendű) IP hálózat a jelenlegi in-

⁷⁶ NSF: National Science Foundation – Nemzeti Tudományos Alapítvány.

⁷⁷ MCI Microwave Communication Incorporation – Mikrohullámú Kommunikáció Egyesület.

⁷⁸ DoD: Department of Defense – Védelmi Minisztérium.

⁷⁹ HPC: High Performance Computing – Nagyteljesítményű Számítástechnika.

⁸⁰ WIDE: Writing In Digital Environments – Digitális környezetben írt. A Michigan Állam Egyetemének számítástechnikai kutató központja.

⁸¹ ATM: Asynchronous Transfer Mode – Aszinkron átviteli mód.

⁸² OC: Optical Carrier – optikai vivő. Az optikai kábeleknél alkalmazott sebesség meghatározás.

⁸³ CBQ: Class Based Queuing – egy módszer hierarchikus forgalomszabályozásra.

ternet több problémáját (lassú vagy változó átviteli kapacitás, véges számú IP címek stb.) kívánja megoldani és új hálózati alkalmazásokat bevezetni.

Az internet2 négy fő kutatási és fejlesztés projektből áll: *QBONE*⁸⁴, *I2DVN*⁸⁵, *I2DSI*⁸⁶ és *Abilene*⁸⁷.

A *QBONE* hálózat célja a minőségi szolgáltatás (QoS) vizsgálata IP protokoll felett, valamint olyan hálózat kialakítása, ami képes fejlett alkalmazások futtatására, multiplex hálózati eszközök implementálására, emellett jól skálázható, adminisztrálható és mérhetőek a szolgáltatásai. A fejlesztés keretén belül még új operációs rendszerek és felhasználói szoftverek kidolgozása is folyik.

Az *I2DVN* hálózat az Internet2 digitális videó hálózata, ami nem más, mint a találkozási pontja számos digitális videózással kapcsolatos kezdeményezésnek. A kutatási tevékenység magába foglalja az élő videó átvitel, a videó konferenciaadások és a „videó kérésre” (video on demand) szolgáltatások fejlesztését.

Az *Abilene* egy nagyteljesítményű gerinchálózat, amely lehetővé teszi az Egyesült Államok egyetemeinek és kutatólaboratóriumainak, hogy nagyteljesítményű internetes alkalmazásokat fejlesszenek. Ez a hálózat vált az elérhető legfejlettebb gerinchálózattá az Internet2 projektben résztvevő egyetemek számára. A hálózat támogatja az olyan fejlett alkalmazásokat, mint a virtuális laboratórium, a digitális könyvtár, a távoktatás és a távjelenlét (tele-immérió), valamint a fejlett hálózati képességek.

Az *I2DSI* hálózat magába foglalja a megbízhatóság, skálázható tárolási struktúra fejlesztését, valamint a vBNS-hez és az Abilenehez hasonló fejlett hálózatok jó tulajdonságainak a lehető legjobb kiaknázását, az internetes tartalmak és programszolgáltatások tükrözését.

Magyarország a Nemzeti Információs Infrastruktúrafejlesztési Program (NIIF) keretén belül már csatlakozott az Internet2 különböző projektjeihez (HUNGARNET⁸⁸).

Bár az áttérés az Internet (1)-ről a 2-re kimondhatatlanul magas összegbe kerül, mégis remélhető, hogy az évtized végére az átlag fogyasztó számára is elérhető lesz egy nagyobb sebességű internet.

⁸⁴ QBONE: Quality Bone – minőségi gerinchálózat.

⁸⁵ I2DVN: Internet2 Digital Video Network- Internet2 Digitális Video Hálózat.

⁸⁶ I2DSI: Internet2 Distributed Storage Initiative – Internet2 Szétosztott Tárolás kezdeményezés.

⁸⁷ Abilene: Advanced Networking for Leading-edge Research and Education – Fejlett hálózati szolgáltatás vezető kutatóknak és oktatóknak.

⁸⁸ Az NIIFI által létrehozott és üzemeltetett magyar kutatói, felsőoktatási és közgyűjteményi hálózat.

2.2. AZ INTERNET PROTOKOLL FEJLESZTÉSE

Az internet a kezdeti kísérleti hálózatból egy világot átfogó rendszerré növekedett a 90-es évek elejére. Kezdetben nem volt egységes protokoll rendszere, ami szabályozta volna az összehangolt működést. Ehhez képest volt előrelépés az *NCP*⁸⁹ *protokoll* bevezetése. Ez a protokoll a 80-as évek elejére már elavult, ezért kidolgozták a TCP/IP protokollt, amire 1983. január 1-én átállították az egész internetet⁹⁰. *A jelenleg használt internet protokoll a 4-es verziószámot viseli (IPv4).*

Az internet rendkívül gyors növekedése számos problémát hozott a felszínre, ami elsősorban a kiosztható címek számának elfogyásában és a hiányos szolgáltatásokban jelentkezett. Habár a jelenlegi protokollrendszer még ki tudja szolgálni a világ több százmillióra duzzadt felhasználóinak táborát, de számos ponton elérte a lehetőségeinek határát, ami a továbblépésben már csak korlátozott. Különböző kiegészítésekkel, kisebb-nagyobb módosításokkal, pótfunkciók hozzáadásával a korlátok tágíthatók és ezzel meg lehet hosszabbítani a használatát. Hosszú távú megoldást csak egy újabb verzió kidolgozása és bevezetése jelenthetett. Ennek az új verzióknak a neve IPv6, vagy közkeletűbb nevén *új generációs internet protokoll*⁹¹ (IPng).

2.2.1. Az IPv4 problémái

A legfontosabb probléma, amelynek kapcsán felvetődött egy új IP verzió kifejlesztése, az IPv4 címterének korlátozottsága volt. Az IPv4 32 bites címhosszúsága elvileg 2^{32} , azaz kb. 4 milliárd egyedi IP-cím előállítását teszi lehetővé, de az osztályozás és a cím disztribúciós eljárás pazarló volta miatt rendkívül sok cím kihasználatlan marad, elvész. Ennek a korlátnak a feloldására több megoldást is kidolgoztak, de végleges megoldást csak az IPv6 hozhat.

A hagyományos „számítástechnikai” felhasználáson túlmenően új piaci szegmensek is megjelentek. Az egyik a szórakoztatás hálózatra települése, „video on demand”, internet IP-rádió vagy IPTV, stb. A másik a különböző intelligens berendezések megjelenése. Várhatóan valamilyen processzor fogja vezérelni a legtöbb világító, hűtő és fűtőberendezést, vagyonvédelmi, katonai eszközöket, stb. és ezek is természetesen távolról is vezérelhetőek lesznek. Az IPv4 protokoll legfontosabb hibái a következők:

- ☞ kicsi a címtér, pontosabban fogalmazva kevés a hálózati szám;

⁸⁹ NCP: Network Control Protocol – hálózatvezérlési protokoll.

⁹⁰ Ez volt az ún. flag-day. 1000 hoszt kapcsolódott akkor a hálózatra.

⁹¹ Internet Protokoll Next Generation – IPng.

- ☞ a routing táblák túl nagyméretűek;
- ☞ kevés lehetőség van egy IP csomag tartalmának vagy feladójának hitelesítésére;
- ☞ nem nyújt minőségi szolgáltatást (QoS);
- ☞ nem vesz figyelembe olyan szempontokat, mint a mobilitás vagy az autokonfiguráció.

2.2.2. Az IPv6 jellemzői

Az új protokoll az IPv6 (a 6-os számú verzió), vagy IPng (IP next generation - újgenerációs IP), ami az IPv4 szerves továbbfejlesztése, vagyis nem teljesen új protokoll, hanem evolúciós lépés. A négyes utáni hatos verziószámot az indokolja, hogy az 5-ös verziószámot az internet *Stream Protocol* (ST) kapta meg, ami az IP-vel azonos szinten működő, erőforrásokat folyamatok számára lefoglaló és így adatot továbbító protokoll. Ez nem az IP kiegészítése garanciákat nyújtó elemekkel, hanem az IP mellé rendelt új mechanizmus.

2.2.2.1. Kiterjesztett címzési és útvonal választási lehetőségek

Az egyik legfontosabb probléma volt az IPv4 címtérnek korlátozottsága. Az IPv6-nál 32 bit-ről 128 bit-re növelik az IP-cím hosszát. Első pillantásra az IP címzés ekkora mértékű kiterjesztése kissé túlzónak tűnik. Láthatjuk, hogy a 32 bit korlátozottá vált, de a 128 bitre való kiterjesztés azt jelenti, hogy $3 \cdot 10^{38}$ cím létezik, ami annyit jelent, hogy $6 \cdot 10^{23}$ cím esik a Föld felületének minden egyes négyzetméterére. A legkevesebbé hatékony hierarchikus címkiosztás megvalósítása mellett, a legóvatosabb becslések szerint is több ezer cím jut egy négyzetméterre. Lehetségesnek tűnik, hogy az elkövetkezendő néhány évben minden elektronikus árucikk kap egy IP címet. Emellett még talán fontosabb, hogy lehetőség van a hierarchikus címzési architektúra kialakítására, szemben az IPv4 címzési rendszerével. A hierarchikus címkiosztás lehetővé teszi, hogy az útvonalválasztás gyorsabbá váljon az útvonalak csoportosíthatóságának (aggregáció) köszönhetően.

2.2.2.2. Címzési architektúra

Az IPv4-hez hasonlóan az IPv6-ban is hálózati interfészekhez rendelnek IP-címeket, nem gépekhez, vagy csomópontokhoz. *Egy interfésznek több IP címe is lehet.* Az IPv4-ben csak az útvonalválasztóknak (router) lehetett több címe, az IPv6-ban ez nem csak általános, de bizonyos funkciók ellátásának alapvető feltétele.

2.2.2.3. Címtípusok

Az IPv6-os címek három csoportba gyűjthetők annak alapján, hogy a hálózaton hogyan történik egy csomag továbbküldése:

- ☞ *Unicast cím* (egyedi cím): Leginkább ez a típus hasonlít az IPv4-ben használatos címekre. Az ilyen címekre küldött csomagok arra az egyetlen interfészre fognak megérkezni, amelyhez ez a cím van hozzárendelve. Az unicast cím egyértelműen meghatároz egy interfészt (az adott területen belül: globálisan, linkre vagy site-ra korlátozottan), és egy hozzá tartozó hosztot.
- ☞ *Anycast cím* (bárki cím): Anycast címek interfész-csoportokhoz rendelhetők. Ezeknek nem feltétlenül kell ugyanazon a gépen lenniük, sőt akár egymástól távol is lehetnek. Az ilyen címre küldött forgalmat legalább egy interfész meg fogja kapni az ezzel a címmel rendelkező csoportból. Az útválasztáskor általában a legközelebbi interfészre irányítják ezeket a csomagokat.
- ☞ *Multicast cím* (csoport cím): Multicast címek is interfész-csoportokhoz rendelhetők, hasonlóan az anycast címekhez, azonban itt a csoport minden tagja megkapja a küldött csomagokat. Az IPv4-es broadcast címek helyett az IPv6-ban ilyen címek használatosak.

2.2.2.4. Automatikus konfiguráció

Nagyméretű – sok hosztot kiszolgáló – lokális hálózatokban a hálózat menedzselése igen sok munkát jelent. Az egyik leggyakoribb és legfáradságosabb munka a hosztok hálózati konfigurálása (IP címek, default gateway /GW/, netmask, DNS, stb. beállítása). Az automatikus konfiguráció célja a csomóponti eszközök egyszerű és gyors felcsatlakoztatása a hálózatra. Lehetővé tenni, hogy a hálózatra kapcsolt berendezés (számítógép) *beavatkozás nélkül megkaphassa a főbb hálózati paramétereiket*: hálózati cím, valamint egyéb konfigurációs paramétereiket. Az IPv6 tervezésénél fontos szempont volt az automatikus konfigurálás lehetőségének biztosítása. Ennek eredményeképpen az IPv6 segítségével hálózatra kapcsolódó hosztok plug-and-play⁹² módon használhatók.

2.2.2.5. Mobilitás

Napjainkban már sok mobil IP készülék létezik, és a közeljövőben várható ennek a szegmensnek az előretörése. Mivel a mobil készülékek változtatják földrajzi helyzetüket, mindig más és más ponton kapcsolódnak a hálózatba, ezért esetükben a hagyományos útvonalválasztás nem alkalmazható, hiszen az a hálózat topológiáján alapszik. Az

⁹² Plug and play – önmagától működő.

IPv6 tartalmazza, illetve támogatja azokat a mechanizmusokat, amelyek segítségével a mobilis hosztok is kiszolgálhatók.

Minden mobilitásra képes eszköznek van egy *honos hálózata (home network)*. Az eszköz „otthon van”, ha ehhez a hálózathoz csatlakozik, ekkor globális címmel elérhető, működése nem különbözik a nem mobil eszközökétől. Mobillá válás esetén az eszköz lecsatlakozik honos hálózataról és rácsatlakozik egy idegen hálózatra (*foreign network*). Globális címével továbbra is elérhető, mert a *honi ágense (home agent)* az idegen hálózatra továbbítja az ide érkező forgalmat. Ehhez a mobil csomópont egy ideiglenes címet kap (*care-of address*) kap, amelyről értesíti ágensét (ügynökét), aki ide fogja irányítani a forgalmat.

2.2.2.6. Biztonság (security)

A mai korszerű hálózatok, illetve alkalmazások egyik legkényesebb problémája a biztonság megteremtése. Számos olyan eset van, amikor nagyon lényeges, vagy kulcsfontosságú a hitelesítés lehetősége, az adatforgalom védelme, stb. Az IPv4 protokoll készlet még nem tartalmaz szabványos, beépített, hatékony mechanizmust a biztonsági feladatok egységes megoldására. Az IPv6 az *IPSec* eljárást standard szolgáltatásként kínálja, kész eszközt nyújtva ezáltal a legtöbb biztonsággal kapcsolatos feladathoz.

Az IPng protokollban *két beépített opció van*, amelyek biztonsági szolgáltatásokat nyújtanak. Ez a két opció használható külön-külön, vagy együtt különböző biztonsági szintek elérésére a más és más biztonsági igényekkel rendelkező felhasználói csoportok számára.

Az első mechanizmus az ún. *hitelesítő fejléc (Authentication Header)*, amely hitelességet és integritást biztosít (titkosság nélkül) a datagramok számára. Ez a mechanizmus használható a hálózati támadások egy jelentős részének a visszaverésére. A másik protokoll az *elzárt biztonsági csomag (Encapsulating Security Payload)*. Ez a mechanizmus integritás ellenőrzést és titkosságot nyújt a datagramoknak, vagyis védi az üzenet tartalmát a lehallgatások ellen, valamint korlátozott védelmet nyújt a forgalmi adat-analízisen alapuló támadások ellen.

2.2.2.7. QoS támogatás és nagy sávszélesség igényű alkalmazások

Az adat- és hang-kommunikáció, az audió és videó rendszerek napjainkban tapasztalható gyors integrációja eredményeként egyre több olyan új, multimédiát is használó alkalmazás jelenik meg az informatikában, amelyek a kielégítő minőség biztosítása érdekében az adathálózattól igénylik a QoS (Quality of Service) támogatását (garantált sáv-

szélesség), illetve nagy mennyiségű adat igen gyors továbbítását. Az IPv6 fokozott támogatást nyújt ezen igények kielégítéséhez, egyebek között az úgynevezett „flow” bevezetésével.

Az IPv6 tervezésekor gondoltak arra is, hogy az egyes csomagokat különbözőképpen kell kezelni, azaz szükséges a teljesítményt optimalizálni, megbízhatóságot, költséget, s biztosítani – amennyire a hálózat forgalma megengedi – a késleltetést.

2.2.2.8. Áttérés az IPv6ra

Az IPv6 még egyelőre nem terjedt el teljes körűen. Ennek egyik oka az, hogy az IPv6 lehetőségeinek jó része az IPv4 keretei között is többé-kevésbé megvalósíthatónak bizonyultak. Az adatvédelem és biztonság terén szintén komoly előrelépést jelentett az *IPsec*, a *GRE (Generic Routing Encapsulation)* és *IP tunneling* segítségével megvalósítható *VPN (Virtual Private Network)* bevezetése. A *CIDR (Classless Inter-Domain Routing)* az útvonalválasztást tette egyszerűbbé, amennyiben, legalábbis az újabban kiosztott hálózati címek hierarchikus struktúrában, kisebb táblázatokkal kezelhetők. A címtartomány kimerülését nagyon jól enyhíti a NAT (Network Address Translation), ahol a tűzfallal vagy útvonalválasztóval leválasztott hálózatban lokális érvényességű címek használhatók, és csak kevés globális cím szükséges. Ugyanilyen haszna van továbbá az *MPLS*⁹³ képességű gerinc hálózatokkal megvalósítható IP-VPN hálózatoknak is. Ezeknél a VPN azonosítóval együtt ugyanolyan hosszú, 32 bites azonosító áll rendelkezésre, mint az IPv4 esetében, de ugyanaz az IPv4 cím több IP-VPN-ben is használható, mert a cím a VPN-en kívül nem látható közvetlenül.

A fő motivációt az IPv6 világméretű elterjedésére és használatára az a szituáció fogja jelenteni, amikor majdnem minden eszköz – a hűtőgéptől a mikrohullámú sütőig az összes mobil eszközünkkel egyetemben – az internethez fog kapcsolódni, és egyedi IP címet fog igényelni.

Egy idő óta jó néhány kísérleti, sőt már kereskedelmi IPv6 hálózat is működik világszerte. Ebben a tekintetben Ázsia, különösen Japán és Dél-Korea viszi a vezető szerepet. Európában is jelentős számú IPv6-tal kapcsolatos kutatási projekt fut. Az USA is kezdi komolyan fontolóra venni az IPv6-ot. Az amerikai védelmi minisztérium 2003-ban írásba is foglalta azon célját, amely szerint a teljes IP hálózatnak IPv6-ra kell átállnia 2008-ra. Magyarországon is lehet IPv6 címtartományt regisztráltatni az NIIF⁹⁴-nél, illetve vannak működő teszhálózatok is (pl. a Budapesti Műszaki és Gazdaságtudomá-

⁹³ Multi Protocoll Label Switching – Sokprotokollós címke kapcsolás.

⁹⁴ Nemzeti Információs Infrastruktúra Fejlesztési Program

nyi Egyetemen). Figyelembe véve a jelenlegi tendenciákat úgy tűnik, hogy az IPv6 világméretű elterjedésével 2008-2010-re lehet számolni.

2.3. A VILÁGHÁLÓ KERESÉSI TECHNIKÁI

Az interneten lévő adatokat, információkat nem rendezi senki sem valamilyen logika szerinti közös tartalomjegyzékbe, ezért azok teljesen rendezetlenül, átláthatatlan összevisszaságban vannak szétszórva szerte a világhálón. Egy konkrét információ megtalálása szinte teljesen lehetetlen lenne, ha nem lennének kereső szolgáltatások.

Köztudomású, hogy az internet folyamatosan növekszik, egyre több a rajta rendelkezésre álló információ. Kérdés, hogy mennyi információ van az interneten és az nagyságrendileg, hogy viszonyul az egyéb forrásokból meríthetőkhöz képest.

2.3.1. Az internet mérete

Az internet valóságos méreteit csak megbecsülni lehet, mert az folytonosan változik. Percről percre újabb és újabb felhasználók csatlakoznak a hálózatra, új oldalak jelennek, illetve szűnnek meg. A méretek megbecslését tovább nehezíti, hogy a jelenleg használt protokoll által biztosított IP címek csak korlátozott mennyiségben állnak rendelkezésre, ami elvileg azt jelenti, hogy belátható időn belül nem lehet majd új szervert, hálózatot, munkaállomást hozzákapcsolni az internethez. A problémára számos megoldás született. Ezek jó része azon az elven működik, hogy nem szükséges minden felhasználó számára külön, a világon egyedi, globális IP címet adni, elég egy lokális cím is. Ezek egy másik gépen keresztül közös globális címen fognak látszani a hálózaton. Ez a megoldás a *NAT*⁹⁵. A módszer hátránya, hogy ezek a bizonyos felhasználók bármit elérnek az interneten, viszont őket egyenként nem lehet kívülről megcímezni. Emiatt a hosztok számának ismerete nem adja meg a gépek valóságos számát, de ismerete alapján meg lehet becsülni a valóságos méreteket.

Az *Internet Systems Consortium*⁹⁶ (ISC) féléves rendszerességgel tájékoztatást ad az internetes hosztok számának alakulásáról (2.1. ábra).

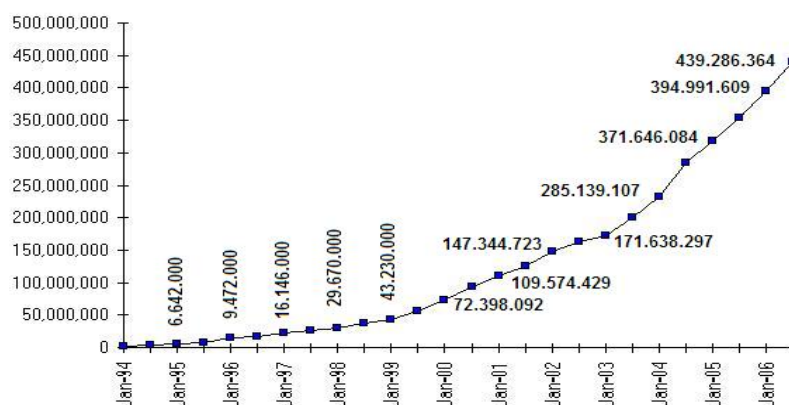
2.3.1.1. Az interneten levő információk mennyisége

Az internet méreteinek becsléséhez hasonlóan nagyon nehéz megmondani azt, hogy mennyi információ található a világhálón. Az Egyesült Államokbeli Berkeley Egyetem

⁹⁵ NAT: Network Access Translation – hálózati címfordítás.

⁹⁶ Az ISC egy nonprofit szervezet, amely az internet infrastruktúrája mellett hálózati szoftvereket és alkalmazásokat fejleszt. (<http://www.isc.org/>).

Információkezelés és Rendszerek Iskolája Kar⁹⁷ oktatói és diákjai kutatásokat⁹⁸ végeztek 1999-ben és 2002-ben, amelynek a célja – többek között – az interneten egy év alatt keletkező új információk mennyiségének megbecslése volt. Becslésük szerint egy év alatt 532.897 TB (1 terabájt = 10^{12} bájt) adat keletkezik az interneten.



2.1. ábra Az internetes hosztok számának alakulás 2006 júliusig⁹⁹

2.3.1.2. Az internetes információk összetétele

Az internetet az információszerzés szempontjából két részre osztják a szakemberek: azokat a weboldalakat, amelyeket a szokványos kereső motorok megtalálnak, a *kereshető* (*surface*), vagy *látható web*nek, míg azokat, amelyeket nem tudnak megtalálni *láthatatlan* (*invisible*), vagy *rejtett* (*hidden*) *web*nek, de újabb elnevezéssel *mély* (*deep*) *web*nek nevezik.

Médium	2002 [terabájt]
Látható (surface) Web	167
Mély (deep) Web	91.850
E-mail	440.606
Azonnali üzenetküldés	274
Összesen:	532.897

2.1. táblázat Az interneten küldött információk mennyisége, TB-ban¹⁰⁰

A 2.1. táblázat szerint az elektronikus levelezésen keresztül küldték a legtöbb információt 2002-ben. A legfontosabb azonban az, hogy a látható, vagyis a kereshető weben folyik a legkisebb forgalom, minden más internetes médium meghaladja annak forgalmát, ezért fontosak azok a törekvések, amelyek a mély web minél nagyobb részének láthatóvá tételére irányulnak.

⁹⁷ School of Information Management and System at the University of California at Berkeley

⁹⁸ A kutatás eredményei a <http://www.sims.berkeley.edu/research/projects/how-much-info-2003/index.htm> oldalon olvasható

⁹⁹ Forrás: <http://www.isc.org/index.pl?/ops/ds/>

¹⁰⁰ Forrás: How much information 2003.

A kísérlet eredményei alapján kijelenthető, hogy az internet egy olyan információforrás, amely *folyamatosan bővül*, és kb. két évente megduplázódik a rajta található adatmennyiség, amely már *több ezer TB nagyságrendű*.

2.3.2. Kereső szolgáltatások

Tim Berners-Lee¹⁰¹ 1991-ben bemutatta a világhálót (World Wide Web, web, WWW), ami a hipermédia információs hálózat üzemszerű formája. Ebben a rendszerben a *hipertext*¹⁰² jóvoltából az információforrásokat *hipertext csatolók (hiperlinkek)* formájában teljesen szabadon lehet egymással összekapcsolni. Az igazi újdonság pedig az, hogy a kapcsolódó és megjeleníthető információforrások nemcsak szövegek, hanem képek, hangok és mozgóképek is lehetnek. A rendszeren belül nem érvényesül semmiféle hierarchikus rendező elv, minden forrás egyszerre több más forrással is összefügghet és fordítva. A weben belül hamarosan kialakultak különféle keresőszolgáltatások, mint például a Mosaic, *AltaVista*¹⁰³, *HotBot*¹⁰⁴, *Lycos*¹⁰⁵ illetve a *Yahoo!*¹⁰⁶, Magyarországon a *HuDir*¹⁰⁷, illetve a *Heuréka*¹⁰⁸ és az *AltaVizsla*¹⁰⁹ (ma *ok.hu*). A web formájában végre megvalósult az általánosan hozzáférhető és az asszociatív gondolkodással összhangban álló felhasználói felület. Benne minden addigit egységesítő (pl. *FTP*¹¹⁰), feldolgozó és keresőeszközt (pl. *WAIS*¹¹¹, stb.) integráltak.

Az elsődleges adatokat tartalmazó dokumentumokat csak a részben belőlük nyert (pl. cím, alcím), részben intellektuálisan megállapított (pl. besorolási adatok, deskriptorok¹¹², jelzetek) másodlagos adatok alapján lehet kereshetően tárolni. A másodlagos adatok egy-egy dokumentumra vonatkozó összessége alkotja a dokumentumleírást vagy dokumentumrekordot. Az interneten is meg kell különböztetnünk az elsődleges dokumentumok szerepét játszó *HTML-dokumentumokat* és a rájuk vonatkozó

¹⁰¹ Tim Berners-Lee (sz. 1955. 06. 08.) 1991-ben még a CERN munkatársaként vett részt a www kifejlesztésében. 1994-ben megalapította a W3C-t, amelynek máig a vezetője.

¹⁰² Olyan speciális adatbázis, amelynek egyes részei egymással tetszőlegesen összekapcsolhatók, és az így létrejövő hiperkapcsok mentén a hagyományos lineáris megoldásoknál (pl. lapozás) gyorsabban kutathatók fel az egymással összefüggő információk.

¹⁰³ Honlapja: <http://www.altavista.com/>.

¹⁰⁴ Honlapja: <http://www.hotbot.com/>.

¹⁰⁵ Honlapja: <http://www.lycos.com/>.

¹⁰⁶ Honlapja: <http://www.yahoo.com/>.

¹⁰⁷ Honlapja: <http://www.hudir.hu/>.

¹⁰⁸ Honlapja: <http://www.heureka.hu/>.

¹⁰⁹ Honlapja: <http://www.ok.hu/index.html>.

¹¹⁰ FTP – File Transfer Protocol – Fájl Átviteli Protokoll.

¹¹¹ WAIS: Wide Area Information Servers – Nagyterjedésű információs szerverek. Információs adatbázisok az interneten. Jelentősége a WWW térhódításával alaposan lecsökkent. Angol mozaikszó.

¹¹² A deskriptorok a szemantikai információkereső nyelvek lexikai egységei (tárgyszavai), vagyis a valamely fogalom leírására, képviselésére kiválasztott szavak vagy kifejezések, szótáraik pedig tezauriszok.

másodlagos vagy *metaadatokból*¹¹³ felépülő *HTML-rekordokat*. HTML-rekordon tehát a HTML-dokumentumról készült, az oldal másodlagos adatait tartalmazó információ-telt értjük. Ez utóbbiakat az internetes keresőszolgáltatások hozzák létre annak érdekében, hogy a felhasználók keresni tudjanak.

A HTML-dokumentumokhoz való hozzáférést biztosító keresőszolgáltatások jelentős része a hirdetésekből tartja fenn magát, és egy részük ingyenes, másik részük használatáért fizetni kell (pl. *Looksmart*¹¹⁴, *Yahoo Search Marketing*¹¹⁵). A keresőszolgáltatások tehát érdekeltek abban, hogy a felhasználók mennél könnyebben és eredményesebben kereshessenek, ezért a választék növelése érdekében óriási adatbázisok alakultak ki. Így érik el, hogy sokan használják őket, ami kihat a hirdetések számára. Mindez emlékeztet a sajtó világára.

A gyűjtőkör szempontjából a szolgáltatások többsége *globális*, azaz – legalábbis elvileg – az egész háló a gyűjtőköre (pl. *Google*¹¹⁶, *AltaVista*, *Yahoo!*), kisebbik része *nemzeti (állami)*, *regionális* vagy *meghatározott nyelvre* korlátozza a gyűjtőkörét (pl. *ok*, *HuDir*). Egy részük *egyetemes*, azaz minden fajta és mindenféle tartalmú HTML-dokumentum a gyűjtőkörükbe tartozik, másik részük csak *speciális* tartalmú (pl. a *WWWomen*¹¹⁷ csak női tárgykörű), vagy speciális típusú (pl. a *MusicSearch*¹¹⁸ csak zenei) HTML-rekordokat szolgáltat. Az utóbbihoz tartoznak a szakterületi információk kapuszolgálatok¹¹⁹ is.

2.3.3. A keresőeszközök csoportosítása

Az online információkereső rendszerben a felhasználónak nem kell ismernie magát a keresőrendszert, ahogy a személyautó vezetőjének sem kell értenie a járműve szerkezetéhez. Olyan felhasználói felületet hoztak létre, amelynek szemiotikai¹²⁰ struktúrája az adott kultúra megszokott, mindennapi jelrendszerének felel meg, azaz a természetes, laikus gondolkodásnak.

Számos kereső eszköz létezik az interneten, hogy melyiket választjuk az általában attól függ, milyen jellegű anyagot keresünk. Más és más keresőt használunk akkor, ha

¹¹³ Adatra vonatkozó adat. Tipikus metaadatok az adatbázisok ábrázolási és szervezési formáit leíró adathalmazok (pl. táblaszerkezet, relációs információk, stb.).

¹¹⁴ Honlapja: <http://www.looksmart.com/>

¹¹⁵ Honlapja: <http://searchmarketing.yahoo.com/>

¹¹⁶ Honlapja: <http://www.google.hu/>

¹¹⁷ Honlapja: <http://www.wwwomen.com/>

¹¹⁸ Honlapja: <http://www.musicsearch.com/>

¹¹⁹ angolul: subject based information gateways – pl. *Goliat* (<http://www.goliat.hu/>).

¹²⁰ Szemiotika – jeltudomány (görög eredetű kifejezés).

magyarul, vagy ha angolul akarunk keresni. De speciális keresőket is igénybe vehetünk, ha például a Usenet hírcsoportokban, vagy FTP-szerverek között keresünk, vagy ha valakinek az e-mail címét akarjuk megtudni.

A legtöbb információ jelenleg a weben található meg, sőt ezen keresztül a más formában létező információk (pl. gopher¹²¹) is elérhetővé váltak. Ezért a webes kereső programoknak van a legnagyobb jelentőségük. A továbbiakban csak a HTML-dokumentumokhoz való hozzáférést biztosító keresőszolgáltatásokkal foglalkozom.

A webes keresőeszközöket a következő csoportba oszthatjuk:

- ☞ kereső motorok;
- ☞ témakatalógusok (internet-katalógusok).

Az interneten folytatott kereséshez külön szoftvereket használnak a szerver-, és külön szoftvereket a kliensgépeken, melyeket szerver-, illetve kliensoldali „navigációs eszközöknek”, egyszerűbben szerver-, illetve kliensprogramoknak is nevezünk. Meg kell különböztetni a kétfajta keresőszolgáltatáshoz szükséges szerveroldali és kliensoldali keresőeszközöket. A szolgáltatói vagy szerveroldalon található a keresőszolgáltatások és a témakatalógusok (böngésző szolgáltatások) programjai és adatbázisai, a felhasználói vagy kliensoldalon pedig visszakereső („nézegető”, viewer, browser) rendszerek.

2.3.4. A keresőmotorok

A keresőmotorok a weboldalakon található *szavakat indexelik*, vagyis *adatbázisba rendezik*. A szolgáltatások leszedői éjjel-nappali üzemben, csatolóról csatolóra haladva indexelik a HTML-dokumentumokat. Különböző algoritmusok alapján megpróbálják kitalálni a lapra legjobban jellemző szavakat és fogalmakat. Jelentős részük a címetek, alcímetek vagy a teljes szöveget indexeli, de közülük sokan a teljes szövegből csak meghatározott számú sort (pl. az első húsz sort), vagy megnézik a 10-nél többször előforduló szavakat, de vannak olyanok is, amelyek a metaadatokat (nem látható jelölések) veszik figyelembe. Léteznek szolgáltatások, amelyek eleve csak a HTML-dokumentumok metaadatait vagy annak kis részét dolgozzák föl. Az indexelést célprogramok végzik, amelyeket *robotoknak* vagy *pókoknak* (*spider*) neveznek. Az összetettebb szempontok szerinti keresést megvalósító programok az *intelligens ügynökök* (*intelligent agent*).

¹²¹ Elosztott fájlkezelő, kliens/szerver elven működő rendszer információforrások egységes, hierarchikus menükön keresztül való eléréséhez

A programok automatikusan figyelik az internetet és próbálják az új lapokat megtalálni. Sok esetben a weblap készítője kéri a keresőmotort, hogy vizsgálja át a webhelyét és építse be a lapokat az adatbázisába. Az internet-katalógusok kisebbik része kizárólag ezen az alapon működik.

Az internetnek egyik legfőbb jellemzője a *folytonos változás*, ami azt jelenti, hogy egyes lapok eltűnnek, mások keletkeznek, vagy a lapokon lévő információ lesz más. Emiatt a *keresőmotorok adatbázisát* időről időre *frissíteni kell*. Az aktualizálás gyakorisága a *hetenkénti, a félévenkénti, sőt évente egyszeri gyakoriság között mozog*. A legtöbb szolgáltatás nem közöl erről semmilyen adatot. Ugyanannak a HTML-dokumentumnak különféle változataiból közelítőleg megállapíthatók az erre vonatkozó adatok. Mennél nagyobb a szolgáltatás, annál kisebb frissítési gyakoriságra lehet számítani.

Egy keresőmotort a *pontossága (precision)*, a *találati aránya (recall)* és a *lefedettsége (coverage)* jellemzi. A pontosság azt adja meg, hogy a találat mennyire fedi a lekérdezést. A találati arány méri azt, hogy a katalogizált és a keresés szempontjából fontos dokumentumok hány százaléka van a találatok között. A lefedettség azt határozza meg, hogy a lehetséges dokumentumokból mennyi található meg a keresőmotor adatbázisában.

2.3.4.1. A dokumentumok rangsorolása

A *rangsorolás* azt jelenti, hogy keresési igény esetén milyen sorrendben jelennek meg az adatbázisban talált egyezések. Ennek azért van nagy jelentősége, mert a listázáskor előbbre álló lapnak közelebb kell állnia a keresett weblaphoz, mint a hátrább állóknak. Kereséskor az elől álló találatokat általában meg szoktuk nézni, hiszen azt feltételezzük, hogy azok a fontosabb oldalak.

A keresőmotorok között azonban jelentős különbségek vannak, mind az adatbázisra, mind pedig a lapok rangsorolására nézve. Ez az oka annak, hogy különböző keresőmotorokkal ugyanarra a keresésre különböző eredményt fogunk kapni.

A rangsorolást befolyásoló faktorok a következők:

- ☞ a kulcsszavak sorrendje;
- ☞ a kulcsszó előfordulási gyakorisága a dokumentumban;
- ☞ a kulcsszó előfordulása a címben;
- ☞ ritka kulcsszavak
- ☞ hivatkozások száma.

Ha egy kulcsszó hamarabb fordul elő, akkor az a dokumentum a rangsorban előbbre kerül. A többször szereplő kulcsszó is a rangsorban való előrejutást segíti elő. A cím kiemelt hely a keresőmotorok szempontjából, ezért ha abban szerepel a kulcsszó akkor az fontosabb dokumentumnak értékelődik. A ritkán előforduló szavak fontosabbnak minősülnek.

A rangsor alakulásába a weblap készítője is bele tud avatkozni. Az egyik ilyen módszer például a nem valós kulcsszavak feltüntetése. Másik módszer, hogy többször is beírják a fontos kulcsszavakat. Ezeket az eljárásokat azért szokták használni, hogy az általuk készített weblap minél előbb legyen egy találati listán, ami komoly reklámértékkel bír.

Manapság csak néhány kereső rendelkezik az egész webet átfogó, naprakész kereső-adatbázissal, a feladat hatalmas erőforrás-igénye miatt. Ebből következően sok keresőoldal csak licenszeli a saját adatbázissal rendelkező keresők (pl. *Google*, *Yahoo!*, *MSN*¹²², *Ask Jeeves*¹²³) szolgáltatásait. Ez némi könnyebbséget jelent a keresés megvalósítása során: például nem kell(ene) külön foglalkozni másik kereső programokkal, mint például a *OK.hu*, elég, ha tudjuk, hogy ez a kereső például a *Yahoo!* kereső adatbázisát használja.

A kereső programokat szakemberek folyamatosan összehasonlítják és erről időnként tanulmányokat tesznek közzé. Egy 2005-ös vizsgálat szerint a négy legnagyobb adatbázissal rendelkező kereső: a *Google*, *Yahoo!*, *MSN*, *Ask Jeeves*. Ezek a keresők az alábbi százalékban tartják nyilván a látható web tartományát adatbázisaikban:

- ☞ *Google* : 76,16%;
- ☞ *Yahoo!*: 69,32%;
- ☞ *MSN*: 61,90%;
- ☞ *Ask Jeeves/Teoma*: 57,62%.

A látható web részét képező weboldalaknak csak a 28,85%-a szerepel a vizsgált négy nagy kereső-adatbázis mindegyikében. Egy tanulmány 2005. január végén 11,5 milliárd oldalra becsülte a négy kereső által beindexelt, vagyis azok számára látható web méretét.

2.3.4.2. *Metakeresők*

A keresőszolgáltatások választéka ma már rendkívül nagy. A számos nemzetközileg ismert rendszeren kívül nagyon sok a *speciális gyűjtőkörű (dedikált)* rendszer, mely csak meghatározott típusú HTML–dokumentumokat dolgoz föl (pl. könyvkiadók leg-

¹²² Honlapja: <http://www.search.msn.com>.

¹²³ Honlapja: <http://www.ask.com/>.

újabb kiadványait, műszaki folyóiratokat, folyóiratcikkeket, Usenet-cikkeket, cégeket), továbbá az olyan rendszer, amely csak meghatározott államon belüli webhelyek HTML-dokumentumait indexeli (a Heuréka és az OK.hu pl. csak a magyarországiakét). A szolgáltatások teljesítménye kisebb-nagyobb mértékben különbözik egymástól, ami a felhasználót arra kényszeríti, hogy a lehető legnagyobb teljesség érdekében több indexelő keresőszolgáltatást is igénybe vegyen, ami meghosszabbítja a keresést. Ráadásul nehéz összehasonlítani az eredményt, mert a találatok külön-külön jegyzékekben jelennek meg.

Ezen hivatottak segíteni a *többszörösen, szimultán* vagy *meta-keresőszolgáltatások* (multiple/parallel/meta search engines). Velük egyszerre több indexelő keresőszolgáltatásban lehet keresni anélkül, hogy a felhasználónak az egyes szolgáltatásokkal külön foglalkoznia kellene. Ez körülbelül olyan, mintha valaki könyvet keres, mégpedig az összes magyarországi könyvtárban, és megkapja találatként, hogy az adott könyv milyen adatok kíséretében található meg az egyes könyvtárakban.

A többszörösen keresőszolgáltatás leszedője a kijelölt keresőszolgáltatásokat a keresőkérdések alapján párhuzamosan fésüli át és a találatokat közös listában jeleníti meg, ami a nagyobb választék mellett a jobb összehasonlítást is elősegíti. Az első ilyen rendszerek 1995-ben készültek. A hatékony működésüket az elosztott rendszerű szervezéssel fokozzák: a részműveletekre felbontott feladatokat egyszerre több számítógép leszedő és indexelő programjaira delegálják. Elég megadni a keresőkérdést, a többszörösen indexelő-szolgáltatás a profiljába fölvetett szolgáltatásokat végignézve kilistázza a találatokat. A jobb minőségű rendszerekben a talált információtételeknél feltüntetik, hogy melyik indexelő szolgáltatásból származnak. Így gyakori, hogy ugyanaz a tétel többször is megjelenik, és összehasonlítható, melyik rendszer szolgáltatja a legfrissebb találatokat.

Metakeresőknek létezik *kliens-*, illetve *szerveroldalon* futó típusa. A kliens oldali programokat fel kell telepíteni a saját számítógépre, ilyen például a *Copernic Agent*¹²⁴ program, ami egyszerre 12 kereső programmal hajt végre keresést, az eredmény listán feltüntetve azt a kereső szolgáltatást ahonnan az adott találat származik. A szerver oldali metakeresők használatához csatlakozni kell a hálózatra és azután lehet futtatni a kereső-programot (pl. *Metacrawler*¹²⁵).

¹²⁴ A program ingyenes változatát a <http://www.copernic.com/> oldalról lehet letölteni.

¹²⁵ Honlapja: <http://www.metacrawler.com/>.

2.3.5. Témakatalógusok

A *témakatalógusok* (browsing services) hierarchikus (ritkábban enumeratív¹²⁶) *osztályozási rendszert alkalmazó* keresőszolgáltatások, amelyek adatbázisa a túlnyomórészt intellektuálisan osztályozott HTML–dokumentumok rekordjait (másodlagos adatokból álló leírásait) tartalmazza, valamint egyéb adatbázisok információtételét. Bennük az osztályok alapján – elsősorban a katalógusban „lapozva” – végezhető a böngészés.

Az ismertebb globális rendszerek közé tartozik például a Yahoo! és az *excite*¹²⁷. A keresőszolgáltatásoknak ez a fajtája jelent meg először, valójában már a web előtt, a Gopherrel egy időben. Magyarországon 1995-től működik a *HUDIR (Hungary. Network)*, illetve a Google¹²⁸ is indított ilyen szolgáltatást.

A témakatalógusoknál *emberek végzik a weblapok katalogizálását*. Végignézik a weblapokat és különböző szempontok alapján azokat katalógusokba, könyvtárakba rendezik. Minden kategóriához kulcsszavakat rendelnek, és a katalógus tulajdonképpen ezek alapján kereshető. Az azonos kulcsszó által indexelt weblapok csoportját nevezzük kategóriának. A kategóriákban már sokkal könnyebb keresni, ha jól megfogalmazzuk a keresésünket. Ha a kategóriákban a dokumentumok száma meghalad egy bizonyos számot, akkor megosztják a kategóriát és alkategóriát hoznak létre.

A legfelső szinten néhány jól áttekinthető és főleg közismert szakterület (főosztály) jelenik meg. Az osztályozási rendszerek többnyire ismeretterületeket tartalmaznak, de vannak földrajzi, időrendi, dokumentumtípusok stb. szerinti rendszerek is.

A témakatalógusokban *intellektuálisan dolgozzák föl* a dokumentumokat, ezért nem készül keresőprogrammal automatikusan a formális adatokról dokumentum-leírás. A formális leírásokat tehát szintén manuálisan kell elkészíteni, hogy létrejöjjön a metaadatokat (szerző, cím stb.) tartalmazó teljesebb másodlagos információ-tétel, amiket az esetek jelentős részében maguk a beküldők, tehát laikusok készítenek el. Az önkéntesen beküldött tételek számos katalógusban többségben vannak, de jóformán minden kereskedelmi célú szolgáltatásban rendelkezésre állnak bejelentési űrlapok.

Különösen az igényesebb szolgáltatásokban előfordul, hogy az osztályozási rendszer valamelyik osztályába besorolt dokumentumhoz még tárgyszavakat vagy deszkriptorokat lehet rendelni. Mivel számos témakatalógusban nemcsak böngészni lehet az osztályozási rendszer hierarchikus szerkezte mentén, hanem természetes nyelv-

¹²⁶ felsoroló, csak mechanikusan rendezett.

¹²⁷ Honlapja: <http://search.excite.com/>.

¹²⁸ Honlapja: <http://directory.google.com/>.

ven is le lehet kérdezni az állományt, a tárgyszavak és deszkriptorok kereshetőbbé teszik a HTML–rekordokat.

2.3.5.1. A témakatalógusok lekérdezése

Általános jelenség, hogy a témakatalógusokban nemcsak a hierarchikus osztályozási rendszerben lapozgatva lehet böngészni, hanem megadható külön ablakban természetes nyelven a keresett szó. Ha ez megegyezik a rendszer valamelyik osztályának nevével, vagy nevének részletével, akkor a kereső rögtön az adott osztálynál találja magát.

Újabban már megfigyelhető az a tendencia, hogy a katalógusokat integrálják az indexelő szolgáltatásokba. A katalógusok adatbázisainak mérete lényegesen kisebb, mint az indexelő szolgáltatásoké. Mivel többnyire intellektuálisan osztályoznak, a teljességre eleve nem törekedhetnek. Annak érdekében, hogy még több releváns adatot szolgáltatassanak, keresőgépet is alkalmaznak.

A nagyobb témakatalógusok egyre több nemzeti/nyelvi változatot is létrehoznak. Ezek jelentős része valójában teljesen önálló. Bennük csak az adott ország, régió forrásait dolgozzák föl és azok az adott ország körülményeihez alkalmazkodó fejlesztések.

2.3.6. Blogkeresők

A *weblog*¹²⁹, web log, webnapló, vagy egyszerűen *blog* egy webes alkalmazás, amely időponttal ellátott bejegyzéseket tárol egy nyilvános weboldalon. A bejegyzések leggyakrabban, fordított időrendi sorrendben vannak. Egy ilyen weblap jellemzően bármely WWW felhasználó számára elérhető, olvasható, sőt bizonyos mértékben néha szerkeszthető is (pl. kommentezés).

A blogok témája sokrétű: lehet egy személyes napló, egyéni vagy csoportos politikai véleménynyilvánítás, üzleti jellegű kommunikáció, illetve tematikus (pl. zenei). A webnaplót vezetők köre az alkalmi szerzőktől (pl. politikusoktól) a nagy írói közösségekig, sőt vállalatokig terjed. Sok weblog teszi lehetővé a látogatóknak nyilvános megjegyzések írását, ami közösségek kialakulásához vezethet. A bloggal kapcsolatos oldalak összességét *blogtérnek* (*blogsphere*) is nevezik.

Mivel a hivatkozások nagyon fontosak a blogokban, ezért a legtöbb támogatja a régebbi bejegyzések archiválását és statikus linkek generálását¹³⁰. A legfrissebb címetek

¹²⁹ A „weblog” a „web” és a „log” szóösszetételből származik. A „blog” kifejezést a server log-gal való összekeverhetőség miatt, ennek elkerülésére kezdték alkalmazni, és hamar elterjedtté vált.

¹³⁰ Ezeket a statikus linkek permalinknek nevezik.

hiperlinkkel és összefoglalóval *RSS*, vagy *Atom XML*¹³¹-formátumban teszi elérhetővé a weblog, hogy az erre a célra fejlesztett programokkal (feed reader) olvashatóak legyenek.

Mivel a blogok írása nagymértékben elterjedt, ezért már kifejlesztettek külön a blogok keresésére alkalmas keresőket is, ami csak a blogokban végez keresést (pl. *Sphere*¹³², *BlogSearchEngine*¹³³, *miner*¹³⁴). A legtöbb blog keresőmotor rendelkezik automatikus ping szolgáltatással, aminek a segítségével, be lehet állítani melyik oldal(ak)nak „szóljon”, ha új bejegyzést adnak egy adott bloghoz. Mivel mára olyannyira elszaporodtak a blogkereső szolgáltatások, hogy egy hétköznapi blogger számára képtelenség nyomon követni, ezért olyan megoldások születtek, mint a *pingomatic*, amelynek a használatával automatikusan meg lehet pingelni az összes nagyobb blogkereső szolgáltatást, így tulajdonképpen leveszi a terhet a blogíró válláról.

2.3.7. A mély web

Van a világhálóknak olyan része, amelyet a kereső programok nem látnak ezt nevezik láthatatlan webnek. Az utóbbi időben inkább a *mély web* elnevezést kezdték használni, mert a valóságban nincs olyan információ, amit nem lehet látni, hanem csak olyan, amit nehezebben lehet megtalálni. Példaként említhető egy menetrend adatbázisa, amely a mély web része, de senki sem nevezné azt láthatatlannak, hiszen rendszeresen használják.

A mély web dokumentumait a keresőszoftverek nem tudják (technikai korlátok) vagy nem akarják (elvi döntések) adatbázisukba beemelni. Különböző becslések találhatók a mély web nagyságáról. A legtöbb becslés a látható webnek 400-550-szeresére értékeli, de van, aki ettől is nagyobbra tartja. A mély web oldalai lefedik életünk szinte valamennyi területét. Forgalmuk sokszorosa a kereshető web oldalainak.

2.3.7.1. A mély web részei

Az indexelő keresők robotjai (spiderek) a linkek mentén haladnak és érik el a weboldalakat. Ezért, ha egy weboldalra nem mutat link, az a kereső számára láthatatlan. Sok weboldal használ adatbázis technológiát. Az ilyen adatbázisokban tárolt adatok lekérdezését valamely lekérdező nyelv (pl. az SQL¹³⁵) teszi lehetővé. A lekérdezés megadásakor valamely karaktersorozatot kell begépelni vagy tulajdonságok egy sorozatát ki kell

¹³¹ eXtensible Markup Language – bővíthető jelölőnyelv.

¹³² Honlapja: <http://www.sphere.com/>.

¹³³ Honlapja: <http://www.blogsearchengine.com/>.

¹³⁴ magyar blogkereső szolgálat <http://miner.hu/>.

¹³⁵: Structured Query Language – Strukturált lekérdező nyelv.

választani, amire az indexelést végző robotok képtelenek. Ezek a valódi mély weboldalak.

Nem átlátható web (Opaque Web). Azon dokumentumok tartoznak ide, amelyek könnyen elérhetőek lennének a begyűjtő robotprogramok számára, de valamely okból mégis kimaradnak a keresőszoftverek adatbázisaiból. Ennek oka lehet a begyűjtés mélységének korlátja, az egy domainről begyűjthető dokumentumok számának korlátja, a frissítési periódus alatt megváltozott, megjelent oldalak, a dokumentumszigetek.

Magán web (Private Web). Az interneten át elérhető, de különböző megoldásokkal a nyilvánosan látható oldalak közül kizárt anyagok, az intranet hálózatokon található dokumentumok tartoznak ide.

Szabadalmazott web (Proprietary Web). Olyan adatbázisok, dokumentumok tartoznak ide, amelyeket tartalomszolgáltatók állítanak elő, és ezért regisztráció után előfizetéssel vagy külön díjért tekinthetők csak meg.

Az **igazán láthatatlan web (Truly Invisible Web).** A nem HTML formátumú dokumentumokat (audio, video, képek) nehezen vagy sehogy sem értelmezik a keresőszoftverek, ezért nem is építik be az adatbázisukba. Számos egyéb formátumra (pdf, flash, office fájlok, stb.) is igaz, hogy időigényes indexelni őket, ezért számos keresőszoftver nem foglalkozik az ilyen típusú dokumentumokkal. Elvi okból kerülnek kizárára a dinamikusan generált weboldalak, valamint a valós idejű tartalmak.

2.3.7.2. Kísérletek a mély web feltárására

A mély web részét képező weboldalak elérését segítik a rendszerezett hivatkozások gyűjteményei, mint pl. az internetes katalógusok. Mivel a mély web teszi ki az internet nagyobb részét, ezért nem lehet lemondani arról, hogy ezt is kereshetővé ne tegyék. Ennek megvalósítására már született néhány megoldás.

Metakereső megoldások. Cégek egész sora kínálja adatbázis alapú weboldalakon való keresését a mély web jelenség egy lehetséges megoldásaként (*ProFusion*¹³⁶, *The Invisible Web*¹³⁷, *Complete Planet*¹³⁸, *DirectSearch*¹³⁹).

A **web oldalak optimalizációja**¹⁴⁰. A kereső oldalakon meg nem jelenő dinamikus lapok problémáját részben az okozza, hogy bizonyos karaktereket tartalmazó *URL*¹⁴¹-

¹³⁶ Honlapja: <http://www.profusion.com>.

¹³⁷ Honlapja: <http://www.invisibleweb.com/>.

¹³⁸ Honlapja: <http://aip.completeplanet.com>.

¹³⁹ Honlapja: <http://www.directsearch.com>.

¹⁴⁰ Szoftver vagy hardver beállítások hozzáigazítása a konkrét működési környezethez és jellemzőkhöz a lehetséges legnagyobb teljesítmény elérése érdekében

¹⁴¹ URL: Uniform Resource Locator – egységes forráshatározó.

eket a kereső pókok figyelmen kívül hagynak, ezért a dinamikus oldalakat is a pókok által kedvelt statikus linkként próbálják meg feltüntetni.

Tartalomjegyzék –TOC¹⁴² - készítés. Kereső-barát linkeket készítenek a dinamikus oldalak URL-jeiből.

Könyvtár oldal (directory site) építése. Vannak eszközök, amelyekkel olyan directory site építhető, amely automatikusan szűri és kategorizálja a beérkező tartalmat.

Rejtett webet is felfedező pók készítése. Kereső oldali megoldások közül az egyik leghatékonyabb lehetne olyan kereső pók építése, amely be tudja járni a mély webet, és képes az ott talált információk kinyerésére. A *Stanford University*¹⁴³ két kutatója által épített, *HiWE (Hidden Web Exposer)* névre hallgató pók részben képes erre.

2.4. AZ INTERNETES KERESÉS FEJLESZTÉSE

A számítógépes hálózat építésének kezdetén egy (ARPANET), majd több számítógépek közötti hálózat is kialakult, de ezek mindegyike saját kódolással és jelentéstartalommal bírt, ezért a különböző hálózatok között nehéz volt az átjárás. A TCP/IP protokoll elterjedése lehetővé tette, hogy a világ bármely részén lévő számítógép össze tudjon kapcsolódni bármelyik másik számítógéppel annak *szemantikai*¹⁴⁴ szintjén. Ezáltal kialakult a számítógépes együttműködésnek egy olyan szemantikai szintje, ami *lehetővé tette a számítógépek közötti adatsorozatok és állományok cseréjét*, de az egyes gépek még továbbra is elkülönültek maradtak, így csak egyedi formátumú adatok és állományok cseréjét lehetett megvalósítani.

A következő nagy fejlődési lépcső a *hiperszöveg protokollok* kialakítása és egységesítése volt. Az új protokolloknak köszönhetően az interneten összekötött számítógépeken tárolt jelentős mennyiségű állományok között már jobban el lehetett igazodni, mert azokban egymásra lehetett hivatkozni. A *http*¹⁴⁵ protokoll és a *html*¹⁴⁶ leírónyelv gyorsan terjedt. Ezek a protokollok azonban nem teszik lehetővé az automatizációt, mert még mindig egyedi, és elsősorban emberi fogyasztásra szánt megjelenítésre készültek.

¹⁴² TOC - Table Of Contents: tartalomjegyzék

¹⁴³ Honlapja: <http://www.stanford.edu/>.

¹⁴⁴ A szemantika a számítógépes jelölési rendszerek és az azok által jelölt fogalmak, tárgyak közötti kapcsolatot jelenti. Eredeti jelentése – jelentéstan.

¹⁴⁵ Hypertext Transfer Protocol (HTTP) Hiperszöveg Átviteli Protokoll.

¹⁴⁶ HyperText Markup Language - hypertext jelölő-nyelv.

2.4.1. A szemantikus web

A világháló legújabb fejlődési lépcsője a *szemantikus*¹⁴⁷ *web*, amelynek az alapkoncepciójára a World Wide Web atyja, Tim Berners-Lee tett javaslatot 1999-ben. Ennek legfőbb építőköveit a *World Wide Web Consortium (W3C)* dolgozta ki és szabványosította. A szemantikus web célja, hogy a számítógépeket, az egyszerű gépi intelligenciákat hívja segítségül az internetes keresés jelenlegi problémáinak a megoldására. Ehhez azonban az információkat számítógépek által értelmezhető formára kell hozni.

A szemantikus webben olyan metaadatokkal látják el az internetre kikerülő információkat, ami olvashatóvá teszi őket más gépek számára, ami a különböző interakciók automatizálását teszi lehetővé. Így megvalósulhat az, hogy egy keresőben rákeresünk például a „sport” szóra, akkor az általunk bejárt online útvonalak nyomon követésével (URL tracking), a kereső találati listája már nemcsak azt veszi figyelembe, hogy jól szerkesztett-e a honlap, sűrűn és magas relevancia-értékkel szerepel-e rajta a kulcsszó, sok jó minőségű link mutat-e rá, hanem az általunk kedvelt témák alapján módosítja a találati toplistát.

A fenti keresési technika megvalósításához a *szemantikus web alapeszköze az RDF*¹⁴⁸ nyelv, amellyel egységesen reprezentálhatók olyan állítások, mint például:

- ☞ ezt a dokumentumot Példa Géza készítette;
- ☞ Példa Géza a ZMNE munkatársa;
- ☞ A ZMNE egy egyetem.

Ezeknek az állításoknak egy része a metaadatok körébe tartozik, míg más állítások egyfajta globális tudásbázis apró részleteiként foghatók fel. Ezeket az állításokat összekapcsolva azután különféle logikai következtetéseket lehet elvégezteni a számítógépekkel, így segítve a felhasználókat. Egyes kérdésekre automatikusan választ kaphatunk, újfajta kereséseket végezhetünk, vagy hasznosabb eredménylistát kaphatunk a kereséseinkről. A szemantikus web eszközkészlete az emberi kommunikációban is alkalmazható: állításainkat, megjegyzéseinket ezzel a módszerrel mások által írt tetszőleges weblaphoz is hozzárendelhetjük.

A szemantikus web másik fontos eleme a *tudásreprezentáció*, vagy más néven az *ontológia*, ami hasonló szerepet tölt be, mint az XML az adatreprezentációban, csak az ontológia az adatok mellé még az adatok közötti összefüggések törvényszerűségeit is

¹⁴⁷ Szemantikus: a számítógépes jelölési rendszerek és az azok által jelölt fogalmak, tárgyak közötti kapcsolatot leíró.

¹⁴⁸ RDF: Resource Description Framework - Forrásleíró keretrendszer

leírja. A W3C által 2004-ben elfogadott *OWL*¹⁴⁹ nyelv ajánlása válik a szemantikus web legfontosabb leíró nyelvéné. Az OWL-t olyan alkalmazások számára fejlesztették ki, amelyek célja már nem az információ ember számára történő pusztán bemutatása, hanem az információ tartalmának gépi feldolgozása. Az OWL nyelv nagyobb mértékben segíti a web-tartalom gépi értelmezését, mert bővített szókészletet, illetve formális szemantikát kínál.

2.4.2. A szemantikus web modellje

A szemantikus web összképét az egymásra épülő több szintes modellel lehet leírni. Ennek szintjei a következők:

- ☞ *Az első szinten a Unicode és URI*¹⁵⁰ *ajánlások biztosítják az egységes nemzetközi karakterformátumot és az egységes címezhetőséget. A szemantikus webben belül az URI-k használata jóval túlmutat a megszokott weblapcímeiken. Egy személyt is azonosíthat mondjuk az e-mail címét tartalmazó URI, vagy egy intézetet a weblapjának a címe. Tulajdonságokat, állításfajtaikat is URI-k azonosítanak, így egyes állítások olyan módon fogalmazhatók meg, hogy az más állításokkal az URI-k mentén összekapcsolhatóvá válnak.*
- ☞ *Az XML réteg biztosítja, hogy a szemantikus web definíciókat egyéb XML alapú szabványokkal integrálhassuk. Az XML a névtér- és séma-definíciókkal együtt lehetővé teszi, hogy egy XML dokumentum a létrehozásakor alkalmazott szintaktikai és szemantikai szabályokra hivatkozzon. Ezáltal egy önleíró, érvényesíthető dokumentumformátum jön létre.*
- ☞ *Az RDF és RDFS*¹⁵¹ *réteg biztosítja, hogy állításokat tehesünk az objektumokról URI-k segítségével, és olyan szótárakat készítsünk, amelyekre ezen URI-k hivatkozhatnak. Az első RDF és RDFS specifikáció 1999-ben készült el, de végleges W3C-ajánlássá csak 2004 februárjában vált. Az RDF számos alkalmazásban jelen van: könyvtári katalógusok, világméretű címtárak, hírek, szoftverek, tartalmak gyűjteményei, személyes zene-, fénykép- és eseménygyűjtemények. Ezek mindegyike XML-t használ a köztes adatcsere leírására. Az RDF specifikáció egy erős keretrendszert biztosít a weben keresztüli tudás cseréjére;*

¹⁴⁹ OWL: Web Ontology Language – Web Ontológia Nyelv

¹⁵⁰ Uniform Resource Identifier – Egységes Erőforrásazonosító. Minden erőforrás a WEB-en (HTML dokumentum, kép, videoklip, program, stb.) egy címmel rendelkezik, amelynek formátumát az URI határozza meg. Az URI-t használhatjuk: link létrehozásra, képek, objektumok vagy appletek beágyazására egy dokumentumba, metaadatokra való hivatkozásra, stb.

¹⁵¹ RDF Schema: RDF séma.

- ☞ Az *ontológiai réteg* lehetőséget ad bonyolultabb szótárak felépítésére. Itt már a különböző fogalmak közötti összefüggések is megadhatók, hasonlóan egy tezauruszhoz. Az ontológia definiálja a terminológiát, amellyel egy adott tudás-terület leírható és kifejezhető. Az ontológiákat az emberek, adatbázisok és alkalmazások használják, akik egy bizonyos témára vonatkozó vagy adott tárgykörben levő információkat szeretnék megosztani egymással. Az ontológiák magukba foglalják az adott terület alapfogalmainak számítógép által használható definícióit és a közöttük lévő kapcsolatot is. Kódolják egy terület tudását és a területet áthidaló tudást is. Így teszik újrahasználhatóvá a tudást. Ehhez fejlesztették ki az OWL nyelvet, amely a szemantikus web legfontosabb leíró nyelve, amellyel strukturált, webalapú ontológiák definiálhatóak, melyek az adatok gazdagabb integrációját és interoperabilitását biztosítják a meghatározásokat készítő közösségek között. Az OWL az RDF-re és az RDF-sémára épül és további szókincsgyűjteményekkel egészül ki, hogy a tulajdonságokat és osztályokat leírhasssa;
- ☞ A *digitális aláírás réteg*, amely a dokumentumok hitelességét igazolja.

2.4.2.1. Az egyes szintek kapcsolatai egymással

A szemantikus web a mai világháló továbbfejlesztésének tekinthető. Fokozatosan a számítógép által is megérthető leírásokat kapcsolnak azokhoz az adatokhoz és dokumentumokhoz, amelyek korábban már a weben voltak és így alakul ki a szemantikus web. Így az XML, az RDF és az OWL segítségével a web-dokumentumok és adatok olyan globális infrastruktúrája alakul ki, ahol könnyebb és megbízhatóbb az információk keresése és felhasználása.

2.5. ÖSSZEGZÉS ÉS KÖVETKEZTETÉSEK

1. Az internet egy olyan folyamatosan és dinamikusán bővülő információforrás, amelynek a mérete már több tízezer Terabájt nagyságrendű.

Az interneten már több az információ, mint a világ legnagyobb könyvtárában az Egyesült Államok Kongresszusi Könyvtárában. Az interneten lévő információk növekedési ütemének arányait nézve körülbelül két évente megduplázódik azok mennyisége. Ez a tendencia várhatóan a jövőben is folytatódik.

2. Az internet jelenlegi kiépítettsége, az általánosan nyújtott sávszélessége és átviteli sebessége csak nehezen biztosít megfelelő feltételeket olyan új szolgáltatások bevezetéséhez, amelyek nagy, illetve állandó sávszélességet igényelnek (pl. IPTV, VoIP). A sávszélesség és átviteli sebesség problémák megoldására kezdték el 1996-ban az Új Generációs Internet Kezdeményezést (NGII), aminek az eredményeként napjainkra már 10 Gbit/s átviteli sebességet is képesek elérni a kutatói hálózatokban.

A kutatói hálózatokban elért átviteli sebesség még jó ideig nem realizálható az általánosan használt hálózatokban. Jelenleg az a cél, hogy 100 Mbit/s átviteli sebességet érjenek el 2010-re. Az NGII kereteiben folyó fejlesztések lehetővé teszik a megbízható összeköttetést (QoS), a nagy átviteli sebességet (>Gbit/s) és nagy sávszélességet megkövetelő szolgáltatások jövőbeni bevezetését, illetve továbbiak kifejlesztését.

A fejlesztések további eredményeként várható, hogy egybe integrálódnak az eddig külön álló információsugárzó, kommunikációs, tájékoztató, szórakoztató, oktató, kereskedelmi és kulturális szolgáltatások. Így az internet része lesz, illetve a hálózaton keresztül is elérhetővé válik a rádió- és televízió műsorsugárzás, a telefon szolgáltatás (vezetékes és mobil egyaránt), újságok, könyvek, áruvásárlás, filmnézés, oktató anyagok, stb., vagyis az élet minden területét átfogó dolgok. Ezek mellett mind a rejtett technikai eszközökkel végzett, mind a nyílt forrású információszerzés leggyorsabb, legolcsóbb, leghatékonyabb és legnagyobb terepévé válik. Ez indokolja, hogy már most elkezdjük minél több tapasztalatot szerezni annak minél hatékonyabb felderítésében, kiaknázásában, felhasználásában, a szükséges források, adatok, tények minél gyorsabb és biztosabb megtalálásában.

3. A jelenleg használt internet protokoll (IPv4) már sok esetben korlátozza az internet szolgáltatásainak további fejlődést. Nem biztosít elég IP címet és nem támogat sok olyan szolgáltatást, amely napjainkban már mindennapos igényként merül fel. Az IPv4 problémáinak kiküszöbölésére kezdték el kifejleszteni a 90-es évek elején az IPv6 (IPnG) protokollt.

Az új protokollra való áttérés folyamatosan megy végbe, 2010-ig nagy valószínűséggel mindenhol befejeződik ez a folyamat. Az IPv6-ra való áttérés után, egyebek mellett, már lehetővé válik a technikai eszközök hálózaton keresztül történő ellenőrzése és irányítása. A katonai harceszközök (robotok), berendezések, valamint különböző gépek, eszközök és berendezések a távolból is irányíthatóak lesznek.

4. Mivel az interneten elérhető adatokat senki sem rendszerezi, ezért nagyon nehéz az információk megtalálása. Ebben nyújtanak segítséget a különböző internetes keresési technikák. A kereső eszközöknek két típusa van: a kereső motorok és a témakatalógusok. Az előbbiek automatikusan kutatják és indexelik az internetet, az utóbbiak intellektuális módon osztályozzák a dokumentumokat.

A témakatalógusok szerepe fokozatosan csökkeni fog, mert nehéz az egyre több anyag besorolását elvégezni. A keresés végzésének leghatékonyabb módja a metakeresők használata, mert azok egyszerre több kereső program adatbázisában is elvégzik a kutatást, így növelik a keresés gyorsaságát és hatékonyságát.

5. A kereső eszközök találatait – üzleti és egyéb megfontolásból – olyan körülmények befolyásolják, amelyek nem feltétlenül segítik a felhasználót abban, hogy azt találja meg, amit igazán keres.

A szemantikus webnek nevezett fejlesztések céljaként azt tűzték ki, hogy az internetes kereséseket minél inkább a felhasználók szokásai, és ne más körülmények befolyásolják. A feltett anyagok metaadatokkal való ellátása lehetővé teszi azt, hogy a számítógépek képesek legyenek önállóan összefüggéseket találni az információk között, ezzel segítve a hatékonyabb keresést. A különböző leíró nyelvek és protokollok segítségével a dokumentumok és adatok olyan globális infrastruktúrája alakul ki, ahol könnyebb és megbízhatóbb lesz az információk keresése és felhasználása.

3. AZ INTERNET ALAPÚ INFORMÁCIÓSZERZÉS RENDSZERTECHNIKAI MODELLJE

Az internettel együtt egy új médium, az ún. *kiber-tér* (*cyberspace*¹⁵²) is megszületett. Ez az informatikai világban létező virtuális tér – hasonlóan a valóságos térhez – a háború, az információs háború színtere is egyben. Ezen a harctéren ugyanúgy közdelem folyik a szemben álló felek között, csak más, az információs műveletek eszközeivel. A cél itt is az ellenség megsemmisítése, lefogása, korlátozása, megtévesztése, felderítése, lehallgatása, figyelése a számítógépes hálózaton belül, annak eszközkészletének felhasználásával. A „megsemmisítés” nem emberek, hanem számítógépek ellen irányul, pontosabban az azon tárolt adatok törlésére különböző rosszindulatú programokkal (vírus, trójai programok, férgek, stb.), de speciális parancsokkal akár a merevlemez fizikai tönkretételére is. A lefogást például *DoS*¹⁵³ támadásokkal, a megtévesztést hamis információk közzétételével vagy speciális programokkal, a felderítést, lehallgatást, figyelést kereső és letöltő programok használatával lehet megvalósítani.

3.1. AZ INTERNETES INFORMÁCIÓSZERZÉS JELENTŐSÉGE, SZEREPE

Az internetes információszerezés terepe a kiber-tér, eszköze a számítógép, illetve az azon futó speciális szoftverek. Végrehajtói olyan különlegesen felkészült szakemberek, akik magas szintű nyelvismerettel rendelkeznek, jártasak a számítástechnikában, az adatok kezelésében, értékelésében, az azokból való hasznos információ kinyerésében. Véleményem szerint ez a tevékenység a felderítésen belül egy teljesen új szakterületnek tekinthető, amit *internet alapú nyílt forrású felderítésnek* (INFF) neveztem el (angolul: *Internet Open Source Intelligence – I-OSINT*). Az internet jelenlegi fejlődési ütemét figyelve, várhatóan – a nem is távoli jövőben – az ilyen forrásokból származó információk teszik ki majd a megszerzett információk zömét.

Az információs társadalom kibontakozásával párhuzamosan globális méretekben terjednek az információkultúra digitális technikai vívmányai. E változások lehetőséget kínálnak az információszerezés területén arra, hogy a döntést előkészítők és a döntésho-

¹⁵² William Gibson *Neuromancer* című könyvében használta először ezt a kifejezést a világméretű számítógépes hálózatok összességére, azóta az amerikai felderítésben is elfogadott fogalom lett.

¹⁵³ DoS támadás – DoS attack Denial-of-Service attack – szolgáltatás-leállítási támadás, amikor a (többnyire) webes kiszolgáló szerveret túlterheléssel kényszerítik a szolgáltatás szüneteltetésére.

zók minél szélesebb körű, elfogadható pontosságú, szelektált és közel valószerű, nélkülözhetetlen téma-környezeti és háttér információk birtokába jussanak. Ezekkel képesek leszűkíteni azt a kritikus időt, amely a döntéshez szükséges információk megszerzésétől (beszerzésétől) azok érdemi felhasználásáig tart. Az idővel való helyes gazdálkodás pedig, jól tudjuk, jelentős versenyelőnyt biztosít a piaci szereplők számára. A fejlett kapitalista világban minden társadalmi szektort „piaci szereplőnek” lehet tekinteni, hiszen csaknem minden fejlesztési és fenntartási feladat pályázati rendszerben történik, ahol az időnek és a pontos információnak domináns szerepe van.

Az INFF rendszer funkcióját tekintve a védelemben elsősorban a nemzetbiztonsági és a hadászati vezetés terén képes érdemben javítani az információ-ellátást. Miközben alkalmas lehet arra is, hogy támogassa a határon-kívüli missziós feladatokat ellátó kontingenseink műveleti információ (mission information) igényeit is, ugyanakkor különleges szerepe lehet a médiaipar területén jelentkező általános média tevékenységek, illetve az információs műveletekhez kapcsolódó, különleges média-hadviselési műveletek megszervezéséhez és vezetéséhez szükséges környezeti és háttér információk megszerzése területén.

Tekintettel arra, hogy az NFF lehetővé teszi a társadalomban keletkező információk 80-90%-ának lefedését, nem lehet közömbös számunkra, hogy ennek az óriási információhalmaznak a feldolgozását kézi, részben gépi, vagy automatizált rendszerekkel végezzük. A fenti felismerések vezettek arra, hogy tudományos igényű kutatások keretében, részleteiben megvizsgáljam a számítógépes hálózatechnikára alapozott nyílt információszerezés technikai lehetőségeit, és kidolgozzak egy olyan rendszert, amely megfelel a jelzett kihívásnak.

Az internetet, már mindenki, nap mint nap használja – „kézi üzemmódban” – nyílt információszerezésre, de ez a módszer nem terjed ki minden internetes forrásra, ennél fogva eredménye hiányos és hézagos, korántsem teljes, kevésbé pontos. Emellett az egyéni „internet-bányászat” vagy hobbi szintű „internet halászat” rengeteg időt vesz igénybe és a keresések eredményeként megjelenített weblapok több száz, vagy több ezer találati helyéből kiválasztani az optimálisat egyre megoldhatatlanabb feladatnak látszik gépi segítség nélkül. Az általam tervezett rendszer mindezen hiányosságokat elkerüli, mivel nagy részében automatizált, ezáltal a feldolgozó munkát meggyorsítja és megkönnyíti.

A társadalom minden szegmensét már átfogja az internet világháló. Mivel igen szerteágazó, sok szakterületet felölelő problémáról van szó, kutatásaimat leszűkítettem a

nemzetbiztonsági és védelmi ágazatok internet alapú információigényeinek vizsgálatára. Úgy találtam, hogy lehetőség van arra, hogy ezen ágazatok döntéshozóinak, azokat támogató elemzők, szakma-politikusok és politológusok, továbbá a döntés-előkészítő szakmai kidolgozók számára megoldható az elektronikus, számítógép vezérelt, többfelhasználós, korszerű információszerező és szolgáltató rendszer kialakítása. Ilyennek terveztem saját rendszeremet, amely képességében szinkronban lehet a hasonló feladatú külföldi nagyrendszerekkel.

Áttanulmányozva az egyes országokban kialakított rendszereket, az internet hatékony figyelésének a legjobb modellje a holland rendszerhez hasonló lehet, ahol a keresést centralizáltan, a tájékoztatást pedig decentralizáltan oldották meg. Ezzel a módszerrel csökkenthető az új források keresésére fordítandó idő, mert azt speciális kereső eszközök, szoftverek és speciális ismeretekkel rendelkező szakemberek végzik, akik az általuk összegyűjtött adatbázisra támaszkodhatnak a munkájuk során. Az ismert források segítségével már szétsztható a speciális területek figyelése.

3.2. AZ INTERNET CSATLAKOZÁS

Az internet megfigyeléséhez természetesen szükség van valamilyen csatlakozási lehetőségre. Ezek a csatlakozási lehetőségek megszorodtak az utóbbi időben. A kilencvenes évek elején szinte csak a modemes, *betárcsázós kapcsolat (dial-up)* révén lehetett a hálóra kapcsolódni, de az a maga max. 56 kbit/s-os átviteli sebességével, már alkalmatlan a mai feladatok ellátására. Napjainkban gyorsan terjednek a szélessávú hozzáférési lehetőségek, az ADSL technológia, a kábelmodemes kapcsolat és a vezeték nélküli technikák. A leggyorsabb átviteli sebesség 20 Mbit/s is lehet, és a közeljövőre (2010-ig) ígérik a 100 Mbit/s-os átviteli sebesség elérését.

A fontosabb szélessávú csatlakozási lehetőségek:

- ☞ kábelmodemes kapcsolat (letöltés sebessége maximum 20.480 kbit/s, a feltöltés pedig 1.024 kbit/s lehet);
- ☞ vezeték nélküli kapcsolat (*IrDA*¹⁵⁴, *Bluetooth*¹⁵⁵, *WLAN*¹⁵⁶);
- ☞ ADSL (letöltés sebessége maximum 18.432 kbit/s, a feltöltés pedig 1.024 kbit/s lehet).

¹⁵⁴ Infrared Data Association – Infravörös adatátviteli szabvány

¹⁵⁵ Bluetooth: kis hatótávolságú kommunikációra alkalmas szabványosított eszköz és protokollszöveg, amelynek segítségével jellemzően hordozható kommunikációs eszközök kapcsolhatóak össze.

¹⁵⁶ Wireless LAN: vezeték nélküli helyi hálózat, közismerten WiFi - Wireless Fidelity

3.2.1. A csatlakozás módjának megválasztása

A csatlakozás megválasztásánál három dolog fontos. Az egyik, hogy a tervezett telepítési helyen elérhető legyen a szolgáltatás. A második szempont, hogy lehetőségekhez mérten a legnagyobb sáv szélességet biztosítsa. A harmadik fontos feltétel, hogy az internet szolgáltató (*ISP*¹⁵⁷) rendelkezzen elegendően nagy sáv szélességgel a nemzetközi szolgáltató, illetve a hazai *BIX*¹⁵⁸ felé. A három feltétel teljesülése esetén szűk keresztmetszet már csak a letöltendő webszervernél lehet.

Habár a rendszer feladata elsősorban weboldalak megtalálása és az ott talált információk lementése lesz, ami elvileg nem igényelne nagy sáv szélességet, de sok – akár több száz vagy ezer – oldal folyamatos, párhuzamos figyelése, monitorozása már megköveteli a nagy átviteli kapacitást. Így a rendszer teljes felfutása után lehet, hogy szükség lesz akár több kapcsolat előfizetésére, vagy a szolgáltatóval való optikai összeköttetés kiépítésére (garantált nemzetközi sáv szélesség biztosításával együtt) is az igények kielégítéséhez. A kezdeti időszakban az internetre való csatlakozásnak az egyik kézenfekvő módja az ADSL kapcsolat lehet, mivel csak kevés olyan terület van, ahol ezt a szolgáltatást nem lehet elérni.

3.3. AZ INFF MODELL RENDSZERTECHNIKAI FELÉPÍTÉSE

Az általam javasolt rendszermodell elvi rendszertechnikai kialakítását mutatja a 3.1. ábra. Az ábrán jól megkülönböztethetők a topológiailag elkülönülő alhálózatok (subnet), illetve folyamatábraszerűen követhető az adatok útja az internetes bemenettől a végfelhasználóig, az egyes rendszertechnikailag elkülönülő elemeken keresztül.

A rendszer három alhálózatból épül fel:

- ☞ *adatszerző alhálózat*: ezen belül valósul meg a források keresése és kiértékelése, az adatok automatikus letöltése, majd kiértékelés utáni adatbázisba való mentése;
- ☞ *adatfeldolgozó alhálózat*: ezen belül történik a lementett anyagokból – az információk kinyerése után – a felhasználók igényeinek megfelelő jelentések elkészítése és adatbázisba történő letárolása;
- ☞ *végfelhasználói alhálózat*: a felhasználók ezen keresztül kapják meg vagy érik el az elkészült jelentéseket online módon.

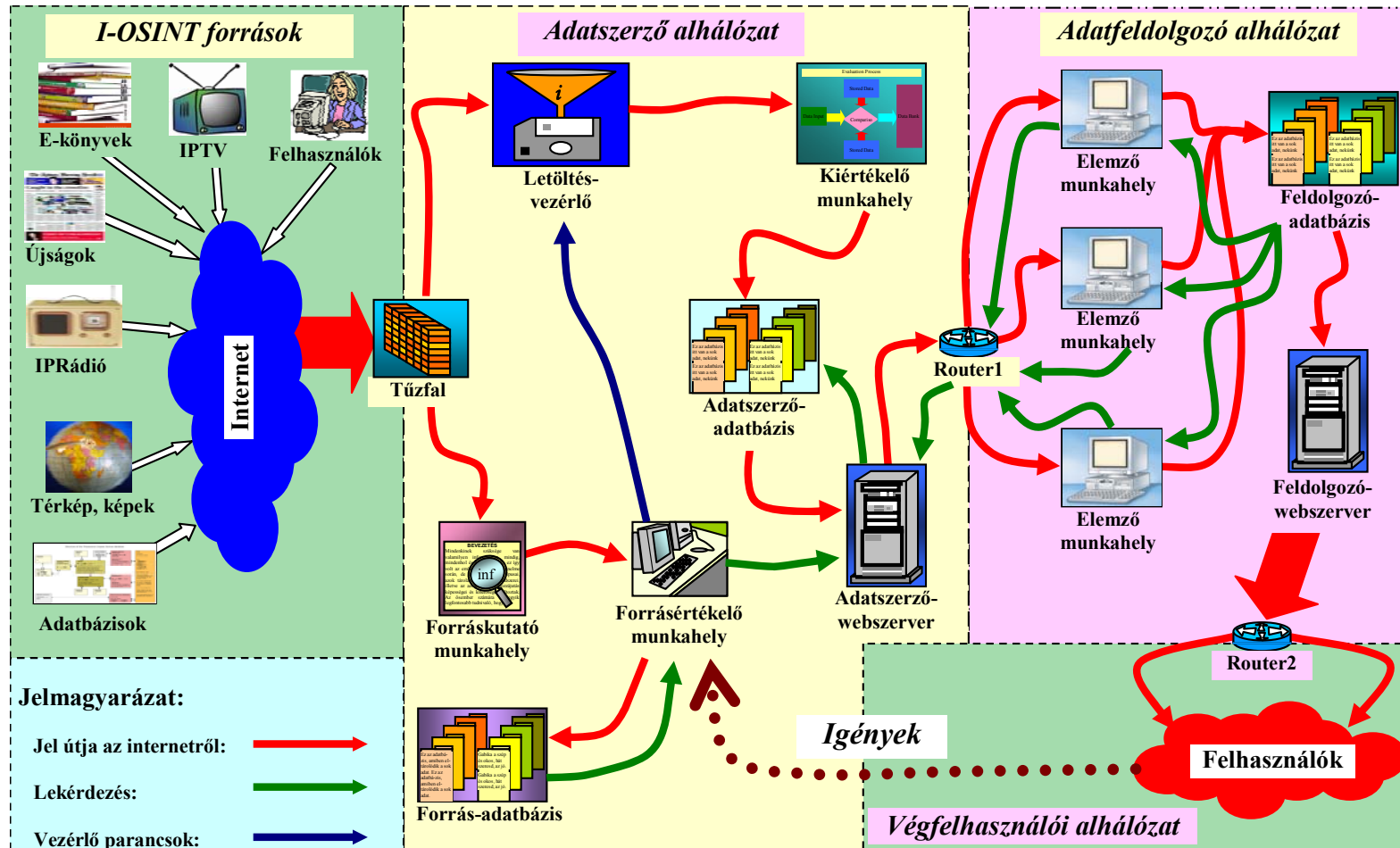
¹⁵⁷ ISP: Internet Service Provider – Internet csatlakozást biztosító cég vagy egyéb intézmény.

¹⁵⁸ BIX: Budapest Internet Exchange – Budapesti adatkicserélő központ egy kapcsolati lehetőség a magyarországi érdekeltségekkel rendelkező internet szolgáltatók részére.

Az alhálózatokra való felosztást adatforgalmi, megbízhatósági és biztonsági megfontolások indokolják. A szétválasztás eredményeként különválik az adatszerzés, az adatfeldolgozás és a végfelhasználók adatforgalma. Mindenki csak abban az alhálózatban férhet hozzá adatokhoz és szolgáltatásokhoz, amelyekben jogosultsággal rendelkezik, illetve az egyik alhálózatban jelentkező bármilyen probléma esetén a másik alhálózat zavartalanul működhet tovább.

Az egyes elemek feladatának rövid leírása:

- ☞ a *tűzfal* védi a belső hálózatot a külső támadásokkal szemben;
- ☞ az adatszerző alhálózaton belül:
 - a *forráskutató munkahelyen* folyik az új források keresése;
 - a *forrásértékelő munkahelyen* kerülnek kiértékelésre a talált források (új, elsődleges vagy másodlagos forrás);
 - a *forrás-adatbázisban* tárolódnak a megtalált források olyan adatai, aminek a segítségével el lehet dönteni, hogy egy megtalált forrás új-e, vagy már ismert;
 - a *letöltés-vezérlő* végzi a kiválasztott források által szolgáltatott anyagok automatikus letöltését;
 - a *kiértékelő munkahelyen* összehasonlításra kerülnek az újonnan letöltött anyagok a már letöltöttekkel, hogy van-e bennük új információ;
 - az *adatszerző-adatbázisban* tárolódnak az elemző munkahelyek számára letöltött anyagok;
 - az *adatszerző-webszerver* szolgáltatásain keresztül érhetik el az elemzők az adatszerző adatbázisban tárolt anyagokat;
- ☞ az adatfeldolgozó alhálózaton belül:
 - az *elemző munkahelyeken* folyik az adatszerző-adatbázisban tárolt anyagokból az információk kinyerése és a jelentések elkészítése;
 - a *feldolgozó-adatbázisba* helyezik az elemzők az elkészült jelentéseket – továbbítás és archiválás céljából –, illetve itt tárolódnak az elemzések elvégzéséhez nélkülözhetetlen alapinformációk;
 - a *feldolgozó-webszerver* teszi lehetővé az adatbázisban tárolt jelentések online elérését a felhasználók részéről.



3.1. ábra Az INFF modell elvi felépítése

3.4. A RENDSZER BIZTONSÁGI KÉRDÉSEI

Mivel az internethez való csatlakozással nem csak információhoz lehet jutni, hanem azon keresztül támadhatóvá is válik a belső hálózat, ezért szükség van biztonsági elemek beépítésére és megfelelő intézkedések megtételére a rendszer védelme érdekében. Ennek elmei a következők:

- ☞ tűzfal;
- ☞ behatolás elleni védelem;
- ☞ vírusvédelem;
- ☞ anonimizálás.

3.4.1. Tűzfal

Tűzfalrendszernek – röviden tűzfal (firewall) – nevezzük azt a *hardver* és *szoftver* együttest, amely megvédi adott szervezet belső (megbízhatónak tekintett) hálózatát a külső (nem megbízható) hálózatról érkező nem kívánatos, illetve jogosulatlan hozzáférésekkel szemben, valamint lehetővé teszi a szervezet számára a külső hálózat szolgáltatásainak kontrollált elérését. A tűzfal egy olyan hálózati eszköz, amely a rajta átmenő forgalmat megszűri és hitelesíti. Minden adatforgalom keresztülfolyik ezeken a gépeken. Működési elvük alapján három különböző csoportba sorolhatók:

- ☞ csomagszűrő tűzfal;
- ☞ proxy alapú tűzfal;
- ☞ hibrid tűzfal.

A *csomagszűrő tűzfal* a legegyszerűbb tűzfal típus, amely a TCP/IP-nek az alapegységein, vagyis a hálózati csomagokon dolgozik. Lényege, hogy minden egyes rajta áthaladó csomagot megvizsgál, és eldönti, hogy továbbítható-e. A vizsgálatot a csomagokban található információk alapján végzi (pl. cél, kiindulás cím, portszámok, stb). Általában az ilyen típusú tűzfalak önmagukban nem nyújtanak megfelelő védelmet. Gyorsak, viszonylag egyszerű felépítésűek, ám finomabb szűrést nem tesznek lehetővé.

A *proxy alapú tűzfalak* már nem a hálózati csomagok szintjén dolgoznak, hanem látják a teljes adatfolyamot, és a tartalmától függően végzik a tevékenységüket. Sokkal kifinomultabb szabályozást tesznek lehetővé, mint a csomagszűrők. Ez a tűzfaltípus úgy működik, hogy a beérkezett adatfolyamot összeilleszti, azt a lehető legmagasabb hálózati rétegben megvizsgálja, majd ha nem talál semmi gyanúsat, továbbítja.

A *hibrid tűzfalak* a csomagszűrők és a proxyk ötvözéséből születtek meg. Ez az a tűzfaltípus, amelyik a vizsgálandó csomagok által tartalmazott információon túl, képes figyelembe venni az eddig érkezett csomagokban lévő információt is. Ezeknek a tűzfalaknak nagy előnyük a sebességük. Biztonsági szempontból azonban hátrányuk az, hogy a csomagszintű manipulációknak nem képesek jól ellenállni.

Hardveres tűzfal megoldások is léteznek, amelyek előnye, hogy egyszerű őket konfigurálni, megbízhatóak, kicsi az energia felvételük, és számos, rugalmasan beállítható szolgáltatást nyújtanak.

Ideális tűzfalrendszer kialakításakor a következő építő elemeket célszerű felhasználni:

- ☞ *szűrt hozzáférésű bástya*: a külvilágból a védett hálózat csak egyetlen gépen keresztül érhető el. Ez a bástya. A többi védett gép csak ennek közvetítésével érhető el. Kifelé a védelem teljesen átlátszó, ezt egy megfelelő csomagszűrő biztosítja.
- ☞ *szűrt hozzáférésű hálózat*: a külvilág egy szabad elérésű hálózatot lát, amire a (nyilvános) szolgáltatásokat nyújtó gépek csatlakoznak. A védett hálózatot egy tűzfalgép kapcsolja erre a hálózatra.

A tűzfalrendszer, az eredeti elmélet szerint, három gépből tevődik össze:

- ☞ csomagszűrő (első fal);
- ☞ alkalmazásszűrő (bástya vagy második fal);
- ☞ csomagszűrő (hátsó vagy harmadik fal);

Az *első fal* feladata a bástya védelme, az erős sánccfal szerepét tölti be. A bástyán futó *proxy tűzfal* nem az operációs rendszer *kerneljében*¹⁵⁹ fut, ezért az operációs rendszer hibáiból adódó alacsonyszintű támadásokat ez a gép védi ki.

A *bástya* a forgalom tartalmát vizsgálja.

A bástya elestekor lép be a képbe a *hátsófal*: a támadó akkor se kaphat hozzáférést a belső hálózat forgalmához, ha a bástyán átjutott. Az első fal és bástya futhat egy számítógépen, de a hátsó fal – funkciója miatt – külön számítógépen kell, hogy működjön.

3.4.2. Behatolás elleni védelem

A tűzfal mellett olyan eszközre is szükség van, amely gyorsan, és a pillanatnyi körülményektől függően tud alkalmazkodni a védelmi igényekhez, akár egy intelligens berendezés formájában, mely méltó ellenfele a felkészült támadóknak. A *behatolás-*

¹⁵⁹ Kernel – rendszermag: egy számítógép operációs rendszerének a lelke, közvetít a programok és a hardver között, valamint menedzseli a gép erőforrásai kihasználását.

érzékelő rendszerek (Intrusion Detection Systems – IDS) olyan szoftveres, vagy hardveres rendszerek, amelyek automatizálják a hálózatban vagy a rendszerben levő események monitorozását (figyelését), gyanús, támadásra utaló jeleket keresve. Feladatuk a már elkezdett behatolás, támadás felismerése. Más hálózatbiztonsági eszközök kiegészítéseként használják, jellemzően tűzfalakkal együtt. A tűzfalak blokkolni tudnak kimenő, illetve bejövő forgalmat, a behatolás-érzékelő rendszerek viszont valós időben észlelni tudnak egy támadást, és ellenlépésként utasíthatják a tűzfalat egy új szabály felvételére, bizonyos forgalom tiltására.

A támadás-érzékelők két fő csoportba oszthatók a detektálási algoritmus szerint. Az első a *rendellenesség (anomália) alapú (anomaly-based) felismerő*, amely a felhasználók viselkedését monitorozza, és a jellemzőit valamilyen módon tárolja (pl. log). A másik lehetséges módszer a *szabály alapú (rule-based) detektálás*. Előzetes tudással dolgoznak, felderítéskor már az ismert támadások szignatúráit nézik.

3.4.3. Vírusvédelem

Tény, hogy a tűzfal nem nyújt teljes biztonságot. Egy olyan rendszerben, amely tele van biztonsági résekkel, teljesen felesleges tűzfalat használni. Célszerű naponta ellenőrizni az aktuális frissítéseket az operációs rendszerhez, böngészőhöz, és minden olyan alkalmazáshoz, amelynek szüksége van internetes kapcsolatra. Sajnos a szoftveres tűzfalak általában nem nyújtanak kellő védelmet a kémprogramoktól, trójai falovaktól, különböző reklámoktól (spam), levélszemektől, stb. A nagyobb biztonság megteremtésének alapvető eszközei a tűzfal mellett:

- ☞ antivírus (antiviral) program;
- ☞ antitrójai (antitrojan) program;
- ☞ anti-routkit programok;
- ☞ adathalászat szűrő (anti-phishing);
- ☞ kéretlen reklám (spam)- és kémprogram (spyware) tisztító;
- ☞ személyes adatok védelmére szolgáló eszköz.

3.4.4. Identitás elrejtése (anonimizálás)

Az interneten az információk szabadon hozzáférhetők bárkinek, akinek van számítógépe és csatlakozása. Sokszor azonban érdemes elrejtetni a kilétünket és szándékainkat, mert ez másnak fontos információ lehet. A névtelen (anonim) szörfölés nem jelent szükségképpen csalást. Sokkal inkább egy lehetőség arra, hogy ne tegyük nyilvánossá a sze-

mélyazonosságunkat, a céljainkat. Ez egyszerűen csak az az eset, hogy „nem fogok válaszolni, hacsaknem kérdezel”. Vannak alkalmak, amikor úgy akarunk valakivel kommunikálni, hogy nem akarjuk használni a valódi nevünket.

Az interneten szörfözők nyomait nagyon könnyű követni. Mindenfelé hátrahagyunk IP számokat, böngészőjelszavakat, cookie-információkat és sokan éppen ezeket akarják megszerezni. A tűzfal nem nyújt teljes biztonságot, ezért érdemes névtelennek maradni és úgy feltűnni, mintha valaki mások lennénk. Ilyenkor, ha egy hacker meg akarja támadni az internet csatlakozásunkat, akkor nem talál semmit. Több különböző eszközt használhatunk arra, hogy szörfözés közben ne derüljön ki az IP-címünk.

Grátis proxy szerverek: a legegyszerűbbek és a legolcsóbbak az ingyenes proxy szerverek¹⁶⁰. Ezek közvetítőként működnek a felhasználó és a web-szerver között. Ahelyett, hogy a cél URL-t közvetlenül hívná meg, a proxyba kell írni a kívánt címet. Ez lekéri, és megmutatja annak tartalmát a böngészőablakban. Így maga a felhasználó a lekérdezett szerver számára láthatatlan marad, az csak a proxy IP-címét fogja „látni”. Alkalmankénti használatra ez az eljárás megfelel, de igazán biztosnak nem lehet lenni a használatukban, kétes eredetük miatt.

Kereskedelmi forgalomban kapható anonimizálók: ezek a proxyk sokkal megbízhatóbbak. Van olyan program (pl. Anonymizer¹⁶¹), amely ezernyi privát proxy számítógépből álló hálózatot használ.

TOR¹⁶² anonimizáló: egy további ingyenes szolgáltatás. Minden felhasználó csatlakoztathatja számítógépét a rendszerhez és így több száz mini proxy egyesülése áll elő. A rendszer maga dönt az adatsomag útvjáról. Minden közbekapcsolt számítógép csak azt tudja, hogy melyik gépre kell a kapott adatokat továbbítani. Egy kém PC-nek így nincs esélye, hogy valamit kitaláljon ebből az adatkeverékből. Minden lekérdezéshez új útvonalat választ a rendszer. Léteznek olyan *anonimizáló böngésző* programok¹⁶³, amelyek a TOR szerverhálózat használják a felhasználók elrejtésének megvalósításához.

Az anonimizálás gyengeségei:

- ☞ a weboldalakhoz már annyi beépített program (feature, plugin) van, hogy ezek segítségével a képzett hackerek képesek az eredeti IP-címhez jutni;

¹⁶⁰ Az ingyenes proxykról átfogó lista található a <http://www.freeproxy.ru/> és a <http://proxy.org/> címen.

¹⁶¹ Honlapja: <http://www.anonymizer.com/>.

¹⁶² The Onion Router : Hagyma átválasztó.

¹⁶³ *Torpark* nevet kapott böngésző a Mozilla Firefox mobil változatára épül, előre bekonfigurált, telepítést nem igényel.

- ☞ vannak olyan webes szolgáltatások, amelyek használatára névtelenül egyáltalán nincs lehetőség, egyes fórumokba anonim vagy fake¹⁶⁴ címmel be se lehet jutni;
- ☞ jelentősen megnövelheti a letöltések idejét.

3.5. AZ ADATSZERZŐ ALHÁLÓZAT

Az adatszerző alhálózaton belül, mint azt a neve is mutatja, történik a konkrét adatszerzés megvalósítása. Az adatszerző alhálózat alkotó elemei:

- ☞ forráskutató munkahely;
- ☞ forrásértékelő munkahely;
- ☞ forrás-adatbázis;
- ☞ letöltés-vezérlő;
- ☞ kiértékelő munkahely;
- ☞ adatszerző-adatbázis;
- ☞ adatszerző-webszerver.

3.5.1. Forráskutató munkahely

A *nyílt forrású felderítés egyik alappillére a hiteles információforrás*, mert a munka elkezdése elképzelhetetlen nélküle. A nyílt információk tömeges mennyisége, valamint az információk szétszórtsága miatt elsősorban nem az információkra kell koncentrálni, hanem a forrásokra, amelyekből a megfelelő információhoz lehet hozzájutni. Fel kell kutatni a forrásokat, majd elemzés és értékelés eredményeként ki kell válogatni azokat, amelyek a szükséges információkat a szakmai követelményeknek megfelelően tartalmazzák. A *fő cél elsődleges*, vagy olyan másodlagos *forrást találni*, amely *elsődleges információkat tartalmaz*. Elsődleges forrásnak számítanak például a nemzetközi és a kormányzati szervezetek, az egyetemek, a kutató intézetek, valamint a hírügynökségek.

Ezen a munkahelyen folyik a források keresése. Az oldalak keresését mind szerver oldali (pl. Google, Yahoo!), mind kliens oldali kereső programokkal (pl. Copernic Agent) lehet folytatni. Célszerű metakereső (pl. Metacrawler) programokat alkalmazni, mert azok egyszerre több keresővel is kerestetnek. A kezdeti időszak után meg lehet vizsgálni annak a lehetőségét, hogy a keresések megvalósítására egy saját fejlesztésű, vagy a polgári életben már alkalmazott, de a saját igényeink szerint átalakított kereső

¹⁶⁴ fake: hamis.

programot használjunk-e, a hagyományos internetes kereső-megoldások helyett, ami a szokásos szolgáltatásokon (dátum, idő, típus, stb. szerinti keresés) kívül képes:

- ☞ bármely országra (domain) szűkíteni a keresést;
- ☞ bármely nyelvre szűkíteni a keresést,
- ☞ adott nyelven beírt kulcsszót több nyelvre átfordítani és így egy kereséssel több nyelvben találatot produkálni;
- ☞ megkülönböztetni kétféle keresést: új forrás vagy új információ keresés;
- ☞ új forrás keresése esetén automatikusan felismerni a már ismert forrást és azt kihagyni a találati listából;
- ☞ egyszerre több keresővel elvégeztetni a keresést;
- ☞ a mély weben belül is keresni.

A mély web keresésével kapcsolatban fontos megjegyezni, hogy annak keresése nem olyan nehéz, mint amilyennek tűnik. A mély webhez tartoznak a következők: speciális adatbázisok, telefonkönyvek, menetrendek, szabványok, törvények, Web-áruházak katalógusai, nem szöveges fájlok (multimédia, grafikai, szoftver, különleges formátumú), jelszóval védett oldalak, dinamikusan változó tartalmú oldalak (pl. hírek, állásajánlatok, szállodai szobák). Vannak olyan *kereső motorok* (search engine), amelyek képesek a mély web tartalmában is keresni (pl. *CompletePlanet*¹⁶⁵ *Deep Query Manager*¹⁶⁶) bizonyos tartalmak után (képek, hírek). A mély web oldalak szakkönyvtárakban is megtalálhatók, amelyen belül már kereshetők.

A jövőben várhatóan folytatódik az a tendencia, miszerint a szolgáltatók egyezséget kötnek a nagyobb kereső szolgáltatókkal, hogy azok kereshessenek az adatbázisaikban. Ezzel a megoldással az adatbázisokban való keresés önmagában ingyenes lesz, de a hozzáférésért már fizetni kell, hasonlóan a már ma is létező könyvismertetőkhöz. A mély web keresése tulajdonképpen *kettős szintű keresési* feladatot jelent. Az első szint: *megkeresni* azt az *adatbázist*, amelyben a szükséges információkat meg lehet találni. A második szint: magában az *adatbázisban megkeresni* a konkrét információt.

Információigény beérkezése esetén először értelmezni kell a kérést, hogy milyen információkra van szüksége a felhasználónak. Második lépésként meg kell vizsgálni, hogy a rendelkezésre álló anyagokból meg lehet-e válaszolni a kért információkat. Amennyiben ez nem lehetséges, akkor a forrás adatbázisból ki kell választani azokat az internetes weboldalakat, amelyek a legnagyobb valószínűséggel képesek a kért infor-

¹⁶⁵ Honlapja: <http://aip.completeplanet.com/>

¹⁶⁶ Honlapja: <http://www.brightplanet.com/deep-query-manager/dqm/deep-query-manager.html>

mációt megadni, illetve azokat, amelyek kiegészítő információkat adhatnak. A források paraméterei ezután a *letöltés-vezérlőbe* kerülnek, ami megvalósítja azok folyamatos figyelését és a talált anyagok letöltését.

Ha nem áll rendelkezésre egyetlen egy olyan forrás sem az adatbázisban, amelyből a kért információ biztosan megválaszolható, akkor szükség van új források keresésére. Az új források keresése nem csak akkor folyik, amikor valamilyen kérésre kell választ adni, hanem azt folyamatosan kell végezni.

3.5.2. A forrásértékelő munkahely

Ezen a munkahelyen folyik a megtalált források értékelése, ami egy olyan szűkítési eljárás, amelynek célja a forrás használhatóságának meghatározása. A forrás értékelésének lépései a következők:

- ☞ az URL (Uniform Resource Locator) vizsgálata;
 - protokoll vizsgálata;
 - számítógép és a domain név vizsgálata;
 - az útvonal vizsgálata;
 - fájlnev és a kiterjesztés vizsgálata;
- ☞ nyomkövetés;
- ☞ a hitelesség értékelése;
- ☞ a pontosság értékelése;
- ☞ objektivitás értékelése;
- ☞ az időszerűség értékelése;
- ☞ a részletesség értékelése.

3.5.2.1. Az URL vizsgálata

A megtalált forrás kiértékeléséhez először meg kell vizsgálni annak URL-jét. Minden egyes weboldalnak van egy saját URL-je, ami csak rá jellemző. Az URL felépítése a következő:

protokoll://számítógép.domain/útvonal/fájlnev.kiterjesztés

A protokoll vizsgálata. Az általánosan használt protokollok a következők:

- ☞ http:// - Hypertext Transfer Protocol – hiperszöveg átviteli protokoll, a web alapprotokollja;
- ☞ https:// - Hypertext Transfer Protocol Secure - biztonságos hiperszöveg átviteli protokoll;

- ☞ ftp:// - File Transfer Protocol – fájl átviteli protokoll. Az internet klasszikus fájlátviteli protokollja, amely adatállományok két gép közötti kétirányú átvitelét teszi lehetővé;
- ☞ news:// - hírcsoport, kommunikációs fórumok összessége, az elektronikus fa-líújságok és a levelezőcsoportok keveréke témakörök szerint;
- ☞ telnet:// - folyamatos, interaktív terminálkapcsolatot lehetővé tevő protokoll két gép között;
- ☞ mailto:// - levél küldése egy címre.

A számítógép.domain¹⁶⁷ adja meg a szerver nevét, ahol az információ megtalálható (általában a szerver IP címe). A domain nevek utolsó tagja általában az ország nevének kétbetűs – az ISO 3166¹⁶⁸ szabványnak (ld. 3. sz. melléklet) megfelelő – rövidítése (pl. „.hu” Magyarországot jelöli). Előfordulnak ettől eltérő legfelső szintű domain nevek¹⁶⁹ 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;
- ☞ gov - szövetségi kormány (government) hivatalok számára;
- ☞ mil - a hadsereg számára.

A gov, illetve a mil az Egyesült Államok kormányzati, illetve védelmi minisztériumi honlapjait jelenti. A nemzetközi gyakorlatban ezt az elvet alkalmazza a többi ország is, de kiegészítve az ország domain nevével (pl. gov.hu).

Az *útvonal* általában egy könyvtár vagy alkönyvtár nevéből áll, ami azt adja meg, hogy a szerver merevlemezén hol kell keresni az információt.

A *fájlnév és a kiterjesztés* jelenti a kívánt fájlt. Ha nincs konkrét fájl meghatározva, akkor „index.html”, vagy „default.html” jelenik meg a böngésző címsorában. A kiterjesztés alapján tudja a böngésző, hogy hogyan kezelje le az adott fájlt. Van néhány fájl, amit képes a böngésző maga is megjeleníteni, de van olyan is, amelyhez külső

¹⁶⁷ Domain: egy számítógép egyedi neve az interneten.

¹⁶⁸ <http://www.iso.org/iso/en/prods-services/iso3166ma/02iso-3166-code-lists/list-en1.html>.

¹⁶⁹ TDL – top-level domain

program meghívása szükséges. Az interneten jellemzően használt fájlkiterjesztések a következők:

- ☞ *html., htm, xml* – weboldalak;
- ☞ *gif, jpg, jpeg, tiff* – képek;
- ☞ *wav, au, aif, mp3, ogg* – hang fájlok;
- ☞ *zip, .rar, .arj, .lha, .ace, .tgz* – tömörített fájlok;
- ☞ *avi, .mpg, .mov*, - videó fájlok;
- ☞ *txt, .doc, .pdf, wri* – szöveg fájlok.

3.5.2.2. Nyomkövetés

Ahhoz, hogy valakinek domain neve legyen azt regisztráltatni kell egy domain név regisztrátornál¹⁷⁰. A legtöbb regisztrátor biztosít ún. *whois*¹⁷¹ funkciót, aminek a segítségével le lehet kérni a domain név tulajdonosának a következő adatait: a tulajdonos nevét, címét, az adminisztrációs és a technikai kapcsolattartó nevét, email címét, cége nevét, cége címét, telefon- és faxszámát, a regisztráció időpontját.

Sok esetben a domain tulajdonosának ismerete nem elégíti ki az információs szükségletet, mert az sokkal fontosabb lehet, hogy hol van a webszerver fizikailag és hogyan kapcsolódik az internetre. A *Unix*¹⁷² alapú operációs rendszerekben a *traceroute* paranccsal kiírathatjuk a célszerverhez vezető hálózati útvonalat. *Windows* alatt a *tracert* parancs való ugyanerre a feladatra. Az operációs rendszerek parancsain kívül programokat is kifejlesztettek erre a feladatra (pl. *SmartWhois*¹⁷³), amelyek minden regisztrációs adattal együtt tüntetik fel a domain tulajdonosát.

Az internet szolgáltatók (ISP) nevei mellett feltüntetik az infrastruktúrában alkalmazott átviteli formát is (pl. T1¹⁷⁴, T3¹⁷⁵, FDDI¹⁷⁶, ATM¹⁷⁷). Abban az esetben, amikor az ISP-nek csak IP címe van megadva, akkor az *Internet Assigned Numbers*

¹⁷⁰ A magyar regisztrátorok listája a <http://www.domain.hu/domain/> oldalon található.

¹⁷¹ Whois – who is – ki az.

¹⁷² Eredetileg a Bell Laboratóriumban a '70-es évek elején kidolgozott operációs rendszer, amely az első magasszintű nyelvben írt rendszerszoftver volt. A kifejezés napjainkban inkább az eredeti Unix rendszer alapelvein nyugvó, de annak jelentősen továbbfejlesztett változatait képező operációs rendszerek gyűjtőfogalmaként használatos. A Unix architektúra jelenleg is létező legismertebb képviselői a Linux, a Solaris, a HP-UX és a különböző BSD-rendszerek.

¹⁷³ Honlapja: <http://www.tamos.com/>.

¹⁷⁴ T1: az AT&T cég elnevezése az 1.544 Mbit/s sebességű hálózati digitális jelátvitelre.

¹⁷⁵ T3: nagy átviteli sebességet (44.7 Megabit másodpercenként) biztosító nagytávolságú hálózati szabvány (az NREN is ezt használja).

¹⁷⁶ FDDI (Fiber Distributed Data Interface) üvegszál hálózatoknál használt adatátviteli technika, mely akár 100 Mbit/s sebességű átvitelt is lehetővé tesz.

¹⁷⁷ ATM (Asynchronous Transfer Mode) fix hosszúságú (53 byte) csomagokkal dolgozó, igen nagy sebességű (10 Mbit/s - 1 Gbit/s) aszinkron hálózati protokoll.

*Authority*¹⁷⁸ oldal segítségével lehetséges annak a meghatározása, hogy kihez tartozik a cím.

3.5.2.3. *A hitelesség értékelése*

Egy oldal hitelessége megállapításához szükséges lépések a következők:

- ☞ az oldalt üzemeltető társaság vagy személy(ek) tisztázása;
- ☞ van-e olyan hivatkozás, amiben a társaság célkitűzéseiről tájékoztatnak;
- ☞ van-e mód ellenőrizni a társaság működésének törvényességét (email címen kívül van-e megadott telefonszám vagy postai cím, aminek útján több információ kérhető);
- ☞ van-e a közreadott anyag(ok)nak nem csak internetes verziója is, aminek a segítségével ellenőrizhető a törvényesség;
- ☞ van-e ismert szerzőtől származó cikk az oldalon, és annak van-e jártassága az adott témában;
- ☞ tisztázható-e, hogy ki a felelős a tartalomért és az anyagokért;
- ☞ meg van-e adva a szerzői jogokkal rendelkező társaság neve.

3.5.2.4. *A pontosság értékelése*

Az oldal pontossága megállapításához szükséges lépések a következők:

- ☞ a közölt tényeket meg lehet-e erősíteni másik forrás(ok)ból is;
- ☞ van-e felelős szerkesztője a publikált anyagoknak;
- ☞ van-e az anyagokban helyesírási vagy tipográfiai hiba (ez nem mutatja meg feltétlenül az információk ellenőrzésének minőségét, de jelezheti az információk pontatlanságát).

3.5.2.5. *Az objektivitás értékelése*

Az objektivitás értékelésének lépései:

- ☞ el van-e különítve az információs tartalom a reklámoktól és a véleményeket tartalmazó részekről;
- ☞ a tényeket és a véleményeket jól láthatóan elkülönítik egymástól.

3.5.2.6. *Az időszerűség értékelése*

Az időszerűség értékelésének lépései:

- ☞ meg lehet-e tudni azt, hogy az oldalt mikor frissítették;
- ☞ meg lehet-e tudni azt, hogy mióta üzemel az oldal.

¹⁷⁸ Honlapja: <http://www.ianna.org>.

3.5.2.7. A részletesség értékelése

A részletesség értékelésének lépései:

- ☞ van-e olyan információ az oldalon, amely elárulja, hogy mennyire részletesek a forrás(ok) anyagai;
- ☞ van-e az anyagoknak nyomtatott változatai is, és az internetes oldal részben vagy egészben megegyezik-e azzal.

Az értékelés folyamatának végeredményeként használhatónak ítélt források paraméterei bekerülnek a *forrás-adatbázisba*.

Az értékelő munka automatizálása igen komoly kihívást jelent a programozók számára, mert olyan nagyfokú humán és gépi intelligenciát kellene az azt végző programba integrálni, ami csak nehezen megoldható feladat. Ez az oka annak, hogy a források értékelése még jó ideig emberi közreműködést, vagyis fejlett humán intelligenciát igényel.

3.5.3. Forrás-adatbázis

A felkutatott és kiértékelt forrásokat csak akkor tudjuk hatékonyan kihasználni, ha egy jól strukturált adatbázisba helyezzük őket. Az adattárban rögzíteni kell az információs csatornák alábbi jellemzőit:

- ☞ címét (pl. www.army.mil);
- ☞ a megjelentetett információk szerzőinek körét (pl: atomtudósok, kormányzati tisztségviselők, katonák, magánszemély, stb.);
- ☞ az információk témaköreit (pl. biztonságpolitika, terrorizmus, haderő, stb.);
- ☞ a témakörökön belüli témákat (pl. terrorista személyek és terrorszervezetek listája, terrorcselekmények, terrorizmussal kapcsolatos aktuális események, stb.);
- ☞ részletességét;
- ☞ mélységét;
- ☞ az érintett térségeket, államokat vagy szervezeteket (Közel-Kelet, Afganisztán, Nemzetközi Vöröskereszt);
- ☞ a frissítési gyakoriságot (naponta, hetente);
- ☞ osztályozását.

Ezekre a paraméterekre az elemzőknek van szükségük, hogy a lehető legrövidebb idő alatt képesek legyenek kiválasztani a felhasználó kérdésének megválaszolására legalkalmasabb forrásokat. Továbbá ajánlott valamilyen módon osztályozni a forrásokat, illetve az azon belül található témaköröket, mivel vannak olyan weboldalak, amelyek

csak néhány témakörben felelnek meg a kritériumoknak, a többiben nem. A megfelelő mennyiségű forrás teszi lehetővé az eredményes feladatellátást.

Az adatbázisban tárolt források meglétét rendszeres időközönként ellenőrizni kell, mert a források megszűnhetnek, más feltalálási helyre kerülhetnek, megváltozhat a rajtuk megjelenő adatok irányultsága, minősége, szakmaisága, stb.

A forráskutató(k) naponta egy-két új forrást is képes(ek) felkutatni, de azok elemzése és értékelése sokkal hosszabb időt vesz igénybe, mivel az olyan komoly és időigényes háttérmunka, ami a hitelesség és pontosság érdekében nélkülözhetetlen. Így ezt a munkát, véleményem szerint, nem elég egy főnek végezni.

3.5.4. Letöltés-vezérlő

A *letöltés-vezérlő* a feladatát a forrásértékelő munkahely irányítása szerint hajtja végre. A forrásértékelő munkahely jelöli ki azokat a weboldalakat, amelyeket a letöltés-vezérlőnek le kell mentenie, valamint azokat, amelyeket folyamatosan figyelnie kell (állandó INFF források), hogy a rajtuk végrehajtott változásokat a leggyorsabban észlelhessük.

Letöltés-vezérlőként olyan programra van szükség, amely rendelkezik olyan letöltési-mechanizmussal, amellyel teljes weboldalakat, vagy azok egy részét le lehet tölteni a távoli gép fastruktúrájának megfelelően. Továbbá fontos, hogy a letöltési mélységet (rekurzíó) és más paramétereket is meg lehessen adni. A programnak mind *HTTP*, mind *FTP* szervereknél alkalmazhatónak kell lennie.

Letöltés-vezérlőként alkalmazható szoftvereket már számosat kifejlesztettek (pl. *wget*¹⁷⁹, *Surfulater*¹⁸⁰, *Internet Researcher*¹⁸¹). A különálló szoftvereken kívül például a Mozilla Firefox böngésző kiegészítő programja a *Spiderzilla*¹⁸² is képes teljes weboldalak letöltésére, tükrözésére. Hogy melyik ezek közül a célnak leginkább megfelelő program, azt célszerű egy vizsgálat után eldönteni.

Számos hírportál lehetőséget nyújt az RSS csatornáira való feliratkozásra, aminek a segítségével automatikus értesítést küldenek az új hírek megjelenéséről. Ezen kívül a Mozilla Firefox böngészőhöz letölthető olyan kiegészítő program (*Sage*¹⁸³), amely lehetővé teszi az RSS anyagok olvasását és a megadott RSS csatornák figyelését. Az RSS

¹⁷⁹ Letölthető például a <http://wget.sunsite.dk/> honlapról.

¹⁸⁰ Honlapja: <http://www.surfulater.com/index.html>.

¹⁸¹ Honlapja: <http://www.zylox.com/home.shtm>.

¹⁸² Honlapja: <https://addons.mozilla.org/firefox/1616/>.

¹⁸³ Honlapja: <https://addons.mozilla.org/firefox/77/>.

csatornákra való feliratkozás lehetőséget biztosít ahhoz, hogy az *információk mindig a legfrissebbek* legyenek, ami a nyílt forrású felderítés alapvető követelménye.

A letöltés-vezérlőben nagyon fontos programtechnikailag megoldani azt, hogy a forrásokon belül csak azt az információt töltsse le, amire éppen szükség van, és a témával nem kapcsolatos adatokat semmiképpen se töltsse le. Ehhez a forráson belüli hatékony keresési technikák alkalmazására van szükség, ami véleményem szerint jól meghatározott kulcsszavas keresésen alapulhat.

A lementett anyagok egy átmeneti tárolóba kerülnek a letöltés-vezérlő merevlemezén. Innen az információk az *adatfeldolgozó adatbázis*ba csak azután kerülhetnek, ha kiértékelésük megtörtént, vagyis olyan új adatokat tartalmaznak, amire szükség van a felhasználók kérésének megválaszolására, vagy egyéb feladat érdekében érdemes azokat elmenteni. Ezzel a módszerrel elkerülhető, hogy az elemző munkahelyeken dolgozókat a rendszer elárassza túl sok fölösleges anyaggal.

3.5.5. Kiértékelő munkahely

A letöltés-vezérlő által letöltött anyagok nem kerülhetnek közvetlenül az adatszerző-adatbázisba. Ezt meg kell előznie az adatok kiértékelésének. A várhatóan igen sok anyagból, ki kell szűrni azokat az információkat, amelyek alkalmasak a felhasználók kérdéseinek megválaszolására, illetve más szempontok alapján fontosak. Előfordulhat, hogy a különböző forrásokból származó adatok tartalma megegyezik egymással, vagy csak kis mértékben különböznek egymástól. Ilyenkor azt az információt érdemes választani, amely a jobb értékelést kapott forrásból származik, illetve abban az esetben, amikor a különbség érdemben ad több információt, akkor csak ezt a többletet szükséges hozzáadni a másik forrásból származó adatokhoz (a forrás megjelölésével). Az összehasonlítást nem csak az új adatokkal, hanem az adatbázisban tárolt anyagokkal is el kell végezni, mert az is előfordulhat, hogy csak kis mértékű a különbség az újonnan talált és a már eltárolt adatok között. Ebben az esetben csak a különbséget kell lementeni kiegészítésként a már az adatbázisban lévőhöz.

A kiértékelő munka – hasonlóan a forrásértékeléshez – nehezen automatizálható feladat, de vannak a munkának olyan fázisai, amelyeknél az nem jelent nagy nehézséget. Így könnyen megvalósítható az anyagok tartalmának összehasonlítása és azok azonosságának vagy különbségének megállapítása. A probléma akkor van, ha ugyanaz az információ két különböző megfogalmazásban kerül letöltésre, mert ilyenkor egy számítógépes program csak a nagyfokú azonosságot tudja megállapítani, de azt, hogy két kü-

lön anyagról, vagy ugyanannak az anyagnak két különböző változatáról van-e szó azt csak emberi intelligencia képes eldönteni.

A kiértékelési döntéseket nem számítógép hozza, mert azok meghozatalához nem állnak rendelkezésre még hatékony szoftverek. A kiértékelési fázis feladata a feldolgozott és gyűjtött anyagok alapján elvégzett szakértői vizsgálat elvégzése. Ez magába foglalja az addig összegyűlt anyagok, információk összegzését, elemzését, értékelést vagy újabb információk kinyerését információfúzió segítségével. Hozzá tartozik még az információ rendszerezése és adatbázisba tölthetőségének előkészítése, és bizonyos következtetések levonása is. A kiértékelést csak arra alkalmas, tapasztalt és szakértelemmel rendelkező állomány végezheti. A kiértékelés után az értékelt anyagok bekerülnek az adatszerző-adatbázisba.

A 3.1. ábrán csak egy kiértékelő munkahely van ábrázolva, de a valóságban ennél több, legkevesebb kettő munkahelyre van szükség.

3.5.6. Adatszerző-adatbázis

A kiértékelés után mentésre kerülő anyagok az *adatszerző-adatbázis*ba kerülnek. A tárolás előtt meg kell oldani azt a problémát, hogy jelenleg a weblapok kódolása még korántsem egységes, ami a nyelvek betűkészletének sokszínűségéből fakad. Vannak ugyan törekvések arra, hogy olyan kódrendszert alakítsanak ki, amelyik minden különleges karaktert kezelni tudjon, ez az *Unicode*¹⁸⁴ (*UTF-8*)¹⁸⁵ kódolás. Az UTF-8 kódolást azonban még nem mindenki használja. Ez főként a régebben készült honlapok esetén jelent problémát. Az UTF-8 kódolás alkalmas arra, hogy a világ jelenleg használt nyelveihez tartozó karaktereket egyszerre lehessen kezelni, ezért a letöltött fájlokat – ha az szükséges – UTF-8 kódolásúvá kell konvertálni a letárolásuk előtt. Így az adatok feldolgozásakor az elemzőknek már nem kell foglalkozni a megjelenítés problémájával.

Nem csak a kódolást, hanem az adatrögzítés formátumát is szükséges egységesíteni. Minden adatot (szöveg, kép, hang, videó, térkép, stb.) csak egységesített kiterjesztésben célszerű tárolni (pl.: szöveget *txt* vagy *rtf*, képet *gif* vagy *jpg*, hanganyagot *mp3*, videót *mpeg-2* vagy *avi* kiterjesztésben), aminek a kiválasztásánál mindig alkalmazkodni kell a felhasználókhöz.

Az adatszerző-adatbázis a legkritikusabb része a rendszernek. Ennek képesnek kell lennie rendkívül nagymennyiségű adat strukturált, rendezett tárolására, valamint

¹⁸⁴ Unicode: egységes kódtábla.

¹⁸⁵ UTF-8-bit: Unicode Transformation Format – 8 bites Unicode átalakítási formátum.

lehetővé kell tennie az adatok megadott feltételek szerinti (tipizált) lekérdezhetőségét, mindezt gyorsan, megbízhatóan, az adatvesztés lehetőségének a minimálisra való leszorításával.

Ennek a problémának a megoldására a legcélszerűbb komplett adatbázisrendszert (pl. *Oracle*, *DB/2*, *MS-SQL MySQL*, *PostgreSQL*) vásárolni olyan szakmai cégektől (pl. *Oracle*¹⁸⁶, *IBM*¹⁸⁷, *Microsoft*¹⁸⁸), amelyeknek ebben sok éves tapasztalatuk van, és képesek az igényeknek megfelelő rendszer kialakítására, annak rendszeres karbantartására, szervizelésére, valamint a változások megkövetelte továbbfejlesztésére. Ez nyilvánvalóan nem a legolcsóbb megoldás, de a legcélravezetőbb, mert így biztosan megfelelően lehet kialakítani az adatbázist, hiszen azt arra felkészült szakemberek készítik el az igényeknek megfelelően.

Az adatbázisban való sikeres keresések végrehajtásához és az összefüggések megtalálásához az elmentett anyagokat minél több paraméterét (metaadatát) szükséges megadni:

- ☞ feltalálási helyének neve (pl. The Federation of American Scientists /FAS/);
- ☞ feltalálási helyének URL-je (pl. <http://www.fas.org/>);
- ☞ feltalálási hely szolgáltatója (pl. Federation of American Scientists);
- ☞ szerző (pl. Henry Kelly);
- ☞ szerző ismert adatai (pl. PhD, president of FAS);
- ☞ kiadó (publisher);
- ☞ a mű eredeti címe (title);
- ☞ keletkezésének ideje;
- ☞ letöltésének ideje;
- ☞ témája, témaköre (subject);
- ☞ rövid tartalma (description);
- ☞ nyelvezete (language);
- ☞ azonosító (identifier);
- ☞ jogok (rights);
- ☞ közreműködő (contributor);
- ☞ kapcsolatok más dokumentumokkal (relation);
- ☞ kulcsszavak.

¹⁸⁶ Honlapja: <http://www.oracle.com/>.

¹⁸⁷ Honlapja: <http://www.ibm.com/>.

¹⁸⁸ Honlapja: <http://www.microsoft.com/>.

A kulcsszavak automatikusan generálhatók olyan programmal, amely az anyagban előforduló leggyakoribb szavakat, illetve különleges szavakat választhat ki kulcsszónak.

3.5.7. Adatszerző-webszerver

A megfelelő struktúrával kialakított adatbázis lehetővé teszi, hogy az abban tárolt anyagokat valamilyen megadott feltétel(ek) alapján lekérdezhessük. Minden adatbázis szolgáltató rendszer lehetőséget biztosít arra, hogy webes felületen keresztül lehessen megvalósítani a tárolt adatok lekérdezését. Emellett arra is lehetőség van, hogy magát az adatbázist is karbantartsuk webes eszközökkel. Az adatszerző alhálózaton belül a *webszerver*nek az a feladata, hogy a szolgáltatásai segítségével az adatszerző-adatbázisban különböző célú kereséseket és lekérdezéseket lehessen végrehajtani. Ezeket a feladatokat mind a kiértékelő munkahely, mind az adatfeldolgozó alhálózat részeként telepített *elemző munkahelyek* végre tudják hajtani. Az előbbi ezeken felül képes az adatbázis karbantartásának végrehajtására is a webszerver szolgáltatásain keresztül.

A megfelelő biztonság kialakítása érdekében a webszervert a kiértékelő munkahely kivételével csak az elemző munkahelyek érhetik el, de azoknak felügyeleti vagy egyéb vezérléssel, irányítással, karbantartással kapcsolatos jogosultságaik nincsenek. A webszerver csak a belső hálózat felé lát el feladatokat, az internettől teljesen el van választva, a tűzfal segítségével.

3.6. ADATFELDOLGOZÓ ALHÁLÓZAT

Ezen az alhálózaton belül történik az adatok feldolgozása, illetve a feldolgozás eredményeként – a végfelhasználók igényeinek megfelelő – jelentések, összefoglalók, háttértájékoztatók, elemzések elkészítése és eltárolása a feldolgozó adatbázisba. Az elkészült jelentések azonnal hozzáférhetők a végfelhasználók számára, amint tárolásra kerülnek a feldolgozó adatbázisban. A jelentésekhez és egyéb elkészült anyagokhoz való hozzáférést a feldolgozó-webszerver biztosítja.

Az adatfeldolgozó alhálózat rendszertechnikai elemei a következők:

- ☞ elemző munkahelyek;
- ☞ feldolgozó-adatbázis;
- ☞ feldolgozó-webszerver.

3.6.1. Elemző munkahelyek

Az *elemző munkahelyeken* folyik a megszerzett információk feldolgozása. Az egyes munkahelyek fizikai kialakítása között nincs különbség. Az eltérés a munkavégzés nyelvében, témakörében, feladatában lehet, ami lehetővé teszi a párhuzamos, több irányú, vagy több témájú, egyidejű feldolgozást. Így egy végfelhasználó több igényének, vagy több felhasználó külön-külön igényeinek a kiszolgálása is lehetséges.

Az elemzőknek – megfelelő jogosultságaik alapján – lehetőségük van az adatszerző-adatbázis elérésére és abban, a részükre megszabott feladatnak megfelelő, kereső, lekérdező, elemző feladatok futtatására. A sikeres kereső-lekérdező munka eredményeként képesek lehetnek az információk közötti összefüggések feltárására, következtetések levonására, minek alapján összeállítják a kívánt formátumú és tartalmú jelentéseiket, amelyeket a feldolgozó-adatbázisba mentenek.

A jelentés mérete attól függ, hogy az kinek készül. Ebben egy ökölszabály:

- ☞ vezetőknek, parancsnokoknak (döntéshozóknak) csak rövid téma-összefoglaló;
- ☞ tanácsadóknak (törzstiszteknek) komplett összefoglaló, amely az adott témára fókuszál, annak háttérének ismertetésével;
- ☞ elemzőknek általában nagyon apró részletekre is szükségük van, hogy kialakíthassák az adott témával kapcsolatos véleményüket.

A jelentéseket minden esetben el kell látni rövid téma-összefoglalóval és tartalomjegyzékkel, illetve csatolni kell a hivatkozásokat, vagyis azokat a weboldalakat, dokumentumokat és szerzőiket, térképeket, képeket, valamint táblázatokat, amelyek felhasználásra kerültek.

3.6.2. Feldolgozó-adatbázis

Ez az adatbázis hasonló szerepet tölt be ebben az alhálózatban, mint az adatszerző-adatbázis az adatszerző alhálózaton belül. A különbség abban van, hogy ebben az adatbázisban nem csak a letöltött anyagok kerülnek tárolásra, hanem olyanok is, amelyek az elemzések sikeres elvégzéséhez szükségesek, és többlettudást biztosítanak.

Az elemzők által végrehajtott vizsgálatok, összehasonlítások, értékelések és elemzések eredményeként előállított összefoglaló *jelentések* az adatbázis egyik részébe kerülnek. A jelentéseket csak *magyar nyelven* szabad elkészíteni, mert a felhasználóknak nincs idejük a fordítással foglalkozni, így meg kell oldani az idegen nyelvű anyagok

lefordítását is. Ehhez már számos jól működő fordítóprogram áll rendelkezésre, még különleges nyelvek esetén is, de azok elsősorban csak angolra tudnak fordítani.

Az adatbázis másik részébe olyan *háttéranyagok* kerülnek feltöltésre, amelyek az elemzésekhez elvégzéséhez nyújtanak segítséget. Ezek a következők:

- ☞ lexikális adatokat;
- ☞ általános adatokat (földrajzi, éghajlati, kulturális, népességi, stb.);
- ☞ hadigazdasággal, haditechnikával, hadszíntéri objektumokkal kapcsolatos adatokat;
- ☞ haderőkre vonatkozó adatokat.

Ezekre az adatokra azért van szükség, mert sok esetben nem lehet az újonnan le-töltött anyagokat értelmezni kiegészítő információk nélkül.

3.6.3. Feldolgozó-webszerver

A feldolgozott információkat el kell juttatni a felhasználókhöz. Ennek a legegyszerűbb és leggyorsabb módja az elektronikus továbbítás, mivel az információs társadalomban már feltételezhető a felhasználókról, hogy rendelkeznek internet elérési lehetőséggel.

Elektronikus formában két módon lehet eljuttatni az információs igény(ek)re adott válaszokat, illetve a napi események összefoglalóját:

- ☞ elektronikus levelező rendszeren;
- ☞ feldolgozó-webszerver szolgáltatásának igénybevételével.

Az *elektronikus levelező rendszeren* keresztül való terjesztés egy védett „személyre szóló” információterjesztési forma. Hátránya, hogy a felhasználónak folyamatosan figyelnie kell a levelező rendszerét. Mivel a levelező rendszeren keresztül általában csak korlátozott mennyiségű információ küldhető, ezért függ a fogadó fél szolgáltatásának tároló kapacitásától.

A felhasználók – megfelelő jogosultság megadása után – képesek lehetnek a *feldolgozó webszerver szolgáltatásain keresztül* az elkészült, és a feldolgozó-adatbázisba elmentett jelentésekhez való hozzáférésre, azok olvasására, vagy letöltésére, illetve elmentésére. A rendszert kombinálva az RSS technikával a végfelhasználók a jelentések elkészülte után azonnal értesítést kaphatnak az anyagok elkészültéről. Megfelelő naplózás segítségével nyilván lehet tartani, hogy melyik felhasználó, mikor, milyen jelentést nézett meg, töltött le, ami fontos adminisztrációs és biztonsági követelmény.

3.7. VÉGFELHASZNÁLÓI ALHÁLÓZAT

Ez az alhálózat tulajdonképpen nem tartozik szorosan a rendszermodellhez, de logikailag feltétlenül a részének kell tekinteni, mivel a rendszer célja a felhasználók igényeinek kiszolgálása. A felhasználók a router2 (útvonalválasztó) keresztül, mind az internetről mind a szervezet saját belső hálózatáról (intranet), online módon kérdezhetik le az elkészült jelentéseket, háttéranyagokat, összefoglalókat, stb.

3.8. ÚTVONALVÁLASZTÓK

Az adatszerző alhálózat és az adatfeldolgozó alhálózat között lévő útvonalválasztó (*router1* – R_1), illetve az adatfeldolgozó alhálózat és a végfelhasználói alhálózat között lévő útvonalválasztó (*router2* – R_2) nem tartozik topológiailag egyik alhálózathoz sem. Feladatuk az alhálózatok egymással való összekötése, illetve az egyes alhálózatok forgalmának szétválasztása, valamint a biztonság megvalósítása azáltal, hogy az adott alhálózat webszerverét másik alhálózatból csak az arra jogosultaknak teszik lehetővé.

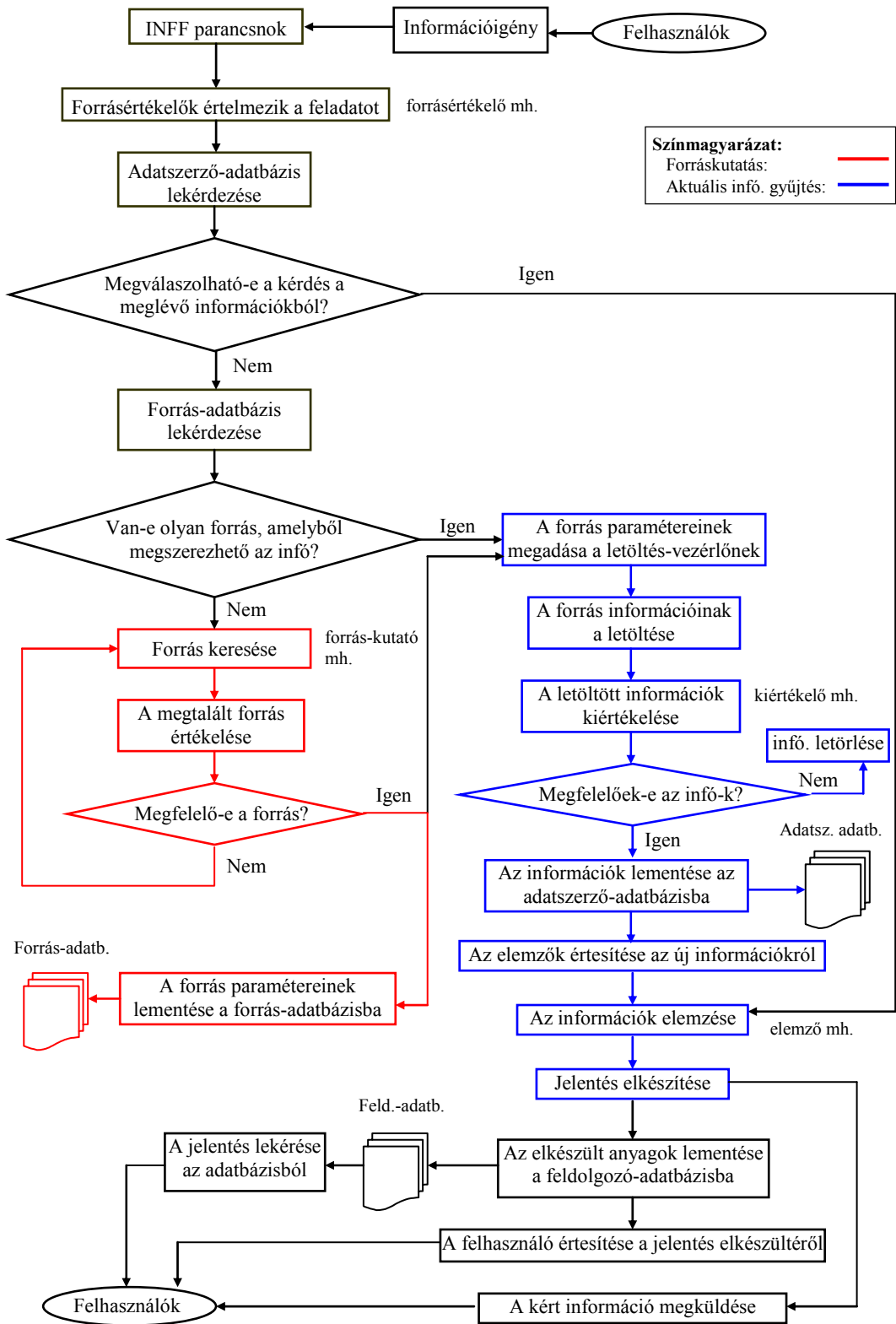
3.9. A RENDSZER MŰKÖDÉSE

Attól függően, hogy milyen feladatot kell ellátnia a rendszernek, három működési folyamatot lehet megkülönböztetni, amelyek a következők:

- ☞ felhasználóktól érkező információigény kielégítése;
- ☞ forráskutatás;
- ☞ aktuális információk gyűjtés és összefoglaló jelentés készítése.

A 3.2. ábrán látható folyamatábra egyben szemlélteti a modell minden működési folyamatát, színekkel megkülönböztetve azokat egymástól.

Amikor a *felhasználóktól információigény érkezik* be az INFF szolgálathoz, akkor azt először a vezető (parancsnok) értelmezi, majd kiadja a feladatot a forrásértékelőknek. Azok megvizsgálják az adatszerző-adatbázisban tárolt anyagokat, hogy van-e közöttük olyan, amiből esetleg ki lehet elégíteni az igényt. Amennyiben már van megfelelő információ, akkor végre kell hajtani annak feldolgozását, majd a jelentés elkészítését. Ha nincs, akkor megvizsgálják a forrás-adatbázist, hogy abban vannak-e olyan ismert források, amelyekből az információ valószínűleg megszerezhető. Ha nincs, akkor el kell kezdeni a megfelelő források felkutatását.



3.2. ábra A rendszer működésének folyamatábrája

Ha van – egy vagy több – megfelelő forrás, akkor annak (lehetőség szerint több mint egy, de tíznél nem több) a paraméterei bekerülnek a letöltés-vezérlőbe, ami automatikusan letölti a forrásban található információkat. A letöltött információkat felhasz-

nálás előtt ki kell értékelni. A szükségtelen vagy megegyező információkat ki kell szűrni, a szükségeseket el kell tárolni az adatszerző-adatbázisba, amiről értesítést kapnak az elemzők.

Az elemzők az adatszerző-adatbázisból kiveszik az anyagot és elemzés, értékelés, összehasonlítás és esetleges kiegészítés után elkészítik a felhasználóknak szóló jelentést, amit a feldolgozó-adatbázisban tárolnak el. A jelentés elmentésével egyidőben a felhasználó értesítést kap annak elkészültéről, vagy igény szerint részére az el is küldhető. A jogosult felhasználók az értesítés megkapása után bármikor elolvashatják, vagy akár le is tölthetik az elkészült jelentéseket.

A megbízható információszerezés alapja a megfelelő források megtalálása, ezért a forrásokat folyamatosan keresni kell. A *forráskutatás* folyamata a 3.2. ábrán piros színnel van jelölve. A források folyamatos keresése a forráskutató munkahely feladata, amit kereső programok felhasználásával végeznek a forráskutatók. A megtalált források közül a megfelelőt a forrásértékelő választja ki a 3.5.2. pontban ismertetett szempontok alapján. A forrás fontosabb paraméterei a forrás-adatbázisban kerülnek eltárolásra.

A napi *aktuális információk (current informations) gyűjtése* fontos lehet az események jobb megértése, illetve a világban történő változások követésének céljából. Ennek a folyamata a 3.2. ábrán kék színnel van jelölve.

A munka a források adatainak a letöltés-vezérlőbe történő betöltésével kezdődik. A források információi automatikusan letöltésre kerülnek, amik között sok azonos is lehet, vagy olyan, ami számunkra érdektelen. Ezek kiszűrését a kiértékelő munkahelyen dolgozó kiértékelők végzik. A releváns információk az adatszerző-adatbázisba kerülnek letárolásra, amiről az elemző munkahelyek értesítést kapnak. Ezután az információkat az elemzők elemzik, értékelik, összehasonlítják a korábbi ismeretekkel és elkészítik a megadott méretű és formátumú jelentést, amit a feldolgozó-adatbázisba mentenek el.

3.10. A SZÜKSÉGES GÉPPARK

3.10.1. Számítógépek konfigurációja

A számítógépek mai gyors fejlődése mellett nehéz előre megmondani azt, hogy milyen legyen a gépek konkrét *konfigurációja*. A mai fejlődési tendenciák alapján annyit nyugodtan ki lehet jelenteni, hogy a nagyobb számítási feladatot végző számítógépeket többmagos processzorral, több GB-os operatív memóriával, gyors és nagykapacitású (>300 GB) merevlemezekkel kell ellátni. Vannak a rendszernek olyan elemei, amelyek-

nek a megvalósításához nem szükséges nagy számítási- vagy tárolókapacitás, ezeket meg lehet valósítani egymagos processzorral, kevesebb memóriával és tárolókapacitással is (pl. tűzfal, elemző munkahely).

Megjelenítő eszközként – a munkavédelmi előírásoknak megfelelően – csak *LCD*¹⁸⁹ monitort szabad alkalmazni, amelyek legalább 19 colos átlóval rendelkeznek, mert ezek biztosítanak kellően nagy felbontást, továbbá nem bocsátanak ki káros sugárzást, s a képük sem villózik, így nem károsak a szemre. Az elemző tevékenységet jelentős mértékben elősegítheti a több képernyős munkahely kialakítása, esetleg intelligens digitális óriás falitábla alkalmazása.

Archiválás megvalósítása. A kisebb méretű adatmennyiség archiválására jelenleg a legalkalmasabb eszközök a CD-ROM kapacitását (800 MB) jóval meghaladó *DVD-ROM*-ok. Az egyrétegű lemez kapacitása 4,7 GB, kétrétegű esetén ennek a duplája. A nagyobb kapacitású, de még csak kevéssé elterjedt *HD-DVD*¹⁹⁰ (15-30 GB), illetve a *Blu-ray*¹⁹¹ (50 GB) eszközök még nagyon drágák. A CD/DVD-ROM-ok alkalmazhatók még az adatok eljuttatására olyan felhasználókhöz, akik valamiért nem csatlakoznak az internetre, illetve kisebb méretű anyagok archiválására.

A nagyobb mennyiségű adatok biztonsági mentésének végrehajtására (például teljes merevlemez anyagának lementése) a *mágnesszalagos egységek (DAT egység)* alkalmasak, mivel ezeknek van ehhez megfelelő tárolókapacitása (jellemzően: 150-330 GB).

3.10.2. Számítógépek száma

A számítógépek száma teljes mértékben attól függ, hogy mekkora költségvetésből lehet gazdálkodni, mert az internet figyelését akár egy számítógéppel is meg lehet oldani, de kérdés, hogy érdemes-e. A megfelelő megoldás természetesen az lenne, ha a 3.1. ábrán ábrázolt minden rendszertechnikai elemet legalább egy külön géppel lehetne megvalósítani, mert csak így biztosítható a rendszer megfelelő megbízhatósága. Így egy elem esetleges meghibásodása esetén nem a teljes rendszer esne ki, hanem csak egy része, ami az esetleges végzetes hibák valószínűségét jelentősen lecsökkentheti.

Az egyes feladatokhoz szükséges számítógépek számát a 3.2. táblázat mutatja.

¹⁸⁹ LCD: Liquid Crystal Display – folyadékkristályos kijelző.

¹⁹⁰ HD DVD High Definition Digital Versatile Disc – Nagyfelbontású digitális rugalmas lemez. Támogatói: Intel, HP, Microsoft, NEC, Sanyo és a Toshiba.

¹⁹¹ Blu-Ray – Kék sugár (utalás a lézer színére). Támogatói: Apple, Dell, HP, Hitachi, LG, Mitubishi Electronics, Panasonic, Pioneer; Phillips, Samsung, Sharp, Sony; TDK, Thomson, Twentieth Century Fox, Disney, és a Warner Brother's Entertainment.

Munkakör	Darabszám
Tűzfal	1-3
Forráskutató munkahely	1-2
Forrásértékelő munkahely	1-2
Forrásadatbázis	1
Letöltés-vezérlő	1
Kiértékelő munkahely	2-3
Adatszerző-adatbázis	1
Adatszerző webservert	1
Elemző munkahely	2-3
Feldolgozó adatbázis	1
Jelentés webservert	1
Összesen:	13-19

3.1. táblázat A szükséges számítógépek száma

3.10.3. Operációs rendszer

Tökéletes operációs rendszer nincs. Mindegyiknek vannak kisebb-nagyobb hibái, esetleg túl drága, vagy egyszerűen elavultnak tekinthető. Az operációs rendszerek kiválasztásánál fontos szempont, hogy támogassák a PC-ekkel kapcsolatos karbantartási, üzemeltetési tevékenységet, biztosítsák a felhasználók részére azt, hogy bármelyik munkaállomásról bejelentkezve a körülményekhez képest ugyanazt a környezetet kapják, ugyanazokkal a programokkal. Biztonságos legyen, mind illetéktelen behatolásokkal szemben, mind véletlen felhasználói műveletekkel szemben. Lehetőleg minél kisebb legyen a hardver igénye.

Mivel nem nagy és bonyolult hálózat kialakítása a cél, hiszen csak a személyzet tagjai alkotják a felhasználókat, így a hálózat fizikai komplexitása, szoftverek használatának engedélyezése, illetve a hálózat menedzselése nem olyan nagyságrendű feladat, mint egy nagyobb vállalati, vagy akár egy nagy egyetemi hálózatnál, viszont fontos szempont a rugalmasság, hogy a kor mindenkori követelményeinek képes legyen minél kisebb anyagi és szellemi ráfordítással megfelelni.

A következő fontos szempont az adatbiztonság és adatvédelem kérdése. Mivel az internethez kapcsolódva folyamatosan fennáll annak veszélye, hogy illetéktelenek férhetnek hozzá a rendszerünkhöz, ezért rendkívül fontos, hogy a belső hálózat védett legyen. Védni sokféleképpen lehet, de csak UNIX-alapú rendszerekkel érdemes, mivel beépítve tartalmazzák az összes – adatvédelemhez és adatbiztonsághoz tartozó – funkciót, amelyekkel nagy biztonsággal lehet védeni a hálózatot. Nem véletlen, hogy az internetes, illetve az internettel bármilyen kapcsolatban álló szervereken túlnyomórészt UNIX alapú rendszereket használnak.

3.11. HUMÁN ERŐFORRÁS

A vizsgálataim alapján javasolt szervezet létszáma nagymértékben függ a rendszertechnikai kialakítástól és a műszakok számától, amennyiben csak egy műszakban működik, akkor kisebb, ha három műszakban működik, akkor nagyobb a humán erőforrásigény.

Javaslatom szerint a rendszer megfelelő működtetéséhez szükség van legalább:

- ☞ *egy fő vezetőre* (parancsnok), aki irányítja az INFF szervezetet;
- ☞ *egy fő* hálózatépítésben és kezelésben jártas szakemberre (*rendszergazda*), aki irányítani tudja a rendszer kiépítését, telepíteni tudja a programokat, szervezi a karbantartást és a teljes rendszer üzemeltetését;
- ☞ *egy-két fő* keresési technikában jártas szakemberre (*forráskutató*), aki(k) ért(enek) az interneten való forráskereséshez;
- ☞ *egy-két fő* forráselemzésben jártas szakemberre (*forrásértékelő*), aki(k) képesek a források értékének meghatározására;
- ☞ *két-három fő* információ kiértékelésben jártas szakemberre (*értékelő*), akik képesek a nagytömegű adatokból kiválasztani a felhasználók kérésének megfelelő információkat és komoly nyelvismerettel is rendelkeznek;
- ☞ *egy fő* adatbázis kezelésben, karbantartásban jártas szakemberre (*adatbázis-adminisztrátor*), aki kezeli, felügyeli, karbantartja a különböző adatbázisokat;
- ☞ *két-három fő* jelentéskészítésben jártas, nyelvismerettel rendelkező szakemberre (*elemző*), akik képesek az adatbázisban lévő anyagokból a felhasználók igényének megfelelő típusú, mélységű, összetételű és formátumú jelentések összeállítására.

Az interneten ma már olyan sok információ található, hogy nem az a probléma, hogy találunk-e egyáltalán valamit, hanem az, hogy az összegyűjtött anyagot fel tudjuk-e dolgozni a megadott határidőn belül. Ez indokolja, hogy az értékelők és az elemzők létszáma a lehető legnagyobb legyen. Az értékelők hatékony munkavégzése megakadályozhatja, hogy az elemzőkre túl nagy mennyiségű információ zúduljon, mert kiszűrhetik a nagymértékben azonos vagy a kérdés megválaszolása érdekében érdektelen információkat. Az értékelő és elemző feladatok ellátásához nem elsősorban számítástechnikai ismeretekre, hanem jó elemző-értékelő készségre, kiváló nyelvtudásra, fordítókészségre és jelentéskészítői gyakorlatra van szükség.

Az INFF rendszer humán erőforrásigénye:

Munkakör	Létszám /fő/
Vezető (parancsnok)	1
Rendszergazda	1
Forráskutató	1-2
Forrásértékelő	2-3
Kiértékelő	2-3
Adatbázis-adminisztrátor	1
Elemző	2-3
Összesen:	10-14

3.2. táblázat *A tervezett munkakörök és az azokat ellátók létszáma*

Abban az esetben, ha a rendszer több műszakban működik, akkor a forráskutató, a forrásértékelő, a kiértékelő és az elemző munkakörök számát a műszakok számának megfelelően meg kell növelni.

3.12. A SZOLGÁLTATOTT INFORMÁCIÓK MENNYISÉGE

A szolgáltatott információ mennyisége attól függ, hogy mennyi csatornát, oldalt kívánunk folyamatosan figyelni, mivel ezek számával együtt nő az információk tömege. A letölteni kívánt anyagok típusa is befolyásolja a fájlok méretét. A csak szöveges tartalmú fájlok a legkisebbek (pl. csak 5 MB-ot tesz ki Shakespeare összes műve ebben a formában) és a videók a legnagyobbak (jellemzően: tömörítve is 1 perc kb. 10 MB). Természetesen a különböző weboldalakon eltérő mennyiségű információt tárolnak.

A napi aktuális (current) eseményekről készülő jelentések készítéséhez szükséges információmennyiség hozzávetőleges számvetéséhez letöltöttem az ORIGÓ ITTHON rovatát a Mozilla Firefox böngésző Spiderzilla nevű kiegészítése segítségével. Az eredmény: 159 MB anyag; több mint 2000 html oldal; az oldalak mérete: 70-90 kb-ot; a legrégibbi oldal 1999. áprilisi 21-ei; adott napi cikkek száma: 6, azok együttes mérete 0,5 MB. Ez azt jelenti, hogy forrásonként durván 0,5-2 MB új anyag keletkezésével lehet számolni naponta abban az esetben, ha csak html oldalakat töltünk le. Így például, ha 100 forrást figyelünk meg egyszerre, akkor a napi új információ 50-200 MB-ra tehető, ami évente kb. 20-100 GB anyag keletkezését és annak tárolását és feldolgozását jelentheti. Ez az adatbázis méreteinek meghatározásához adhat támpontot. Ezek a számok erősen túlzóak, mert nem valószínű, hogy egy forrásnál minden új hírre szükség van.

3.13. A MEGSZEREZHETŐ INFORMÁCIÓK TÍPUSAI

A rendszer első közelítésben, a szöveges és képi dokumentumok figyelésére lehet alkalmas, mivel ezek megtalálása és figyelése technikailag könnyen megvalósítható, mert már léteznek olyan programok, amelyek ezt képesek megvalósítani.

A rádióadások vétele nagyon egyszerűen megvalósítható és a hanganyag tömörített tárolása technikailag kidolgozott. Gondot a leszedett anyagok kiértékelése, végighallgatása, feldolgozása jelent, mert nagyon sok időt vesz igénybe azokat végighallgatni. Hangkereső szoftverek egyelőre még csak kísérleti stádiumban vannak.

Az IPTV adások sikeres vételéhez feltétlenül szükséges a minél szélesebb sávú internet csatlakozás. A TV műsorok vétele és tömörítése is egyszerűen megvalósítható. Itt is az értékelés és feldolgozás jelenti a problémát, mert gépesítése még nem megoldott.

3.14. A MEGFIGYELÉSRE ÉRDEMES OLDALAK

A rendszer kezdeti működése idején, amíg még nem áll rendelkezésre megfelelő mennyiségű elsődleges forrás, addig elsősorban a hírportálok figyelésével lehet elkezdni a munkát, mert ezek segítségével folyamatosan figyelemmel lehet kísérni az aktuális (current) eseményeket. A kereső rendszer fejlődésével, és az ismertté vált források szaporodásával, folyamatosan lehet bővíteni a megfigyelt oldalak számát, így kiterjeszteni a figyelést a katonai, gazdasági, pénzügyi, tudományos és kulturális jellegű információkat tartalmazó oldalakra is.

Külön említést érdemel a *blogok figyelése*. A mai média-központú világban a politikusok felismerték azt, hogy az elektronikus napló segítségével gyorsan, közvetlenebb hangnemben, a hivatalos tájékoztatástól némileg eltérő szemszögből tájékoztathatják a híveiket. Az ilyen személyes oldalakon elejthetnek a politikusok és egyéb prominens személyiségek olyan információkat, amelyek nagyon fontosak lehetnek az információszerezés szempontjából. Az Egyesült Államokban már külön weblog monitoring részleggel működtetnek a CIA¹⁹² és az NSA¹⁹³ számára.

A rendszer legfontosabb tulajdonsága lehet – hatékony kereső eljárás alkalmazásával –, hogy megtalálja azokat – az akár csak rövid ideig a hálózaton lévő és nem fel-

¹⁹² CIA: Central Intelligence Agency – Központi Hírszerző Ügynökség. Honlapja: <https://www.cia.gov/>.

¹⁹³ NSA/CSS: National Security Agency/Central Security Service – Nemzetbiztonsági Ügynökség/Központi Biztonsági Szolgálat. Honlapja: <http://www.nsa.gov/>.

tétlenül hivatalos – anyagokat, amelyek hasznos információkat tartalmaznak az éppen a felderítés fókuszában lévő területekről.

3.15. A RENDSZER ALKALMAZÁSÁNAK SZINTJEI

A rendszer alkalmazható *makró szinten*, ami lehet országos, ágazati, vagy műveleti szint. Ezt teszi lehetővé, hogy a teljesen kiépített rendszer képes globális és egyéb internet alapú nyílt információk összegyűjtésére, értékelésére, szelektálására, tárolására, átalakítására és nem utolsósorban azok gyors feldolgozására még nagy tömeg esetén is.

A rendszer kisebb méretekben, vagyis *mikró szinten*, megvalósított (mobil eszközbe szerelt) változata alkalmas lehet a Magyar Honvédség külföldön szolgálatot teljesítő kontingenseinek regionális vagy lokális internet alapú információval való ellátására és az adott területen egyéb módon folytatott felderítés kiegészítéseként. Ilyen képességekre abban az esetben lehet szükség, amikor nem a NATO, hanem az Európai Unió, az ENSZ vagy más szervezet keretein belül hajtunk végre valamilyen műveletet. Ilyen esetben – *képesség felajánlás formájában* – a MH biztosíthatná ezt a képességet a többi ország kontingensei számára.

Amennyiben nincs lehetőség egy misszió során, létszám és más okok miatt, a komplett mikro rendszer kiküldésére, akkor legalább egy forráskereső munkahelyet érdemes lenne kialakítani a kontingensen belül. Ez a forráskereső munkahely, amelyhez az internetes csatlakozást és az IP címet értelemszerűen az adott ország internet szolgáltatója biztosítja, kibővítheti a honi (makro) forráskutatás képességeit, mert olyan IP címet és szolgáltatót alkalmaz, amely nem okoz feltűnést ahhoz képest, mintha itthonról folytatnánk ezt a tevékenységet. A rendszer többi feladatát, az internetes kapcsolat révén, már a hazai rendszer végezheti el.

3.16. A RENDSZER KIALAKÍTÁSÁNAK VÁRHATÓ KÖLTSÉGEI

A teljes rendszer kialakításához szükséges számítógépek, az adatbázis gépeit kivéve, nem igényelnek olyan különleges kialakítást, ami jelentősen megnövelné az árakat. Komolyabb költséget az adatszerező adatbázis gépe(i) képviselhetnek. A többi számítógép beszerzése, véleményem szerint, hazai forrásokból is megoldható lehet. Természetesen ezeknél is törekedni kell arra, hogy olyan konfigurációk kerüljenek kialakításra, amelyek az adott feladat ellátásához megfelelő számítási- és tárolási kapacitással ren-

delkeznek. A kialakítástól függően a számítógépek kb. 6-10 millió forintból megvalósíthatók (adatbázis: kb. 1 millió, 1 szg.+monitor: kb.: 400-500 ezer Ft). Ehhez jönnek még a szükséges szoftverek árai, ami szintén több millió forintot tehet ki. Költséggént merül még fel a rendszer elhelyezéséhez szükséges helyiség kialakítása, abban az elektromos- és számítógépes hálózat kiépítése, valamint az internet csatlakozás előfizetése. A teljes rendszer beindításának költsége kb. 30-50 millió forintra tehető.

A rendszer nagy előnye, hogy csak a kialakítása követel ekkora összeget, a fenntartása ennél sokkal kisebb költségekkel jár, mert a számítógépeket évekig nem kell újra cserélni és a jól bevált szoftvereket sem kell minden évben lecserélni csak azért, mert új jelent meg, ami alig valamivel tud többet az előzőnél. Működtetése során esetleg további költséggént jelentkezhet néhány új forrásra vagy internet kapcsolatra való előfizetés.

3.17. ÖSSZEGZÉS ÉS KÖVETKEZTETÉSEK

Az információs műveletek új hadszíntere a kiber-tér, ez a számítógépek alkotta virtuális világ, amelyben a valóságos világhoz hasonlóan folyik a küzdelem. Ez a harc-tér egyre fontosabb lesz, mert a világháló nagy léptékben fejlődik mind a méretére, mind a szolgáltatásaira nézve.

Napjaink sürgető feladata, hogy rendelkezünk megfelelő eszközökkel a kiber-térben való információszerzésre, vagyis a virtuális világból valóságos információk kinyerésére. Az internetről való információszerzés nem új dolog, de a honi információszerzésben, véleményem szerint, még nincs a jelentőségének megfelelő szinten kezelve. Az információszerzésnek ez a módja a számítástechnika, a távközlés és a média konvergenciájának, vagyis az infokommunikációnak köszönhetően rövid időn belül – kevesebb, mint tíz év – domináns szerephez fog jutni az összes információgyűjtő eljárás között. Ez az új szerep kényszeríti ki azt, hogy ezt a területet külön információszerző tevékenységként, javaslatom szerint INFF-ként, kezeljük, (hasonlóan ahhoz, ahogy a rádiófelderítés kivált egykor a híradó szolgálatból).

Az INFF eredményes végzéséhez szükség van egy olyan rendszer kialakítására, amely képes az internetről való információszerzésre úgy, hogy a kereső, értékelő és elemző feladatokat arra felkészült szakértők végzik, nagyban támaszkodva a számítástechnika adta automatizálási lehetőségekre. A rendszer segítségével lehetővé válik az események közel valós idejű követése, a megszerzett információk gyors feldolgozása, a felhasználók igényeinek rövid határidőn belüli kielégítése, a napi események folyamatos

figyelése, a tanácsadók és elemzők megfelelő összefoglalókkal és háttéranyagokkal való ellátása a nap 24 órájában online elérési lehetőségen keresztül.

A disszertációban kidolgozott INFF rendszer – adottságainál fogva – képes lehet a fenti kívánalmaknak megfelelni. Viszonylag olcsón, hazai alapon kivitelezhető, aránylag kevés emberi erőforrást igényel, az információigényekhez rugalmasan alkalmazkodik, miközben kapacitása technikailag gyorsan bővíthető. Meggyőződésem, hogy egy ilyen rendszer előbb-vagy utóbb bekerül a nemzetbiztonsági, nemzetvédelmi információszerezés és -szolgáltatás fegyvertárába.

Kidolgozott modellem rendszerbeállításával a Magyar Köztársaság nemzetbiztonsági és katonai védelme egy olyan nélkülözhetetlen, nyílt információ-szolgáltató rendszerrel gazdagodik, amellyel jelenleg érdemben nem rendelkezik. Bevezetésével felzárkózhatunk a fejlett országok hasonló rendszereihez. Tartós lemaradásunk ezen a téren nem indokolt, és sokáig nem védhető.

4. AZ ÉRTEKEZÉS TÉZISEI ÉS VÉGKÖVETKEZTETÉSEI, ÚJ TUDOMÁNYOS EREDMÉNYEK

1. A nyílt információk szerepe változásának tézise: Az emberi társadalom fejlődésének új korszaka az információs társadalom, amelynek a kialakulását a társadalom anyagi szellemi és termelési módjában beállt változások tették lehetővé. Az információ szerepe át- és felértékelődött az eddigi társadalmakban betöltött szerepéhez képest. Az információs társadalom talaján a társadalom gyorsuló ütemben fejlődik tovább, amelyet tudástársadalomnak nevezünk. Ebben az új formációban a nyílt információk és a tudás szerepe tovább növekszik.

Következtetés: A mai társadalmi és világpolitikai helyzetben előtérbe kerülnek a nyílt információk, amelynek okai: a hidegháború befejeződése után felgyorsult demokratizálódási folyamatok, illetve a nyílt forrású információk mennyiségi túlsúlyba kerülése, az egyéb forrásból származó információkkal szemben. A nyíltság felé való eltolódást erősíti az a felismerés, hogy az információs társadalomban az információk minősítése kevésbé nyújt előnyt a versenytársakkal szemben, mert megnöveli a költségeket. Helyette a gyorsaság kerül előtérbe. Aki gyorsabban szerzi meg és dolgozza fel az információt, az időfölénybe kerül, és elsőként aknázhatja ki a lehetőségeket. Az információszerezés nyílt forrásokból oldható meg a leggyorsabban, mert azok minden pillanatban keletkeznek és folyamatosan elérhetők. Számos esetben csak ilyen forrás áll rendelkezésünkre. Kutatásaim szerint, napjainkban a szükséges információk akár 90%-a is megszerzhető nyílt forrásokból.

2. A digitális elektronikus információk szerepe változásának tézise: Az információt nem elég megszerezni azt fel is kell tudni dolgozni. Az olyan nem elektronikus formában rendelkezésre álló információk, mint a nyomtatott anyagok, a filmek (mikrofilmek), papíralapú képek és térképek feldolgozása nehézkes, időt rabló munkafolyamat. Ehhez hasonlóan az analóg elektronikus jelek (pl. magnó és videó felvételek) feldolgozása is bonyolult, eszköz-, létszám- és időigényes feladat.

Következtetés: A digitális elektronikus információk feldolgozása – más formátumokkal összevetve – egyszerűbb, olcsóbb, gyorsabb és nagymértékben automatizálható. Ez az oka annak, hogy a nem digitális elektronikus formájú anyagok előállítására fokozatosan csökkenő tendenciát mutat (erre a nyomtatott sajtó a legjobb példa). Emellett minden országban indultak olyan projektek, amelyek a régebbi dokumentumokat, köny-

veket, újságokat, képeket, filmeket, hang és videó anyagokat átalakítják, illetve megőrzik digitális elektronikus alakban is. Ez azt jelenti, hogy az analóg rendszerek és az általuk előállított analóg információk fokozatosan és relatíve megszűnnek, átadva helyüket a digitális elektronikus információknak.

3. A számítógépes hálózatoknak az információátvitelben betöltött szerepe változásának tézise: A megszerzett, majd feldolgozott információt valamilyen módon el kell juttatni a felhasználókhöz. A nem elektronikus információterjesztés térben és időben korlátozott, a továbbítás távolságának növekedésével arányosan növekszik annak költség- és időigénye, másolás és sokszorosítás során információtartalma torzul, elvész. Az analóg elektronikus információk terjesztése könnyebb és olcsóbb, de azok is nagymértékben torzulhatnak és veszteséget szenvedhetnek az átviteli csatornában. A digitális elektronikus jelek – megfelelő eljárások alkalmazásával – mentessé tehetők az átviteli torzulásoktól, hibáktól és veszteségektől, emellett másolás és sokszorosítás során nem veszítenek a tartalmukból, minőségükből. Emiatt az információ továbbításának legolcsóbb és leghatékonyabb módja a jelek digitális elektronikus formában történő átvitele.

Következtetés: A digitális elektronikus jelek átvitelére az egyik legalkalmasabb közeg az internet – a számítógépes világháló –, mert annak kiépítettsége olyan szintet ért el, hogy rajta nagy mennyiségű felhasználóhoz (számuk már meghaladja az 1 milliárdot) lehet egyidőben, közel fénysebességgel, veszteségmentesen, az eredetivel megegyező minőségben információt eljuttatni. Ez az átviteli közeg mentes a hibáktól, a decentralizált kialakításának köszönhetően képes megszakításmentes, állandó sáv szélességű összeköttetést biztosítani és az információtovábbítás költségei függetlenek az átviteli út nagyságától.

4. Az internetnek a nyílt információk elérésben betöltött szerepe változásának tézise: Az internet sokáig csak a fejlesztésében résztvevő kutatóknak nyújtott segítséget a munkájukhoz és az ismereteik megosztásához. Ez mindaddig így maradt, amíg a számítástechnikai eszközök drágák voltak és a hálózat használata külön felkészültséget igényelt. A számítógépek árának a 80-as évek végétől kezdődő rohamos csökkenése és széles körben való elterjedése felkeltette az igényt a hálózathoz való csatlakozásra. A tömeges használatot a világháló 1991-es bevezetése és az üzleti életnek a fejlesztésre gyakorolt anyagi hatása tette lehetővé. A beáramló pénzekből elindultak a hálózat szol-

gáltatásainak, protokolljainak, átviteli sebességének és sávszélességének növekedését eredményező fejlesztések.

Következtetés: Az interneten keresztül olcsón, gyorsan lehet az információt eljuttatni a felhasználókhhoz. Ez indokolja a nyílt forrású információk terjesztésének fokozatos áttevődését erre a hálózat alapú csatornára. A nagy adatbázis-szolgáltatók már a 90-es évektől hálózati úton is elérhetővé tették anyagaikat. A hagyományos médiánál is megkezdődött az internetre való átköltözés: újságok, rádió-, televízióadások érhetők el a hálózaton keresztül. Könyvtárak adatbázisai, könyvek, képek, műholdas képek, térképek stb. mind elérhetővé válnak a világhálón keresztül is. A tendencia az, hogy az információk eleve digitális elektronikus formában születnek és egyből felkerülnek a világhálóra. Megjelennek az internetes rádió és televízió műsorok, az internetes újságírás, elektronikus könyvtárak, elektronikus kereskedelem stb. Ezek a változások azt jelentik, hogy az internet válik a legnagyobb nyílt forrású információt szolgáltató médiummá.

5. Az internet fejlesztéseinek hatása az információ átvitelében és keresésében tézise. Az internet fejlesztése – többek között – arra irányul, hogy gyorsabbá és megbízhatóbbá tegye az adatátvitelt. Ehhez az átviteli közegek, átviteli technikák és a különböző protokollok fejlesztését végzik, amelynek eredményeként terjedőben vannak a szélessávú internet kapcsolatok, amelyek új szolgáltatások bevezetését teszik lehetővé. Az interneten lévő anyagok rendezetlenek, nincsenek a keresést segítő valamilyen tartalomjegyzékkel ellátva, ezért az információk megtalálása nehézséget jelent. Az információk megtalálásában nyújtanak segítséget a keresési technikák.

Következtetés: A szélessávú átviteltechnikák elterjedése felgyorsult az utóbbi időben. Ma már lehetséges előfizetni 20 Mbit/s-os átviteli sebességű kapcsolatra, ami már lehetővé teszi olyan szolgáltatások használatát, mint az IPTV. 2010-re várható a 100 Mbit/s-os átviteli sebesség általánossá válása. A kutatói (akadémiai) hálózatokban 10 Gbit/s-os sebességet is elérnek már. Az IPv6 protokollal – egyebek mellett – biztosítható a megbízható beszédátvitel és a megfelelő számú IP cím kiosztása, így az internetre költözhet a telefonszolgáltatás (a mobilt is beleértve), valamint megvalósulhat a különböző eszközök, berendezések, készülékek távirányítása. A fejlesztések tendenciájának vizsgálata alapján az internet válik a legfontosabb információterjesztő és kommunikációs eszközzé. A keresési technikák fejlesztése során arra törekednek a kutatók, hogy olyan protokollokat, számítógépes nyelveket és szabványokat fejlesszenek ki, ami hatékonyra teszi az információ keresését, illetve az azok közötti összefüggések felisme-

rését gépi úton. A szemantikus web adta lehetőségek olyan globális infrastruktúra kialakulását teszik lehetővé, ahol könnyebb és megbízhatóbb az információk keresése és felhasználása.

6. Az internetes információszerzés szerepe felértékelődésének tézise: A számítógépes hálózat kialakulásával létrejött a kiber-tér, a számítógépek alkotta virtuális világ. Ebben a térben ugyanúgy folyik a küzdelem a versenytársak között, mint a valós világban. Az információkkal végzett műveletek nagy része áttevődik az internetre, ezért szükség van egy olyan rendszer kialakítására, amely képes az internet folyamatos figyelésére és az arról való, nagymértékben automatizált információszerzésre.

Következtetés: A fent már említett fejlesztések, valamint a felhasználók számának rohamos növekedése hatására az interneten elérhető információk már több tízezer TB (10^{12} bájtt) nagyságrendet érnek el, ezzel a legnagyobb információforrássá vált a világháló. Ez és a várható további fejlődés kényszeríti ki – véleményem szerint –, hogy az internet felhasználásával megvalósított információszerzést már külön információszerzési szakterületként kezeljük, mert az speciális szaktudást és komoly felkészültséget követel meg annak művelőitől. Javaslatom szerint ezt „internet alapú nyílt forrású felderítés”-nek (INFF) (angolul: Internet Open Source Intelligence – I-OSINT) javaslom elnevezni. Ez az új szakterület új és megnövelt információszerzési képességet jelent, amely értékes módon egészíti ki a korábbi információszerzési rendszereket.

7. Az INFF rendszermodell szükségességének tézise. Az információk mennyiségének növekedése már lehetetlenné teszi azt, hogy egy egyszerű felhasználó átlássa az internet egészét. A hatékony információszerzést ma már csak az erre felkészült szakemberek tudják megvalósítani speciálisan kialakított eszközrendszer és szoftverek alkalmazásával.

Következtetés: A kutatásaim során nem találtam olyan szakirodalmat, cikket, publikációt, amely azt tárgyalta volna, hogy hogyan kell megoldani az internetről való információszerzést rendszertechnikai szempontból. Az ismereteknek ezt a hézagát kívántam kitölteni az általam javasolt INFF rendszermodell kidolgozásával. Ezen belül elemeztem az internet hálózatra kapcsolható információkereső, -feldolgozó, -szelektáló rendszereket és azok alapján kialakítottam azt a modellt, amely szelektív, többlépcsős, rugalmas, könnyen továbbfejleszthető, valós idejű rendszer, amely alkalmas az internetről történő információszerzésre. Megítélésem szerint az általam kidolgozott rendszer-

modell (INFF modell) segítségével elfogadható költségek mellett, eredményesen lehet kiaknázni az interneten lévő információkat. A modell megvalósításával olyan előnyökhöz juthatunk, amelyekkel jelenleg nem rendelkezünk, miközben a fejlett országokban már ilyen eszközöket sikeresen alkalmaznak. Az INFF rendszer bevezetésével a döntéselőkészítők (elemzők és kidolgozók), illetve végső soron a döntéshozók olyan időelőnyre tesznek szert, és olyan tájékozottsági szintet érhetnek el, amilyennel a versenytársak nem rendelkeznek. Ezért egy ilyen rendszer megvalósítása – véleményem szerint – országos és nemzetbiztonsági érdek. A rendszer megvalósítása nem ütközik elháríthatatlan technikai vagy biztonsági akadályokba. A nyílt információk kezelése akár azt is lehetővé teszi, hogy a rendszer zárt feladatú polgári vállalat keretében működjön. Az INFF rendszer az információigényeket illetően rendkívül rugalmas, figyelmének fókuszát könnyen és gyorsan lehet új feladatokra átirányítani. A rendszer biztosítja, hogy a rendkívül nagy tömegű információk közül – értékelés után – csak azok kerüljenek feldolgozásra, amelyekre a vezetésnek igénye van. Képes az információk egyidejű, párhuzamos feldolgozására, elemzésére és értékelésére. A rendszer egyik igen fontos előnye, hogy ugyanazon információkat digitális úton azonos időben képes szétküldeni az érintett feldolgozó vagy döntési végpontok felé. A rendszer eszköz-költségkihatása a feladathoz mérten szerény, üzemeltetési költsége nem túlzott, létszámigénye korlátozott. Külön előnye, hogy hazai beszerzésekkel és kivitelezéssel megvalósítható. Az INFF rendszerek a fejlett országokban már eredményesen működnek, jelentős versenyelőnyt, tájékozottságot, döntési fölényt és megbízható nemzetbiztonságot lehetővé téve az érintett ország vezetői számára.

4.1. ÚJ TUDOMÁNYOS EREDMÉNYEIM

1. Összefoglaltam és rendszereztem, a nyílt forrású felderítésen belül, az internet alapú nyílt forrású felderítés helyével, szerepével, végzésével és rendszertechnikai megvalósításával kapcsolatos ismereteket. Ez a munka egy meglévő hiányt pótol a nyílt forrású felderítéssel kapcsolatos magyar szakirodalom területén.
2. Az internet fejlettségét, szolgáltatásait, technikáit, eszközrendszerét és fejlesztésének tendenciáit elemezve felismertem és bizonyítottam, hogy az internet alapú információszerzés eszköztára adaptálható a nyílt információszerzésre.
3. A nyílt információszerzéssel kapcsolatos eljárásokat, a megvalósításának eszközrendszerét, a végzéséhez szükséges informatikai, nyelvi, technikai, felderítő szakmai

ismereteket elemezve, bizonyítottam, hogy az internet alapú nyílt információszerzés olyan speciális felkészültséget és ismereteket követel az azt végrehajtó állománytól, ami indokolja annak új szakterületként való kezelését a nyílt információszerzésen belül.

4. Kidolgoztam egy internet alapú információszerző rendszermodellt, ami – rendszer-technikai kialakításának köszönhetően – egyaránt képes a megfelelő információforrások felkutatására, az információk automatikus gyűjtésére, azok szelektálására és feldolgozására, valamint – a kiépített hálózati kapcsolatok révén – a feldolgozott anyagoknak a felhasználókhöz való eljuttatására.

4.2. TOVÁBBI KUTATÁSI ÉS FELHASZNÁLÁSI LEHETŐSÉGEK

Az értekezésben megfogalmazott tézisek és bemutatott eredmények a nyílt forrású, azon belül az internet alapú információszerzés, folyamatosan ad új lehetőségeket a kutatásra, mivel a felgyorsult világ gyors fejlődést hoz az élet minden területén, amit mindig követni kell. A jövőbeli kutatásokat több főirányban javaslom folytatni:

- ☞ a források keresésének automatizálása;
- ☞ olyan intelligens kereső program kifejlesztése, ami képes elemezni, felismerni a vizsgált oldalak tartalmát és ez alapján eldönteni, hogy az tartalmaz-e releváns információt;
- ☞ az internet újonnan megjelenő szolgáltatásainak vizsgálata;
- ☞ a rendszer adatszerző adatbázisának minél hatékonyabb kialakítása, struktúrájának fejlesztése;
- ☞ az idegen nyelvű anyagok gépi fordítási lehetőségeinek kutatása;
- ☞ a források és az információk értékelési folyamatának automatizálása;
- ☞ a jelentések készítésének minél nagyobb mértékű gépesítése.

Az általam kidolgozott rendszermodell megfelelő alap lehet a kialakítandó rendszer megvalósításához.

Az egyetemi oktatásban bevezetendő nyílt forrású információszerzéssel kapcsolatos tananyaghoz, ajánlott irodalomként.

A kutatásaim eredményei alapján kidolgozott rendszer minden olyan szervezetenél alkalmazható, amely információszerzéssel foglalkozik. Mivel a rendszer kiépítése rugalmasan változtatható, ezért egyaránt alkalmas kisebb és nagyobb méretekben való realizálásra.

Amennyiben a számítógépeket a mai legmodernebb eszközökből valósítják meg, akkor a rendszer hosszú ideig – több évtizeden át – képes lehet kiszolgálni az információszerzési igényeket, a szükséges fejlesztési és karbantartási feladatok elvégzése mellett.

A rendszer kisebb méretekben megvalósított változata (mobil eszközökbe telepítve) alkalmas lehet a Magyar Honvédség külföldön szolgálatot teljesítő kontingenseinek információval való ellátására, valamint a hazai felderítő rendszer kiegészítésére.

Budapest, 2007. május 25.

Ferenczy Gábor Zoltán doktorjelölt

FELHASZNÁLT IRODALOM

1. **Várhegyi, István – Makkay, Imre:** Információs korszak, információs háború, biztonságkultúra, OMIKK, Budapest, 2000.;
2. **Intelligence Exploitation of the Internet**, SACLANT, Norfolk, 2002.;
3. **NATO Open Source Intelligence Handbook**, SACLANT, Norfolk, 2001.
4. **Allied Joint Intelligence, Counter Intelligence and Security Doctrine (AJP-2.0)**, NATO PfP ;
5. **Joint Publication 2-01:** Joint and National Intelligence Support to Military Operations, Department of Defense, 2004.;
6. **FM 3-13 (FM 100-6):** Information Operations: Doctrine, Tactics, Techniques, and Procedures Headquarters, Department of the Army, Washington DC. 2003.;
7. **Joint Publication 3-13:** Information Operation, Department of Defense, 2006.;
8. **Peter L. Burnett Jr.:** Information Operation, U.S. Army War College Carlisle Barracks, Pennsylvania, 2002.;
9. **Lévay, Gábor:** A nyílt források felhasználásának lehetőségei a hírszerző munkában, Felderítő Szemle, III. évfolyam 3. szám, 2004.
10. **Lévay, Gábor:** A nemzeti OSINT, Felderítő Szemle, III. évfolyam 4. szám, 2004. december;
11. **Open Source Intelligence Resources for the Intelligence Professional**, The 434th Military Intelligence Detachment, 200 Wintergreen Avenue New Haven, 1994;
12. **Lévay, Gábor:** A nyílt információs hírszerzés [OSINT] lehetőségei a katonai műveletek tervezésének támogatásában, Felderítő Szemle, IV. évfolyam 1. szám, 2005.;
13. **Steele, Robert D.:** Searching for Bin Laden: The Use of Intelligence in the War on Terror or How NOT to Spend the Taxpayers' Treasure, <http://www.oss.net/WAC> (2006. 11. 02.);
14. **Ferenczy, Gábor:** Nyílt források felhasználása a felderítésben, Bolyai Szemle, Budapest, 2000. 11. 02. különszám;
15. **NATO Open Source Intelligence Reader**, SACLANT, Norfolk, 2002.;
16. **Tompkins, Alan D.:** The Internet: Resource or Quagmire??. Military Intelligence Professional Bulletin, USAR, 1997 <http://www.fas.org/irp/agency/army/mipb/1997-1/tompkins.htm> (2006. 10. 04.) ;

17. **Operations Security Intelligence Threat Handbook:** Open Source Collection, The Interagency OPSEC Support Staff 1996,
<http://www.fas.org/irp/nsa/iOSS/threat96/part06.htm>; (2006. 09. 02.);
18. **Robert, David Steele:** Open Source Intelligence (OSINT), Strategic Intelligence on OSINT, version 2.4 2006.;
19. **Stalder, Felix – Hirsh, Jesse:** Open Source Intelligence. First Monday 7. kötet 6. szám, 2002. június;
20. **Tanenbaum, Andrew S.:** Számítógép-hálózatok, Panem Kft., Budapest, 2003;
21. **Kis, Balázs:** Winternet, Szak Kiadó Kft., Budapest, 2001.;
22. **Rády, Ferenc:** A központi szolgáltatások jelene és jövője az új törvény szellemében, Könyv, Könyvtár, Könyvtáros, 2. évf. 9. sz. 1998. szeptember;
23. **Hideg, Róbert:** Bizniszkémek Üzleti hírszerzés törvényesen, Piac és Profit 2002. 10. szám;
24. **Mikulás, Gábor:** Az információ- és tudás piac új szereplője: az információbróker, Marketing & Menedzsment, 2003.;
25. **Mikulás, Gábor:** Adatok, információk és a tudás bázisa, Figyelő, 2002. október 31.;
26. **Zimányi, Magdolna:** A Magyar Elektronikus Könyvtár, Magyar Tudomány, 2001/2;
27. **Gillin, Paul:** The coming collapse and rebirth of newspaper journalism, BtoB Magazine, 2006. 12. 11.;
28. **Kotsis, Márk:** Sajtó a hypertexten. Magyar on-line lapok vizsgálata Kodolányi Füzetek, 3. Kodolányi János Főiskola Székesfehérvár, 1999.;
29. **Prim Online:** Szakmai konferencia a tudományos publikációk elektronikus hozzáférhetőségéről, 2006. november 6 <http://hirek.prim.hu/cikk/55665/> (2007. 01. 11.);
30. **Dr. Galántai, Zoltán:** Web 2.0 A web mint platform,
<http://galantai.inno.bme.hu/web2.html> (2006. 10. 01.);
31. **Csordás, Attila:** A feltölthető világ - Web 2.0, Magyar Narancs Online,
<http://mancs.hu/index.php?gcPage=/public/hirek/hir.php&id=12407> (2006. 10. 02.);
32. **O'Reilly, Tim:** What Is Web 2.0, O'Reilly Media, 2005. 09. 30.
<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html> (2006. 10. 02.);
33. **Járolai, József:** Web 2.0, Innen.hu Bt, Budapest, 2006. 1-2. old.
<http://webni.innen.hu/Web20> (2006. 10. 03.);

34. **Ferenczy, Gábor:** Az Internet mint a felderítés adatforrása, Bolyai Szemle, Budapest, 2001. X. évf. 1. szám;
35. **Lyman, Peter – Varion, Hal:** HOW MUCH INFORMATION 2003?, University of California, Los Angeles 2003;
36. **Stagel, Stef:** Next Generation Internet in Europe, Querverbindungen, Agentur für Öffentlichkeitsarbeit GmbH, Berlin, 1999.;
37. **Gulli, Antonio – Signorini, Alessio:** The Indexable Web is more than 11.5 billion pages, <http://www.cs.uiowa.edu/~asignori/web-size/> (2006. 11. 22.);
38. **Jagannathan, Prasanna Kumar:** Global Next Generation Internet Initiatives, 1999. <http://www.cs.wustl.edu/~jain/cis788-99/ftp/testbeds/index.html> (2006. 11. 21.);
39. **Siwamogsatham, Siwaruk:** 10 Gigabit Ethernet, Washington University, Washington, 1999.;
40. **Seetharaman, Srinivasan:** IP over DWDM, Washington University, Washington 1999.;
41. **Song, Chuck:** vBNS - Very High Bandwidth Backbone Network Service for US Research and Education and Its Role in the Next Generation Internet, MCI WorldCom, 1999.;
42. **Braun, David:** Two Project Lay a Fast Track To The Future, TechWeb, 1997. <http://www.techweb.com/wire/news/aug/0825sequel2.html> (2006. 12. 04.);
43. **Narayan, Vishy:** NREN /NGIX –W, NASA Ames Research Center Jan 26, 1999.;
44. **Kratz – Ackermanb - Hanss, Corbato:** NGI and Internet2: Accelerating the Creation of Tomorrow’s Internet IOS Press, Amsterdam 2001.;
45. **Internet2 Consortium:** Abilene Network, 2005. február <http://www.internet2.edu/pubs/200502-IS-AN.pdf> (2006. 11. 23.);
46. **DoD HPC Modernization Program:** Tools for Discovery Annual Report, 2002., http://www.hpcmo.hpc.mil/Htdocs/ANNUALREPORT/2002_hpcmp_annual_report.pdf (2006. 11. 19.);
47. **Teitelbaum, Ben:** Internet2 QoS Testbed QBone, Internet2 Member Meeting 1998. szeptember 28., <http://qos.internet2.edu/wg/calendar/Sep98GeneralMtg/Teitelbaum.pdf>;
48. **Schatt, Stan:** Hogyan működik az ATM? Panem-McGraw-Hill, Budapest, 1998.,
49. **Farkas, Károly:** IPv6 – A jövő Internet protokollja?, Computer Engineering and Networks Laboratory (TIK), Zürich, 2004.;

-
50. **Kumm**: Az IPv6, PROG.HU - On-line Fejlesztői Portál, 2001.10.08.,
<http://www.prog.hu/cikkek/230/Az+IPv6.html> (2006. 12. 07.);
 51. **Gábor, Tamás, - Temesvári, András**: IPv6, <http://ipv6.ircnet.hu/index.php>;
 52. **Ferenczy, Gábor**: Az új generációs Internet, Bolyai Szemle, Budapest, 2002. XI. évf. 3. szám;
 53. **Ungváry, Rudolf**: A tartalom szerinti információkeresés az Interneten 1-2. Tudományos és Műszaki Tájékoztató, 47. évf. 2000. 1-2. szám;
 54. **Tarr, Bence**: Keresés az Interneten, Panem Kft., Budapest, 2001.;
 55. **Ferenczy, Gábor**: Internetes keresési technikák, Bolyai Szemle, Budapest, Különszám 2001. 11. 07.;
 56. **Járol, József**: Kereső adatbázisok mérete, innen.hu Bt, Budapest, 2005.;
 57. **Surányi, Gábor Mihály**: A mély háló (Deep Web), BME Informatikai Központ Axelero Internet Rt. 2003.;
 58. **K. Princz, Mária**: Tájékoztató a weben, Debreceni Egyetem Műszaki Főiskolai Kar, Debrecen, 2005.;
 59. **Cohen, Laura**: Internet Tutorials The Deep Web, University at Albany, Albany, 2006.;
 60. **Dr. Rutkovszky, Edéné – Rutkovszky, Ádám**: A láthatatlan web keresése, Debreceni Egyetem TTK, <http://nws.iif.hu/ncd2003/docs/ehu/EHU-61.htm> (2006. 12. 19.);
 61. **Sherman, Chris and Price, Gary**: The Invisible Web. Uncovering Information Sources Search Engines Can't See, Independent Publishers Group, Los Angeles 2001.;
 62. **Raghavan, Sriram –Garcia-Molina, Hector**: Crawling the HiddenWeb, Computer Science Department Stanford University, Stanford, 2001.;
 63. **Berners-Lee, Tim**: Semantic Web Road map, 1998.
<http://www.w3.org/DesignIssues/Semantic.html>, (2007. 01. 02.);
 64. **Varga, László Zsolt**: A web-alkalmazások új technológiái, MTA SZTAKI, Budapest 2004.;
 65. **Szerey, Péter – Lukácsy Gergely – Benkő Tamás**: A szemantikus világháló elmélete és gyakorlata, Typotex KFT, 2005.;
 66. **Herman, Iván**: Szemantikus Web: Bevezető W3C, Budapest, 2002.;

-
67. **Fülöp, Csaba – Kovács, László – Micsik, András:** A metaadatsémák és a szemantikus web: egységesítés és specializáció a metaadatok világában, Tudományos és Műszaki Tájékoztatás 51. évf. 7. szám 2004.;
68. **Manola, Frank – Miller, Eric:** Az RDF bevezető tankönyve, W3C Ajánlás, 2004. február 10., <http://www.w3c.hu/forditasok/RDF/REC-rdf-primer-20040210.html> (2006. 10.29.);
69. **Brickley, Dan:** Az RDF Szókészlet Leíró Nyelv 1.0: RDF Séma, W3C Ajánlás, 2004. február 10., <http://www.w3c.hu/forditasok/RDF/REC-rdf-schema-20040210.html>, (2006. 10.29.);
70. **McGuinness, Deborah L. - van Harmelen, Frank:** Az OWL Web Ontológia Nyelv – Áttekintés W3C Ajánlás, 2004. február 10., <http://www.w3c.hu/forditasok/OWL/REC-owl-features-20040210.html>, (2006. 10.30.);
71. **Ferenczy, Gábor:** A jövő felderítő eszköze a számítógép, Bolyai Szemle, Budapest, 2003. XII: évf. 4. szám;
72. **Dr. Haig, Zsolt:** Számítógép-hálózati hadviselés rendszere az információs műveletekben, Bolyai Szemle, 2006. XV. évf. 1. szám, Budapest;
73. **Dr. Várhegyi, István – Dr. Ványa László:** Hadviselés az információs hadszíntéren. Trendek, új eljárások és eszközök, Bolyai Szemle, 2006. XV. évf. 1. szám, Budapest;
74. **Dr. Várhegyi, István – Sallai József:** Információs műveletek, vezetési hadviselés I-II., BJKMF főiskolai jegyzet, Budapest, 1998.;
75. **Pulai, András – Sziklássy, Fábián – Tóth, Péter, Udvarosi, H. Vilmos:** Véd magad az Interneten Biztonságtechnika a számítógépes hálózatokon, Kossuth Kiadó, 1997.;
76. **ProfQki:** Hardveres és szoftveres tűzfalak, PC Fórum, 2004.05.24. <http://pcforum.hu/cikkek/109/Hardveres+es+szoftveres+tuzfalak.html>, (2006. 09.29.);
77. **Gyimesi Judit:** Elosztott behatolásérzékelő rendszer lehetőségei, alkalmazása, BME-TMIT, 2003.;
78. **ProfQki:** Spyware, malware, backdoor, firewall, PC Fórum, 2004.05.04. <http://pcforum.hu/cikkek/107/Spyware+malware+backdoor+firewall.html>, (2006. 09.28.);

-
79. **Antrophus**: Főszerepben a biztonság 1. rész: tűzfalak, C-Enter, 2004. január 5. www.c-enter.hu/center/0117326.html, (2006. 10. 01.);
 80. **Dr. Kovács, László**: Terrorizmus a digitális hadszíntéren, Bolyai Szemle, 2006. XV. évf. 1. szám, Budapest;
 81. **Kyas, Othmar**: Számítógépes hálózatok biztonságtechnikája, Kossuth Kiadó, 2000.;
 82. **Avantgarde**: Time To Live On Network, <http://www.avantgarde.com/xxxxttln.pdf>, (2006. 10. 29.);
 83. **Iasi, Anthony F. – Bremner, Lynn M. – Servati, Al**: Az intranet bibliája I.-II. Kossuth Könyvkiadó, 1997;
 84. **Internet Systems Consortium (ISC)**: Internet Domain Survey, <http://www.isc.org/index.pl>, (2006. 11. 29.);
 85. **NATO Intelligence Exploitation of the Internet**, SAACLANT, 2002.;
 86. **Lévay, Gábor**: Internetes weboldalak elemzése, Felderítő Szemle, V. évfolyam 1. szám, 2006. március;
 87. **Vinter, József**: Nyílt forrásból származó hírszerzési információk feldolgozási folyamata, ZMNE diplomamunka, 2006. július;
 88. **Zakon, Robert H'obbes'I**: Hobbes' Internet Timeline v8.2/Growth, <http://www.zakon.org/robert/internet/timeline/#Growth> (2006. 11. 09.);
 89. **Heti Világgazdaság**: Az internetezők száma meghaladta az egymilliárdot, 2006. május 15., <http://hvg.hu/Tudomany/20060515internet.aspx> (2006. 11. 17.);
 90. **MiMi.hu**, <http://www.mimi.hu/index.html>;
 91. **PC Fórum**: Szótár, <http://pcforum.hu/szotar/> ;
 92. **Internet szótár**, <http://www.hqsoft.hu/szotar/>;
 93. **Új Magyar Évezred**: Internet értelmező kyszótár, <http://home.iae.nl/users/nickl/mikrosz.html> ;
 94. **whatis?com**: Computer Dictionary, <http://whatis.techtarget.com/whome/0,289825,sid9,00.html>;
 95. **Wikipedia**, <http://www.wikipedia.org/>.

AZ ÉRTEKEZÉSEL KAPCSOLATOS PUBLIKÁCIÓIM

1. **Ferenczy, G.** Támadások az Interneten keresztül I. Vírusok, ZMNE Fórum, Budapest, 1999. III. évf. 11-12. szám (15-18);
2. **Ferenczy, G.** Támadások az Interneten keresztül II. Hálózati támadások, támadók és védelem, ZMNE Fórum, Budapest, 2000. IV. évf. 1. szám (15-18);
3. **Ferenczy, G.** Támadások az Interneten keresztül III. Adatvédelem, ZMNE Fórum, Budapest, 2000. IV. évf. 2. szám (13-16);
4. **Pándi, E. – Ferenczy G.** A kormányzati és közigazgatási levelezőrendszer minősített időszak alkalmazásának lehetősége, Bolyai Szemle, Budapest, 2000. IX. évf. 3. szám (53-70);
5. **Ferenczy G.** Network Security Problems, Repüléstudományi Közlemények, Szolnok, 2000. XII. évf. 29. szám (101-114);
6. **Ferenczy G.** Internet Security Problems, International Conference on Advanced Military Technology In The XXI. Century – Concepts Of New Intelligence And Electronic Warfare Systems, Budapest, 200. 05. 16-17. (107-120);
7. **Ferenczy G.** Nyílt források felhasználása a felderítésben, Bolyai Szemle, Budapest, 2000. 11. 02. különszám (117-130);
8. **Ferenczy, G.** Az Internet mint a felderítés adatforrása, Bolyai Szemle, Budapest, 2001. X. évf. 1. szám (72-89);
9. **Ferenczy, G.** Az Internet mint a felderítés új adatforrása, Repüléstudományi Közlemények, Szolnok, Különszám I. 2001. 04. 21. (177-184);
10. **Ferenczy, G.** Internetes keresési technikák, Bolyai Szemle, Budapest, Különszám 2001. 11. 07. (123-138);
11. **Ferenczy, G.** Az új generációs Internet, Bolyai Szemle, Budapest, 2002. XI. évf. 3. szám (91-102);
12. **Ferenczy, G.** A protokollváltás szükségessége az Interneten, Repüléstudományi Közlemények, Szolnok, Különszám 2. 2002. 04. 12-13-14. (173-178);
13. **Ferenczy, G.** A jövő felderítő eszköze a számítógép, Bolyai Szemle, Budapest, 2003. XII: évf. 4. szám (81-102);
14. **Ferenczy, G.** A XXI század felderítő eszköze a számítógép, Tudományos konferencia, Repüléstudományi Közlemények Elektronikus különszáma, Szolnok, 2003. 04. 04. (CD mellékleten);

15. **Ferenczy, G.** A jövő felderítő eszköze a számítógép, Bolyai Szemle elektronikus különszám, 2003. 10. 28-29. (CD mellékleten);
16. **Ferenczy, G.** Az internet új fejlődési iránya: Web 2.0, előkészület alatt;

KONFERENCIA ELŐADÁSOK

1. **Ferenczy, G.** Network Security Problems, Megújuló magyar repülőszakember képzés, tudományos konferencia, Szolnok, 2000. 04. 15.;
2. **Ferenczy, G.** Internet Security Problems, Korszerű katonai technológiák a XXI. században – Az új felderítő, elektronikai hadviselési rendszerek koncepciói, nemzetközi konferencia, Budapest, 2000. 05. 17.;
3. **Ferenczy, G.** Nyílt források felhasználása a felderítésben, Fiatal tudósok konferenciája, Budapest, 2000. 11. 02.;
4. **Ferenczy, G.** Az Internet, mint a felderítés új adatforrása, XX. század haditechnikai forradalmának hatása a XXI. század katonai repülésére, tudományos konferencia, Szolnok, 2001. 04. 21.;
5. **Ferenczy, G.** Internetes keresési technikák, Új kihívások a hadtudomány és a kapcsolódó tudományok területén a XXI. század hajnalán, tudományos konferencia, Budapest, 2001. 11. 07.;
6. **Ferenczy, G.** A protokollváltás szükségessége az Interneten, Future Aviation Technologies, nemzetközi szimpózium, Szolnok, 2002. 04. 13.;
7. **Ferenczy, G.** A protokollok fejlődése az Interneten, „Új kihívások a hadtudomány és a kapcsolódó tudományok területén a XXI. század hajnalán”, tudományos konferencia, 2002. 11. 04.;
8. **Ferenczy, G.** A XXI. század felderítő eszköze a számítógép, 100 éves a géprepülés „A katonai rendszerek repülőgépei, a katonai repülőgépek rendszerei”, tudományos konferencia, Szolnok, 2003. 04. 04.;
9. **Ferenczy, G.** A jövő felderítő eszköze a számítógép „New Challenges In The Field of Military Sciences 2003” nemzetközi tudományos konferencia, 2003. 10. 28-29.

AZ ÉRTEKEZÉSEL NEM KAPCSOLATOS PUBLIKÁCIÓIM

1. **Sallai J. – Szűcs, P. – Ferenczy, G. – Balogh, K.** Műholdas rádiótelefonია I. Bolyai Szemle, 1998. 3. szám (69-83);
2. **Pándi Erik – Ferenczy G.** A hazai zártcélú rádiótávközlő hálózatok egységesítésének technikai lehetősége Bolyai Szemle, Budapest, 2000. IX. évf. 2. szám (54-70);
3. **Ferenczy, G. – Oláh, T.** Rövidhullámú vételtechnikai eszközök, főiskolai jegyzet, Budapest, 1995.;
4. **Ferenczy, G. – Szűcs, P. – Balogh, K.** Rádiólokáció alapjai, főiskolai jegyzet, Budapest, 1998.;
5. **Ferenczy, G. – Sallai, J.** Hírközlés, főiskolai jegyzet, Budapest, 1999.;
6. **Ferenczy, G. – Poroszlai, Á.:** Rádió-vételtechnika, főiskolai jegyzet, Budapest, 2002.

1. SZÁMÚ MELLÉKLET

RÖVIDÍTÉSEK TÁBLÁZATA

Rövidítés	Feloldás	Jelentés
ABILENE	Advanced Networking for Leading-edge Research and Education	Fejlett hálózati szolgáltatás vezető kutatóknak és oktatóknak
ADSL	Asymmetrical Digital Subscriber's Line	aszimmetrikus digitális előfizetői vonal
ALIWEB	Archie Like Indexing for the Web	Archie indexelő a világháló részére
ARPANET	Advanced Research Project Agency Network	Korszerű Kutatásokat Tervező Intézet Hálózata
ATM	Asynchronous Transfer Mode	Aszinkron átviteli mód
AVI	Audio/Video Interleaved	időosztásos hang/videó
BBC	British Broadcasting Corporation	Brit Műsorszóró Vállalat
BLDSC	British Library Document Supply Centre	Brit Nemzeti Könyvtár Dokumentum Ellátó Központ
CBQ	Class Based Queuing	osztály alapú sorba állítás
CERN	Conseil Européen pour la Recherche Nucléaire	Európai Atommagkutató Tanács
CIA	Central Intelligence Agency	Központi Hírszerző Ügynökség
CIDR	Classless Inter-Domain Routing	osztály nélküli domain közti útválasztás
CIS	CompuServe Information Service	CompuServe Információs Szolgálat
COMINT	Communication Intelligence	rádiófelderítés
CNN	Cable News Network	Kábel hírhálózat
DARPA	Defense Advanced Research Projects Agency	Fejlett Kutatások Védelmi Ügynöksége
DNS	Domain Name System	domain név szerviz
DoD	Department of Defense	Védelmi Minisztérium
DOE	Department of Energy	Energia Minisztérium
DoS	Denial-of-Service	szolgáltatás-leállítás
DREN	Defense Research & Engineering Network	Védelmi Kutató és Tervező Hálózat
DSC	Directory of Social Change	Társadalmi Változások Könyvtára
DVD	Digital Versatile Disc	digitális sokoldalú korong
DWDM	Dense Wavelength Division Multiplexing	sűrű hullámhossz osztásos multiplexálás
ELINT	Electronic Intelligence	rádiótechnikai felderítés
ESNET	Energy Sciences Network	Energia Tudósok Hálózata
FAS	Federation of American Scientists	Amerikai Tudósok Szövetsége
FDDI	Fiber Distributed Data Interface	üvegszál as adatelosztó interfész
FTP	File Transfer Protocol	fájl átviteli protokoll
GB	GigaByte	Gigabájt
GIF	Graphic Interchange Format	grafikus átalakítási alak
GRE	Generic Routing Encapsulation	általános beágyazott útválasztás
GW	Gateway	átjáró
HiWE	Hidden Web Exposer	mély web kereső
HPC	High Performance Computing	Nagyteljesítményű Számítástechnika
HTML	HyperText Markup Language	hiperszöveg jelölőnyelv
HTTP	Hypertext Transfer Protocol	hiperszöveg átviteli protokoll
HTTPS	Hypertext Transfer Protocol Secure	biztonságos hiperszöveg átviteli protokoll
HUMINT	Human Intelligence –	ügynöki felderítés
IBM	International Business Machines	Nemzetközi Üzleti Gépek
IDS	Intrusion Detection System	behatolás-érzékelő rendszer
IMINT	Imagery Intelligence	képi felderítés
INIST	L'Institut de l'Information Scientifique et Technique	Tudományos és Technikai Információk Intézete
I2DSI	Internet2 Distributed Storage Initiative	Internet2 Szétosztott Tárolás kezdeményezés

Rövidítés	Feloldás	Jelentés
I2DVN	Internet2 Digital Video Network	Internet2 Digitális Video Hálózat
IO	Information Operation	információs műveletek
IPng	Next Generation Internet Protocol	új generációs internet protokoll
IPSec	Internet Protocol Security	internet protokoll biztonság
IPTV	Internet Protocol Television	internet protokoll televízió
IPv4	Internet Protocol version 4	4-es verziószámú internet protokoll
IPv6	Internet Protocol version 6	6-os verziószámú internet protokoll
ISC	Internet Systems Consortium	Internet Rendszer Társaság
ISO	International Standards Organization	Nemzetközi Szabványügyi Szervezet
ISP	Internet Service Provider	Internet szolgáltató
JPEG	Joint Photographic Experts Group	egyesített fénykép szakértői csoport
ITIRC	IBM Technical Information Retrieval Center	IBM Technikai Információ-visszakereső Központ
LAN	Local Area Network	helyi hálózat
LCD	Liquid Crystal Display	folyadékkristályos kijelző
MASINT	Measurements and Signatures Intelligence	technikai jelfelderítés
MB	MegaByte	Megabájt
MIT	Massachusetts Institute of Technology	Technikai Intézet Massachusetts
MPEG	Moving Picture Experts Group	mozgóképszakértő csoport
MPLS	Multi Protocol Label Switching	Sokprotokollós címkekapcsolás
NASA	National Aeronautics and Space Administration	Nemzeti Légügyi és Űrhajózási Hivatala
NAT	Network Access Translation	hálózati címfordítás
NATO	North Atlantic Treaty Organization	Észak-atlanti Szerződés Szervezete
NCP	Network Control Protocol	hálózatvezérlési protokoll
NGI	Next Generation Internet	újgenerációs internet
NGII	Next Generation Internet Initiative	újgenerációs internet kezdeményezés
NGO	non- governmental organisation	nem-kormányzati szervezet
NIH	National Institutes of Health	Nemzeti Egészségügyi Intézet
NIIF		Nemzeti Információs Infrastruktúrafejlesztési Program
NISN	NASA Integrated Service Network	NASA integrált szolgáltatású hálózata
NIST	National Institute of Standards and Technology	Nemzeti Szabványügyi Intézet
NREN	NASA Research & Education Network	NASA Kutatói és Oktatói Hálózat
NSA	National Security Agency	Nemzetbiztonsági Ügynökség
NSF	National Science Foundation	Nemzeti Tudományos Alapítvány
NSFNET	National Science Foundation Network	Nemzeti Tudományos Alapítvány Hálózata
NTIS	National Technical Information Service	Nemzeti Technikai Információ Szolgálat
OC	Optical Carrier	optikai vivő
OS	Open Source	nyílt források
OSD	Open Source Data	nyílt forrású adat
OSI	Open Source Information	nyílt információforrások
OSIF	Open Source Information	nyílt forrású információ
OSINT	Open Source Intelligence	nyílt forrású felderítés/információ-szerzés
OSSI	Open Sources Solution Incorporation	Nyílt Forrású Megoldások Egyesülete
OWL	Web Ontology Language	web ontológia nyelv
QBONE	Quality Bone	minőségi gerinchálózat
QoS	Quality of Service	minőségi szolgáltatás
PDF	Portable Document Format	hordozható dokumentum formátum
RDF	Resource Description Framework	forrásleíró keretrendszer
RDFS	Resource Description Framework Schema	forrásleíró keretrendszer séma
ROM	Read Only Memory	csak olvasható memória
RSS	Really Simple Syndication	nagyon egyszerű intézés
RTF	Rich Text Format	gazdag szövegfórmátum
SIGINT	Signal Intelligence	rádióelektronikai felderítés

Rövidítés	Feloldás	Jelentés
SQL	Structured Query Language	Strukturált lekérdező nyelv
ST	Stream Protocol	folymat protokoll
STP	Shielded Twisted Pair	árnyékolt sodrott érpár
TB	TeraByte	terabájt (10^{12} bájt)
TCP/IP	Transmission Control Protocol/Internet Protocol –	átvitel vezérlési protokoll/internet protokoll
TDL	top-level domain	felső szintű domain
TOC	Table of Content	tartalomjegyzék
TOR	The Onion Router	Hagyma útválasztó
TXT	Text	szöveg formátum
URI	Uniform Resource Identifier	egységes erőforrás azonosító
URL	Uniform Resource Locator	egységes forráshatározó
UTF	Unicode Transformation Format	egységes kódtábla alak
UTP	Unshielded Twisted Pair	árnyékoltatlan sodrott érpár
vBNS	Very high speed Backbone Network Service	szupergyors gerinchálózati szolgáltatás
VoIP	Voice Over Internet Protocol	Beszédátvitel internet protokoll felett
VPN	Virtual Private Network	virtuális magánhálózat
W3C	World Wide Web Consortium	világháló konzorcium
WAIS	Wide Area Information Servers	Nagykiterjedésű információs szerverek
WAN	Wide Area Network	nagytávolságú hálózat
Wandex	World Wide Web Wanderer	világháló vándora
WIDE	Writing In Digital Environments	Digitális környezetben írni
WLAN	Wireless Local Area Network	vezeték nélküli helyi hálózat
WWW	World Wide Web	világháló
XML	eXtensible Markup Language	bővíthető jelölőnyelv

2. SZÁMÚ MELLÉKLET

IDEGEN SZAVAK TÁBLÁZATA

Idegen szó	Fordítás
anomaly based	anomália alapú
anonymisation	anonimizálás: névtelen böngészés
anti-phishing	adathalászat megakadályozó
antiviral	vírusirtó
antitrojan	trójai irtó
anycast address	bárki cím
authentication header	hitelesítő fejléc
backbone	gerinchálózat
blogsphere	blogtér
browser	böngésző
browsing server	témakatalógus
contributor	közreműködő
cookie	süti
coverage	lefedettség
current information	
cyberspace	kiber-tér
deep	mély
description	rövid tartalom
deszkriptor	a szemantikai információkereső nyelvek lexikai egységei
default	alapértelmezett
dial-up	betárcsázás
directory site	könyvtár oldal
domain	címtartomány
ebook	elektronikus könyv
educational	oktatói
encapsulating security payload	elzárt biztonsági csomag
enumerative	enumeratív: felsoroló, csak mechanikusan rendezett
fake	hamis
feature	jellemző
feed reader	RSS olvasó program
firewall	tűzfal
flow	folyam (adatfolyam)
government	kormány, kormányzat
gray literature	szürke irodalom
hacker	számítógépes betörő
hidden	rejtett
hypertext	hiperszöveg (egyes részei egymással tetszőlegesen összekapcsolható adatbázis)
home agent	honi ügynök, ágens
home network	honos hálózat
identifier	azonosító
information warrior	információs harcos
informetria	az információkeresés adat- és szövegbányászat, az információáramlás mennyiségi szempontú elemzését ötvöző tudományág
international	nemzetközi
issues management	a társadalmat vagy egy részét foglalkoztató és a szervezetet közvetlenül érintő makrokérdések kezelése, illetve ehhez kapcsolódó megelőző szervezeti viszony és kommunikáció szervezése
just in case	minden esetre
just in time	éppen időben
kernel	rendszermag
language	nyelv

Idegen szó	Fordítás
metadata	metaadat (adatbázisok ábrázolási és szervezési formáit leíró adathalmazok)
mission information	műveleti információ
multicast address	csoport cím
newsgroup	hírcsoport
netmask	cím-maszk (a címzéseknél a hálózat és hoszt címének szétválasztására használják)
news	hírek
ontológia	egy adott területre jellemző kategóriákat, illetve a köztük fennálló kapcsolatokat írja le
opaque web	nem átlátható világháló
optimalizáció	beállítások hozzáigazítása a konkrét működési környezethez és jellemzőkhöz a lehetséges legnagyobb teljesítmény elérése érdekében
organization	szervezet
ping	Segédprogram az internetre csatlakozó gép bekapcsolt állapotának lekérdezésére
pingomatic	blogkeresőkre automatikus bejelentkezés
pixel	képpont
plug and play	önmagától működő
plugin	bővítmény
precision	pontosság (keresőmotoré)
private web	magán web
proprietary web	szabadalmaztatott web
protocol	protokoll (szabályrendszer)
proxy	Speciális tűzfal-típus, amely szétválasztja a külső és a védett hálózat közötti kommunikációt.
publisher	kiadó
puzzle	kirakós játék
recall	találati arány (keresőmotoré)
relation	kapcsolat
right	jog
router	útvonalválasztó
rule based	szabály alapú
search engine	kereső motor
semantic	szemantika (a számítógépes jelölési rendszerek és az azok által jelölt fogalmak, tárgyak közötti kapcsolat).
spam	kéretlen reklámlevél
spider	pók (kereső program)
spyware	kémprogram
subject	téma, témakör
subnet	alhálózat
surface	felszín, felület
surface web	látható világháló
szemiotika	jeltudomány
tezaurusz	szótár, amelyben az egyes szavak, mint fogalmak az egymáshoz viszonyított és jelzett értékeik (alá-, fölérendeltség stb.) szerint vannak felsorolva.
title	cím
trace route	nyomkövetés
truly invisible web	igazán láthatatlan web
unicast address	egyedi cím
video on demand	igény szerinti videó lehívás (egy videoszerverről tetszőleges időpontban lehívható, digitális kódolású TV program)
web	háló
webmetria	az interneten való információkeresés, adat- és szövegbányászat, információáramlás mennyiségi szempontú elemzését ötvöz tudományág
weblog	webes napló

3. SZÁMÚ MELLÉKLET

A DOMAIN NEVEK LISTÁJA AZ ISO 3166 SZERINT (ANGOL ORSZÁGNEVEKKEL)

Betű	Kód	Ország	Kód	Ország	Kód	Ország
A	.ac	Ascension Island	.ad	Andorra	.ae	United Arab Emirates
	.af	Afghanistan	.ag	Antigua and Barbuda	.ai	Anguilla
	.al	Albania	.am	Armenia	.an	Netherlands Antilles
	.ao	Angola	.aq	Antarctica	.ar	Argentina
	.as	American Samoa	.at	Austria	.au	Australia
	.aw	Aruba	.az	Azerbaijan		
B	.ba	Bosnia and Herzegovina	.bb	Barbados	.bd	Bangladesh
	.be	Belgium	.bf	Burkina Faso	.bg	Bulgaria
	.bh	Bahrain	.bi	Burundi	.bj	Benin
	.bm	Bermuda	.bn	Brunei Darussalam	.bo	Bolivia
	.br	Brazil	.bs	Bahamas	.bt	Bhutan
	.bv	Bouvet Island	.bw	Botswana	.by	Belarus
	.bz	Belize				
C	.ca	Canada	.cc	Cocos (Keeling) Islands	.cd	Congo, Democratic republic of the (former Zaire)
	.cf	Central African Republic	.cg	Congo, Republic of	.ch	Switzerland
	.ci	Côte d'Ivoire	.ck	Cook Islands	.cl	Chile
	.cm	Cameroon	.cn	China	.co	Colombia
	.cr	Costa Rica	.cs	Czechoslovakia (former - non-existing)	.cu	Cuba
	.cv	Cape Verde	.cx	Christmas Island	.cy	Cyprus
	.cz	Czech Republic				
D	.de	Germany	.dj	Djibouti	.dk	Denmark
	.dm	Dominica	.do	Dominican Republic	.dz	Algeria
E	.ec	Ecuador	.ee	Estonia	.eg	Egypt
	.eh	Western Sahara	.er	Eritrea	.es	Spain
	.et	Ethiopia	.eu	European Union		
F	.fi	Finland	.fj	Fiji	.fk	Falkland Islands
	.fm	Micronesia	.fo	Faroe Islands	.fr	France
G	.ga	Gabon	.gb	United Kingdom	.gd	Grenada
	.ge	Georgia	.gf	French Guiana	.gg	Guernsey
	.gh	Ghana	.gi	Gibraltar	.gl	Greenland
	.gm	Gambia	.gn	Guinea	.gp	Guadeloupe
	.gq	Equatorial Guinea	.gr	Greece	.gs	South Georgia and the South Sandwich Islands
	.gt	Guatemala	.gu	Guam	.gw	Guinea-Bissau
	.gy	Guyana				
H	.hk	Hong Kong	.hm	Heard and McDonald Islands	.hn	Honduras
	.hr	Croatia	.ht	Haiti	.hu	Hungary
I	.id	Indonesia	.ie	Ireland	.il	Israel
	.im	Isle of Man	.in	India	.io	British Indian Ocean Territory
	.iq	Iraq	.ir	Iran	.is	Iceland
	.it	Italy				
J	.je	Jersey	.jm	Jamaica	.jo	Jordan
	.jp	Japan				
K	.ke	Kenya	.kg	Kyrgyzstan	.kh	Cambodia
	.ki	Kiribati	.km	Comoros	.kn	Saint Kitts and Nevis
	.kp	Korea, Democratic People's Republic of	.kr	Korea, Republic of	.kw	Kuwait
	.ky	Cayman Islands	.kz	Kazakhstan		

Betű	Kód	Ország	Kód	Ország	Kód	Ország
L	.la	Lao People's Democratic Republic	.lb	Lebanon	.lc	Saint Lucia
	.li	Liechtenstein	.lk	Sri Lanka	.lr	Liberia
	.ls	Lesotho	.lt	Lithuania	.lu	Luxembourg
	.lv	Latvia	.ly	Libyan Arab Jamahiriya		
M	.ma	Morocco	.mc	Monaco	.md	Moldova
	.mg	Madagascar	.mh	Marshall Islands	.mk	Macedonia
	.ml	Mali	.mm	Myanmar	.mn	Mongolia
	.mo	Macau	.mp	Northern Mariana Islands	.mq	Martinique
	.mr	Mauritania	.ms	Montserrat	.mt	Malta
	.mu	Mauritius	.mv	Maldives	.mw	Malawi
	.mx	Mexico	.my	Malaysia	.mz	Mozambique
N	.na	Namibia	.nc	New Caledonia	.ne	Niger
	.nf	Norfolk Island	.ng	Nigeria	.ni	Nicaragua
	.nl	The Netherlands	.no	Norway	.np	Nepal
	.nr	Nauru	.nu	Niue	.nz	New Zealand
O	.om	Oman				
P	.pa	Panama	.pe	Peru	.pf	French Polynesia
	.pg	Papua New Guinea	.ph	Philippines	.pk	Pakistan
	.pl	Poland	.pm	St. Pierre and Miquelon	.pn	Pitcairn
	.pr	Puerto Rico	.ps	Palestine	.pt	Portugal
	.pw	Palau	.py	Paraguay		
Q	.qa	Qatar				
R	.re	Reunion	.ro	Romania	.ru	Russia
	.rw	Rwanda				
S	.sa	Saudi Arabia	.sb	Solomon Islands	.sc	Seychelles
	.sd	Sudan	.se	Sweden	.sg	Singapore
	.sh	St. Helena	.si	Slovenia	.sj	Svalbard and Jan Mayen Islands
	.sk	Slovakia	.sl	Sierra Leone	.sm	San Marino
	.sn	Senegal	.so	Somalia	.sr	Surinam
	.st	Sao Tome and Principe	.su	USSR (former)	.sv	El Salvador
	.sy	Syrian Arab Republic	.sz	Swaziland		
T	.tc	The Turks & Caicos Islands	.td	Chad	.tf	French Southern Territories
	.tg	Togo	.th	Thailand	.tj	Tajikistan
	.tk	Tokelau	.tm	Turkmenistan	.tn	Tunisia
	.to	Tonga	.tp	East Timor	.tr	Turkey
	.tt	Trinidad and Tobago	.tv	Tuvalu	.tw	Taiwan
	.tz	Tanzania				
U	.ua	Ukraine	.ug	Uganda	.uk	United Kingdom
	.um	United States Minor Outlying Islands	.us	United States	.uy	Uruguay
	.uz	Uzbekistan				
V	.va	Holy See (Vatican City State)	.vc	Saint Vincent and the Grenadines	.ve	Venezuela
	.vg	Virgin Islands British	.vi	Virgin Islands U.S	.vn	Vietnam
	.vu	Vanuatu				
W	.wf	Wallis and Futuna Islands	.ws	Samoa		
Y	.ye	Yemen	.yt	Mayotte	.yu	Yugoslavia
Z	.za	South Africa	.zm	Zambia	.zr	Zaire (non-existent, see Congo)
	.zw	Zimbabwe				

4. SZÁMÚ MELLÉKLET

AZ INTERNET RÖVID TÖRTÉNETE 2000-IG¹⁹⁴

Év	Események
1957	A Szovjetunió fellövi a Szputnyik-1-et, az első mesterséges holdat. Válaszul az USA megalakítja az ARPA-t (Advanced Research Projects Agency).
1961	Leonard Kleinrock (MIT) papírra veti a csomagkapcsolás elméletét.
1962	J.C.R. Licklider és W. Clark (MIT) A Galaktikus Hálózat koncepciója.
1964	Paul Baran (RAND): Az osztott kommunikációs hálózat elve. Csomagkapcsolt hálózatok elmélete.
1965	A TX-2 (MIT) és az AN/FSQ-32 (SDC) számítógép összekapcsolása analóg modemen keresztül ez az internet őse.
1966	Lawrence G. Roberts (MIT): az első terv az ARPANET-re.
1967	Larry Roberts publikálja az ARPANET tervezetét.
1968	Bolt Beranek and Newman (BBN) nyeri a pályázatot az IMP (Interface Message Processor) megépítésére.
1969	Az első csomópont (node) az UCLA-n szeptember 2-án, második csomópont az SRI-n október 1-én, harmadik csomópont az UCSB-n november 1-én, negyedik csomópont az UU-n decemberben. Az első RFC (Request for Comment) Steve Croker (április 7.)
1970	Az ARPANET hosztok elkezdi használni az NCP-t (Network Control Protocol), az első hosztól-hosztig protokollt.
1971	Az ARPANET már 23 gépegyesre bővült. Ray Tomlinson kifejlesztette az első email programot, amellyel üzeneteket lehetett küldeni a számítógépes hálózaton részt vevő gépek számára. Csak érdekesség, hogy az első E-mail üzenet tartalma a következő volt: QWERTYUIOP. A mi-értre egyszerű a válasz: ez az angol-amerikai PC-b billentyűzetének felső sora. Michael Hart elindítja a Gutenberg Project-et, mely céljával tűzte ki, hogy minél nagyobb tudás és minél több könyvet tegyenek ki szabadon kutathatóvá a hálóra. Az első szöveg az USA Függetlenségi Nyilatkozata volt.
1972	40 számítógépből áll a hálózat. Az első számítógépek közötti chat program. Megalakul az International Network Working Group (INWG). RFC 318 a Telnet specifikáció.
1973	Az első nemzetközi kapcsolódás az ARPANET-re: a University College of London (Anglia) és a NORSAR (Norvégia) kapcsolódik a hálózatra. RFC 454: File Transfer specifikációja.
1974	Vint Cerf és Bob Kahn publikálja a Transmission Control Program (TCP) tervezetét.
1975	Steve Walker elkészíti az első levelezési listát (mailing list) MsgGroup néven.
1976	Az AT&T Bell Laboratórium kifejleszti az UUCP (Unix-to-Unix CoPy) protokollt.
1977	A Wisconsin-i Egyetem 100 számítógépes tudósának megépíti a THEORYNET hálózatot Larry Landweber vezetésével.

¹⁹⁴ Forrás: <http://www.zakon.org/robert/internet/timeline/>.

Év	Események
1978	A TCP protokollt felosztják a TCP és az IP protokollra.
1979	Megalapítják a tudományos vitafórumnak szánt Usenet komputeres világhálózatot az Internet égisze alatt, ahol az azonos érdeklődési körűek csoportokba szerveződhetnek. Alapítók: Tom Truscott és Jim Ellis, a Duke Egyetem két végzőse és Steve Bellovin az Észak-Karolina Egyetemről.
1980	Október 27-én a teljes ARPANET hálózat leállt egy üzeneteket kiíró vírustámadás miatt.
1981	Megalakul a Bitnet a New York-i Egyetemen elektronikus posta- és fájlátvitel céljából. (BITNET = Because It's Time NETWORK) Franciaországban megkezdődik a Minitel/Teletel-szolgálat a France Telecom jóvoltából. Az Internet időközben 213 hoszt-gépre bővült.
1982	Kidolgozzák és bevezetik a TCP-t és az IP-t az adatátvitel „csomagolástechnikáját”, miáltal nagy tömegben küldhetők fájlcsomagok. A TCP/IP mind a mai napig élő és működő Internet-szabvánnyá érett. Az „Internet” szó első említése. A hálózat már 235 hoszt-gépből áll.
1983	Január 1-jén a TCP/IP szabvány lesz az Internet egyetemes nyelve. A Wisconsin Egyetem névszerver (Name Server) kezd működni, hogy az egyre kaotikusabb hálózat könnyebben átlátható nevekkel legyen címezhető. Az ARPANET-ből kiválik a MILNET az addigi 113-ból 68-al.
1984	William Gibson megírja a Neuromanc című sci-fi regényét, amelyben először szerepel a "Cyberspace" kifejezés. A hálózaton már 1024 hoszt gép van. Bevezetik a domain-nevek rendszerét. (Domain Name System = DNS). A szovjet irányítás a KremVax rendszer segítségével rácsatlakozik a USENET hálózatra.
1985	Az Information Sciences Institute elkezd a DNS-ek regisztrációját, a DNS-root szolgáltatást. Március 15-én jegyzik be az első domain-t a Symbolics.com-ot. A továbbiak április 24-én: cmu.edu, purdue.edu, rice.edu, berkeley.edu, ucla.edu, rutgers.edu, bbn.com ; május 23.: mit.edu; május 24.: think.com; június: css.gov, mitre.org; július: .uk
1986	Az Egyesült Államok kormány szerve, a National Science Foundation (NSF) megalakítja a saját NSFNET-jét öt szuperszámítógépből, továbbá 56 ezer bit/sec átviteli sebességű hálózattal. Ezzel megkezdődött az Internet lényegét adó tevékenység: a különféle lokális hálózatok közötti kapcsolatépítés. Ekkorra már 5089 hoszt-gépre bővült a hálózat.
1987	E-mail kapcsolat Németország és Kína között - az első levelet szeptember 20-án Kína küldte. Hostok száma 10 ezer felett.
1988	Kanada (ca), Dánia (dk), Finnország (fi), Franciaország (fr), Izland (is), Norvégia (no) és Svédország (se) csatlakozik az NSFNET-hez. Az Internet 56 ezer hoszt-ra gyarapszik. Az NSFNET sebessége elérte a T1-es szintet (1.544 Mbit/s). Jarkko Oikarinen megalkotja az egyik legnépszerűbb csevegő-szolgáltatást, az IRC-t. (=Internet Relay Chat).
1989	A kapcsolattartás műszaki ügyeleteinek felügyeletére megalakul az Internet Engineering Taskforce (ETF).
1990	Beszünteti működését az ARPANET. Mitch Kapor megalakítja Electronic Frontier Foundation (EFF) a világméretű adatkommunikáció szabadságának a védelmére.
1991	Tim Berners-Lee bemutatja a WWW-t, a World Wide Webet a genfi Európai Atommagkutató Tanácsban (CERN), amely az Interneten belül létesült új hálózat. Általa először sikerül grafikus felhasználói felületekkel is operálni. NSFNET sebessége elérte a T3-at (44.736 Mbit/s). Philip Zimmerman megalkotja a biztonságos azonosításra szolgáló PGP (Pretty Good Privacy) szabványt.

Év	Események
1992	Első ízben továbbítanak hang- és képfájlokat a WWW-en. 1136000 host-gép van. A Világbank (WorldBank) is megjelenik on-line. Az „internetes szörfözés” kifejezést Jean Armour Polly megalkotja és bevezeti a köztudatba.
1993	Az NSF megalapítja az InterNIC-et az Internet címek megállapítására és további kiosztására. A Fehér Ház és az ENSZ online formában megjelenik az Interneten. A címek: ENSZ: http://www.un.org/ Fehér Ház: http://www.whitehouse.gov/ Az Internetet több gazdasági vállalkozás és média használja immár 2 056 000 számítógépről. Az első grafikus internet-böngésző, a Mosaic megszületik április 22-én. Később ebből fejlődött ki a Netscape.
1994	Megjelennek az első online ajánlatok, boltok és áruházak. Megkezdzi működését az első hálózati rádióadó, az RT-FM (Las Vegas, USA) és megnyílik az első Cyber-Bank (First Virtual Bank) is az Interneten; valamint a Pizza Hut megkezdte az on-line rendelt pizzák kiszállítását. Arizona állam Canter és Siegel előterjesztésére törvényt alkot a kénytelen levelek (spam) ellen, ugyanis az amerikai vízum-igénylést ilyen kénytelen levelekben kínálták. A gépek száma: 3.864.000.
1995	Márciusban a WWW adatforgalma meghaladja az FTP-jét. A többi on-line szolgálat is lehetővé teszi előfizetői számára az internetre kijutást. Megjelenik a Weben a Vatikán-állam. A Sun Computer kifejleszti az Internet programnyelvét, a Javát és május 23-án kiadja. A VRML-szabvány pedig lehetővé teszi 3-dimenziós modellek megjelenítését. Megjelenik a RealAudio szabvány az on-line rádiózás könnyebbé tételéért. Megkezdte működését a Radio HK az első üzleti, kizárólag csak az interneten szolgáltató adó. Szeptember 14. után a domain nevek ingyenes regisztrációja megszűnik. Az új ár: 50\$ az üzleti domainek alá! Magyar kormányzati szervek is megjelennek online formában. Gépellomány: 6.642.000.
1996	Megrendezik az első nemzetközi vásárt az Interneten. Már telefonálni is lehet, elég széles kapcsolat esetén pedig a videotelefonálás is megvalósítható. Kormányok, kormányhivatalok, kormányzati szervek mind nagyobb számban jelennek meg a Web-en. Júliusban már kb. 10 millió hoszt működött a világhálón. Az első nagy összegért elkelt domain nevet a CNET vette meg. A név: tv.com - és az ár: 15000\$. A Netscape és a Microsoft között elkezdődik a böngésző-háború. Új legfelsőbb szintű domain-nevek: .firm, .store, .web, .arts, .rec, .info, .nom.
1997	Egyre gyilkosabb verseny folyik a világ minden táján az Internetes lehetőségek minél jobb kihasználásáért. A hatalmas forgalombővülésnek azonban ára is van: egyre gyakoribbak a „dugók”, a hatalmas forgalom okozta torlódások. Hosztok száma: 16 millió. A levelezőlisták központi szervere, a LISZT már 71 618 listát tart számon. 101 803 Name Server van a Whois adatbázisban. A Business.com domain nevet már 150ezer \$-ért adják el.
1998	Megalapítják az Internet2 csoportot a vezető hardware- és software-cégek. Céljuk: az Internet sebességének nagyságrendekkel való megnövelése és a forgalmi dugók kialakulási esélyeinek drasztikus lecsökkentése. Angol nyelvterületről a vállalatok sorra jegyeznék be domain-eket Türkmenisztánban (TradeMark = tm). Az amerikai posta (US Postal Service) bevezeti az elektronikus postabélyeget. Megjelenik az elektronikus üzleti modell, az E-commerce és beindulnak az első portálok; valamint megjelenik az új, igen tömör zenei formátum, az MP3.
1999	A „dotkom” (értsd: .COM) cégek árfolyama hihetetlen emelkedést visz végbe az amerikai tőzsdén. Bár a meredek emelkedés miatt pár cég túl nagyra nő és tönkremegy, de az internet egyértelműen az egyik legerősebb húzóágazattá lesz a világban. Y2K-probléma: kérdés, hogy a kétjegyű évszámmal írt programok és gépek kibírják-e a 2000. évet. (Y2K = Year 2000) Február 22-én beindul az első teljes kiszolgálást nyújtó, tisztán internetes bank: First Internet Bank of Indiana.
2000	Egyre olcsóbbá és elérhetőbbé válik a tömegek számára is a világháló. Az elektronikus vásárlások egyre biztonságosabbak, bár a karácsonyi rohamot csak a legnagyobb áruházak bírják áruval. Az év az amerikai „dotkom” cégek jelentős visszaesését hozta. Hostok száma: kb. 50 millió. Május 16-án az Internet2 közössége kiadja az IPv6-os szabványt. November 16-tól jegyezhető új domain-ek: .aero, .biz, .coop, .info, .museum, .name, .pro. Megjelenik az első komoly megosztásra alapuló zenecserélő-szolgáltatás, a Napster.

5. SZÁMÚ MELLÉKLET

A VILÁGHÁLÓ (WWW) NÖVEKEDÉSE (1990-2006)¹⁹⁵

hó/év	oldalok szám	hó/év	oldalok szám	hó/év	oldalok szám
12/90	1	04/99	5.040.663	04/03	40.100.739
12/91	10	05/99	5.414.325	05/03	40.444.778
12/92	50	06/99	6.177.453	06/03	40.936.076
06/93	130	07/99	6.598.697	07/03	42.298.371
09/93	204	08/99	7.078.194	08/03	42.807.275
10/93	228	09/99	7.370.929	09/03	43.144.374
12/93	623	10/99	8.115.828	10/03	43.700.759
06/94	2.738	11/99	8.844.573	11/03	44.946.965
12/94	10.022	12/99	9.560.866	12/03	45.980.112
06/95	23.500	01/00	9.950.491	01/04	46.067.743
01/96	100.000	02/00	11.161.811	02/04	47.173.415
03/96	135.396	03/00	13.106.190	03/04	48.038.131
04/96	150.295	04/00	14.322.950	04/04	49.750.568
05/96	193.150	05/00	15.049.382	05/04	50.550.965
06/96	252.000	06/00	17.119.262	06/04	51.635.284
07/96	299.403	07/00	18.169.498	07/04	52.131.889
08/96	342.081	08/00	19.823.296	08/04	53.341.867
09/96	397.281	09/00	21.166.912	09/04	54.407.216
10/96	462.047	10/00	22.282.727	10/04	55.388.466
11/96	525.906	11/00	23.777.446	11/04	56.115.015
12/96	603.367	12/00	25.675.581	12/04	56.923.737
01/97	646.162	01/01	27.585.719	01/05	58.194.836
02/97	739.688	02/01	28.125.284	02/05	59.100.880
03/97	883.149	03/01	28.611.177	03/05	60.442.655
04/97	1.002.612	04/01	28.669.939	04/05	62.286.451
05/97	1.044.163	05/01	29.031.745	05/05	63.532.742
06/97	1.117.259	06/01	29.302.656	06/05	64.808.485
07/97	1.203.096	07/01	31.299.592	07/05	67.571.581
08/97	1.269.800	08/01	30.775.624	08/05	70.392.567
09/97	1.364.714	09/01	32.398.046	09/05	71.723.098
10/97	1.466.906	10/01	33.135.768	10/05	74.409.971
11/97	1.553.998	11/01	36.458.394	11/05	74.572.794
12/97	1.681.868	12/01	36.276.252	12/05	74.353.258
01/98	1.834.710	01/02	36.689.008	01/06	75.251.256
02/98	1.920.933	02/02	38.444.856	02/06	76.184.000
03/98	2.084.473	03/02	38.118.962	03/06	77.568.868
04/98	2.215.195	04/02	37.585.233	04/06	80.655.992
05/98	2.308.502	05/02	37.574.105	05/06	81.565.877
06/98	2.410.067	06/02	38.807.788	06/06	85.541.228
07/98	2.594.622	07/02	37.235.470	07/06	88.166.395
08/98	2.807.588	08/02	35.991.815	08/06	92.615.362
09/98	3.156.324	09/02	35.756.436	09/06	96.854.877
10/98	3.358.969	10/02	35.114.328	10/06	97.932.447
11/98	3.518.158	11/02	35.686.907	11/06	101.435.253
12/98	3.689.227	12/02	35.543.105		
01/99	4.062.280	01/03	35.424.956		
02/99	4.301.512	02/03	35.863.952		
03/99	4.349.131	03/03	39.174.349		

Megjegyzés: az oldalak tulajdonképpen webszervert jelentenek (egy hoszton több oldal lehet különböző domain vagy port használatával).

¹⁹⁵ Forrás: <http://www.zakon.org/robert/internet/timeline/>

6. SZÁMÚ MELLÉKLET

AZ INTERNET NÖVEKEDÉSE (1969-1997)¹⁹⁶

Dátum	Hosztok	Dátum	Hosztok	Hálózatok	Domainek
12/69	4	07/89	130,000	650	3,900
06/70	9	10/89	159,000	837	
10/70	11	10/90	313,000	2,063	9,300
12/70	13	01/91	376,000	2,338	
04/71	23	07/91	535,000	3,086	16,000
10/72	31	10/91	617,000	3,556	18,000
01/73	35	01/92	727,000	4,526	
06/74	62	04/92	890,000	5,291	20,000
03/77	111	07/92	992,000	6,569	16,300
12/79	188	10/92	1,136,000	7,505	18,100
08/81	213	01/93	1,313,000	8,258	21,000
05/82	235	04/93	1,486,000	9,722	22,000
08/83	562	07/93	1,776,000	13,767	26,000
10/84	1,024	10/93	2,056,000	16,533	28,000
10/85	1,961	01/94	2,217,000	20,539	30,000
02/86	2,308	07/94	3,212,000	25,210	46,000
11/86	5,089	10/94	3,864,000	37,022	56,000
12/87	28,174	01/95	4,852,000	39,410	71,000
07/88	33,000	07/95	6,642,000	61,538	120,000
10/88	56,000	01/96	9,472,000	93,671	240,000
01/89	80,000	07/96	12,881,000	134,365	488,000
		01/97	16,146,000		828,000
		07/97	19,540,000		1,301,000

Hosztok = egy számítógép rendszer regisztrált IP címmel

Hálózatok = regisztrált A/B/C osztályú címek

Domainek = regisztrált domain név

A VILÁG INTERNET HASZNÁLATA 2006 MÁJUSÁBAN¹⁹⁷

Régiók	Internetezők száma	Lakossági penetráció	Részesedés a világból	Növekedés 2000-2005
Afrika	23649000	2.6 %	2.3 %	423.9 %
Ázsia	364271000	9.9 %	35.6 %	218.7 %
Európa	291601000	36.1 %	28.5 %	177.5 %
Közél-Kelet	18203000	9.6 %	1.8 %	454.2 %
Észak-Amerika	227304000	68.6 %	22.2 %	110.3 %
Latin-Amerika, Karibi térség	79963000	14.4 %	7.8 %	342.5 %
Óceánia, Ausztrália	17873000	52.6 %	1.7 %	134.6 %
Világ összesen	1.022.864.000	15.7 %	100.0 %	183.4 %

¹⁹⁶ Forrás: <http://www.zakon.org/robert/internet/timeline/>.

¹⁹⁷ Forrás: <http://hvg.hu/Tudomany/20060515internet.aspx>.