

## **Nyikes Zoltán<sup>258</sup>: A Big Data, mint a kritikus infrastruktúra része**

*Absztrakt: Napjaink információ szükséglete hatalmas. Az információ mennyisége óriási mértékben növekszik, évek alatt megsokszorozódik. A kritikus infrastruktúra és a részét képező kritikus információs infrastruktúra biztosítja az állam és polgárai jólétét. A mindennapi döntéshozatalok, legyenek azok üzleti-, közigazgatási-, vagy magánügyi döntések, egyre több és különböző információra van szükségünk. A publikációban a Big Data és a kritikus infrastruktúra kapcsolata kerül bemutatásra. Az elmúlt időszakban a Big Data, mint fogalom megjelenése és alkalmazása nagy hangsúlyt kapott. A publikáció bemutatja a Big Data jelenséget, fogalmi és gyakorlati szinten. Bemutatásra kerül a Big Data története. Megvizsgálásra kerül az alkalmazásának feltételeit és lehetőségei, mely az elkövetkezendő időszak gazdasági növekedését szolgálhatja. Bemutatásra kerül Digitális Nemzeti Fejlesztési Program, ami a kormányzat és az Európai Unió finanszírozásában létrejövő fejlesztés, amely a magyar társadalom jólétét növelő beruházás, melynek alkalmazásával a Big Data mindenki számára könnyebben elérhető digitális szolgáltatás lehet. Megvizsgálásra kerül Big Data kockázati a biztonsági tényezői, melyek elengedhetetlenek a kritikus infrastruktúra védelmének-, valamint a személyes adatok védelmének hatékony biztosításához.*

### **Mi az a Big Data?**

Az ipari társadalmat a XX. század végén felváltotta az információs, tudáslapú társadalom. Ez a változás jelentősen átalakította a mindennapjainkat. Szinte minden magyar család legalább 2-3 mobiltelefonon előfizetéssel rendelkezik. Az új autók esetében, már-már az alapfelszereltség részét képezi a GPS<sup>259</sup> alapú navigációs rendszerek. A világháló már olyannyira a mindennapjaink részévé vált, hogy teljesen természetessé vált az interneten történő vásárolás, banki ügyintézés, az adóbevallás beküldése, a biztosítások kötése, váltása. Mindezek mellett filmet nézünk, zenét, rádiót hallgatunk, vagy éppen híreket olvasunk rajta. Ezen kívül még nem is került említésre a napjainkban mind népszerűbb közösségi oldalak látogatása, csak hogy néhány lehetőségét említve a napról napra fejlődő világhálónak. Ugyanakkor egy másik digitális megoldás lépett be mindannyiunk nappalijába a digitális tévénézés és rádiózás lehetőségével, ami új korszakot indít el a média világában. A rengeteg digitális eszköz használat a 90-es évektől kezdődően, egyre több és több digitális adat generálódását, felhalmozódását eredményezte. Mai világunkban már a *terabájt* és a *petabájt* fogalmak lassan hétköznapivá váltak, válnak. Aki az információs technológiával foglalkozik, annak már nem idegenek az *exabájt* és a *zettabájt* fogalmak. Ezek már akkora adattömegek, amelyek feldolgozása, tárolása komoly infrastruktúrát és erőforrás-bevonást igényel. Ezen adattömegek, illetve a hozzá kapcsolódó tevékenységek összefoglaló elnevezésekén megjelent a „*Big Data*”, mint „*nagy adat*” fogalom, mely napjainkban egyre divatosabb kifejezéssé kezd válni.

### **Az adatfelhalmozás története**

Ahhoz, hogy megérthessük pontosan mi is az a *Big Data* és melyek a hozzá kapcsolódó tevékenységek tekintsük át az elektronikus számítógép megjelenését követő időszak adat felhalmozódását egészen napjainkig, néhány példán keresztül.

Mindenki által ismert, hogy a digitális adat mennyisége robbanásszerűen növekedett az elmúlt években. Az alábbi idővonal pillanatképei néhány nagyobb amerikai vállalati adatgyűjtés évtizedenkénti bontásban.

258 Óbudai Egyetem Biztonságtudományi Doktori Iskola

259 GPS (Global Positioning System) Globális Helymeghatározó Rendszer

### Az 1950-es évek

Ebben az évtizedben a John Hancock<sup>260</sup> biztosítási cég rendelkezett a legnagyobb felhalmozott digitális adattömege, amit a biztosítottakról gyűjtöttek és tároltak. Ez a cég 1955-ben az egyik úttörője volt az ügyfelek adatainak digitalizálásában és tárolásában. A kétmillió életbiztosítási információt egy UNIVAC<sup>261</sup> számítógép segítségével dolgozták fel a mintegy 600 megabájtnyi adatot.

### Az 1960-as évek

A Holdra-szállás évtizedében az American Airlines<sup>262</sup> fejlesztette ki és alkalmazta a SABRE<sup>263</sup> helyfoglalási rendszert, amit az akkori egyik legnagyobb IBM<sup>264</sup> számítógépes rendszerre fejlesztettek ki. A SARBE elsőik egyike volt az olyan online számítógépes rendszereknek, amely lehetővé tette a légitársaság számára a hatalmas helyfoglalási mátrix nyomon követését. Ez az adatfelhalmozás mintegy 807 megabájtnyi adatot jelentett.

### Az 1970-es évek

Az első mikroprocesszor, az Intel 4004<sup>265</sup> megjelenésének az évtizedében a FedEx<sup>266</sup> szállítványozó cég vezettette be a COSMOS<sup>267</sup> rendszert, ami lehetővé tette a vállalat számára, hogy vizsgálhassa és nyomon követhesse a hatalmas mennyiségű csomagot-szállítványokat az egész világon. Ezek az adatok már 80 gigabájtnyi adatot jelentettek.

### Az 1980-as évek

A személyi számítógépek megjelenésének az évtizedében, bankoknak volt legnagyobb szerepük az adatmennyiségek a növekedésben. Az ATM<sup>268</sup>-ek elterjedésével a bankok a tranzakciós adatok gyűjtésére és elemzésére összpontosítottak az ügymeneteikben. A CitiCorp NAIB<sup>269</sup> ekkor 450 gigabájt adatot tárolt elektronikusan.

### Az 1990-es évek

Az internet megjelenésének évtizedében vált a Wal-Mart<sup>270</sup> cég a legnagyobb amerikai kereskedelmi áruházláncá, és feltételezhető, hogy ez a cég volt a legnagyobb kereskedelmi adatokkal rendelkező áruházlánc a világon, ami 180 terabájtnyi adatot jelentett. A szakértők becslése szerint az Amazon.com<sup>271</sup> a 90-es évek vége felé még csak egy számjegyű terabájtnyi tárolt adattal rendelkezett.

### A 2000-es évek

A 90-es évek internet-robbanása előkészítette a terepet az online<sup>272</sup> vállalatok 2000-es években történő megjelenéséhez, mint a világ „nagy adat” vezetőinek. Ezekben az években vált vezető céggé a Google<sup>273</sup>, amely 25 petabájtnyi adatot kezelt digitalizált formában.

<sup>260</sup> John Hancock Mutual Life Insurance Co., USA-beli székhelyű biztosító társaság

<sup>261</sup> Universal Automatic Computer

<sup>262</sup> American Airlines, Inc.; amerikai légitársaság

<sup>263</sup> Semi-automated Business Research Environment

<sup>264</sup> International Business Machines; USA-beli székhelyű, multinacionális informatikai cég

<sup>265</sup> Egy 4 bites mikroprocesszor, amit az Intel vállalat adott ki 1971-ben.

<sup>266</sup> FedEx Corporation; egy amerikai globális szállítványozó vállalat. A név a cég eredeti légiforgalmi divíziójának, a Federal Expressnek a nevéből származik.

<sup>267</sup> Customer Operations Service Master On-line System

<sup>268</sup> Automated Teller Machine; bankautomata

<sup>269</sup> CitiCorp North American Investment Bank; USA-beli pénzügyintézet

<sup>270</sup> Wal-Mart Stores; USA székhelyű, a világ egyik legnagyobb kiskereskedelmi cége

<sup>271</sup> Amazon.com, Inc.; elektronikus kereskedelemmel foglalkozó amerikai cég

<sup>272</sup> Közvetlen kapcsolaton alapuló információközvetítő, illetve adatcserélő eljárás



1. ábra: SI-prefixumok (készítette a szerző a [4] alapján)

A további példa elgondolkodtató és ijesztő a jövőre nézve. Az NSA<sup>278</sup> által a 2013-as év végén átadásra került új adatközpont Utahban<sup>279</sup> 1 yottabájnyi adat tárolására alkalmas.

Az Intel<sup>280</sup> cég 2013-as felmérése alapján, az interneten percenként 640 terabájt adat továbbítódik, amely 204 millió e-mailt, több mint 3 millió keresést, 6 millió Facebook bejegyzést, 30 órányi videó feltöltést és 1,3 millió videó megnézését jelent a YouTube<sup>281</sup>-on, illetve 50 gigabájnyi adat termelődik percenként a Large Hadron Collider<sup>282</sup> esetében. [3]

Belátható, hogy ez a rengeteg adat olyan tekintélyt parancsoló, hogy érdemben illik vele foglalkozni.

Szimbólum	Név	Jelentés	Érték	
Ki	kibi-	bináris kilo	$2^{10}$	$= 1000^1 \times 1,024$
Mi	mebi-	bináris mega	$2^{20}$	$= 1000^2 \times 1,048 576$
Gi	gibi-	bináris giga	$2^{30}$	$= 1000^3 \times 1,073 741 824$
Ti	tebi-	bináris tera	$2^{40}$	$= 1000^4 \times 1,099 511 627 776$
Pi	pebi-	bináris peta	$2^{50}$	$= 1000^5 \times 1,125 899 906 842 624$
Ei	exbi-	bináris exa	$2^{60}$	$= 1000^6 \times 1,152 921 504 606 846 976$
Zi	zebi-	bináris zetta	$2^{70}$	$= 1000^7 \times 1,180 591 620 717 411 303 424$
Yi	yobi	bináris yotta	$2^{80}$	$= 1000^8 \times 1,208 925 819 614 629 174 706 176$

2. ábra: Bináris prefixumok [készítette a szerző a [5] alapján]

### A Big Data jelenség

Az élet természetes velejárója, hogy mindig új és újabb dolgokat kell kitalálni, vagy éppen régieket átfazonírozva újak beállítani. Az informatikai piac is előszeretettel dob be könnyen megjegyezhető kifejezéseket, melyek mögé trendet és természetesen terméket, szolgáltatást is társít. Korunk egyik ilyen hívószava a *Big Data* lett, amelyről nem kisebb dolgot állítanak a szakemberek, minthogy ez a terület lesz a *21. század olaja*. Azaz az adatok minél teljesebb körű összegyűjtése, strukturálása és elemzése fogja meghatározni és előre vinni a világ gazdaságát.

A *Big Data* definíciója ugyan nem kristályosodott ki még egyetlen elfogadott meghatározásban, de azt nagyjából mindenki elfogadja, hogy a világban egyre nagyobb mennyiségben és egyre változatosabb formában fellelhető komplex adattömegek módszeres begyűjtését, analizálását kell érteni alatta. Az elemzési folyamat végén szerencsés esetben olyan következtetésekre juthatnak a kapott információkból a döntéshozók, amelyekkel nagyobb üzleti értéket képesek felmutatni, mint abban az esetben, amikor nem használnak big data-hoz köthető módszereket. [6]

<sup>278</sup> National Security Agency (Nemzetbiztonsági Ügynökség); az Amerikai Egyesült Államok rádióelektronikai hírszerző szervezete,

<sup>279</sup> Az Amerikai Egyesült Államok 45. állama

<sup>280</sup> Intel Corporation egy eredetileg amerikai vállalat, mely mára nemzetközivé nőtte ki magát. Elsősorban mikroprocesszorairól ismert.

<sup>281</sup> YouTube; egy videómegosztó weblap, ahol a felhasználók videoklipeket tölthetnek fel és nézhetnek meg

<sup>282</sup> Large Hadron Collider; Nagy hadronütköztető a CERN 2008-ban átadott részecskegyorsítója és ütköztetőgyűrűje

A „big” azt jelenti, hogy nagyon sok adatból álló adatbázis vagy adatkészlet. Olyan sokból, hogy azt már nehéz a megszokott, általánosan rendelkezésre álló adatbázis-kezelőalkalmazásokkal manipulálni. Óriási adattömeg, amelynek kezelése (létrehozása, tárolása, feldolgozása, továbbítása, lekérdezése stb.) a technikai lehetőségek határát feszegeti.

A *Big Data* jelenség alatt nem csak az adatok mennyiségének növekedése, a mennyiségi értelemben vett adatrobbanás értendő.

Mindezek alapján a *Big Data* jelenséget a következőképpen írhatjuk le: hatalmas, nagyfokú változatossággal és komplexitással jellemezhető, gyorsan keletkező és szaporodó adattömegek megjelenése, amelynek hasznosítására kevés idő áll rendelkezésre. [7]

#### A 4V modell

Mint ahogy a fentiekben bemutatásra került, sokkal több adat létezik napjainkban, mint akár néhány évvel ezelőtt, és céljá vált e különféle forrásokból származó adathalmazok összesítése, szintetizálása, méghozzá a jobb, sikeresebb üzleti döntések meghozatala érdekében.

„*Volume*” - az információ mennyisége. Az adatok mennyisége napjainkra valóban sokkoló méreteket öltött.

„*Variety*” - az adatok sokfélesége, változatossága. A közösségi média csak egy része a *Big Data*-nak, bár kétségkívül jelentős része. Ide sorolhatóak még a viselkedéstani (fogyasztói magatartással kapcsolatos) adathalmazok, a tranzakciós adatok, a videók, a fényképek, a GPS adatok, a hangfelvételek stb.

„*Velocity*” - az adatok terjedési sebessége. Az adatokhoz való hozzáférés és az adatelemzés sebessége nagyon megugrott, lehetővé téve ezzel az ún. streaming adatok (adattömegek) elemzését.

„*Veracity*” - Az adatok érvényessége. A fentiek egyike sem ér semmit, ha az adatokat nem sikerül valós időben, a vállalatok számára az üzleti döntéseket segítő módon integrálni és szintetizálni. [8]

#### A Big Data kiaknázásához szükséges feltételek

Adatgyűjtés: Egy intézmény különböző adatforrásokból gyűjti össze az adatokat, amelyeket aztán masszívan párhuzamosított rendszerek segítségével - *gyakran grid<sup>283</sup> technológia alkalmazásával* - oszt szét; minden egyes csomópont az adatok egy részhalmazát dolgozza fel.

Feldolgozás: A rendszer ugyanezt a nagy teljesítményű párhuzamosságot használja fel annak érdekében, hogy gyors számításokat végezzen az adatokon minden egyes csomópontban, amelyek az eredményeket jobban felhasználható adathalmazokká csökkentik.

Adatkezelés: A feldolgozandó nagy mennyiségű adatok gyakran heterogének, mivel különböző rendszerekből származnak. Csaknem minden esetben szükség van az adatok értelmezésére, definiálására, tisztítására, gazdagítására és biztonsági célokból történő auditálására.

Mérés: Az alkalmazó szervezeteknek folyamatosan mérniük, monitorozniuk kell adataik minőségét, biztosítaniuk kell adataik összekapcsolhatóságát. A felhasználói követelmények határozzák meg, hogy pontosan mit szükséges mérni és milyen célt érdemes kitűzni az adatok minőségével kapcsolatban.

Felhasználás: Az adatok felhasználásának összhangban kell lennie a feldolgozás eredeti követelményeivel.

Tárolás: Az adatfelhő (data-as-a-service) trend kialakulásával az adatok egyre inkább egy adott helyen találhatóak meg, míg az azokat elérő programok helye nem rögzített. Akár rövid távú kötegelt feldolgozás, akár hosszú távú megőrzés céljából tárolja a szervezet az adatokat, a tárolási megoldásokat tudatosan kell kezelnie.

Adatvaqyon-gazdálkodás: Az adatokkal kapcsolatos felhasználói szempontú irányelveket és felügyeletet öleli fel, amely az előző 6 lépés mindegyikére vonatkozik. [9]

<sup>283</sup> A grid egy infrastruktúra, mely közvetlen hozzáférést biztosít a gépi feldolgozóerőhöz és az adattároló kapacitáshoz is, mely megosztva helyezkedik el a Földön

## Big Data alkalmazása

### Analitika a közszférában

A nagy mennyiségű adatok kiaknázásában rejlő lehetőségekkel és kihívásokkal a közszféra szereplői is szembesülnek – az ő munkájukat, beleértve a kormányzati döntéshozatalt is, szintén hatékonyan támogatják az analitikai megoldások. Jó példa erre a mikroszimuláció, amellyel modellezhető, hogy az egyes intézkedések milyen hatást gyakorolnak a gazdasági egységek – *például egyének vagy háztartások* – szintjén. A *Központi Statisztikai Hivatal* (KSH) pedig a 2011. évi népszámlálás adatainak központi feldolgozása során használja a SAS eszközt<sup>284</sup>. [10]

*Vágújhelyi Ferenc*, a NAV<sup>285</sup> informatikai szakfőigazgatója szerint az adatrobbanás mértékét jól érzékelteti, hogy az adóhatóságnál 1988 és 2000 között összesen gyűlt össze annyi adat, mint a 2013 első négy hónapjában. A feldolgozás már jórészt online folyamat, az évi 80 millió ügyből csak másfél milliót intéznek papíralapon.

A szakértő megerősítette az IBM kutatója, *Robyn Schwartz* trendjelzését a szingularizáció ügyében. Ma már a NAV sem csak aggregált adatokkal dolgozik, hanem a tételes ÁFA esetében például atomizált, nagyszámú eseményekkel is. (A beérkező adatmennyiség a pénztárgépek bekötése miatt nagyságrendekkel megugrott.) A beszerzett petabájtos tároló mellett pedig egyre bonyolultabb matematikai modellekre, elemzőképességekre van szükség. [11]

### Big Data a népszámlálás segítségében

A 2013 év nagy hazai Big Data történetét *Kerékyártó Sándor*, az Educatio ügyvezető igazgatója mutatta be. A magyar népszámlálás 11,4 millió kitöltött kérdőívét ugyanis az Educatio dolgozta fel a KSH számára. A továbblépés lehetősége az adatbázisok összekötésében rejlik. Például a felsőoktatásban végzetek elhelyezkedését figyelő, az Educatio által kezelt diplomás pályakövető rendszer még csak a NAV és az OEP<sup>286</sup> adatbázisaival van összekötve, de ha a foglalkoztatási hivatal adatbázisával is összekapcsolódna, akkor a rendszer nem csak regisztrálni tudná, de munkalehetőséget is ajánlana. [11]

### Big Data a magyar mezőgazdaságban

Olyan környezetben, mint a mezőgazdasági termelés, különösen nagy eredményeket lehet elérni egy hatékony információs rendszerrel, mivel a rendelkezésre álló tudás kihasználása nehézkes. *Mezőgazdasági tudásközpont* létrehozását célzó kutatást indított a *Hewlett-Packard Informatikai Kft.*, a *Széchenyi István Egyetem*, a *Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézete* (MTA SZTAKI), és az *eNET Internetkutató és Tanácsadó Kft.*, a projekthez a konzorcium több, mint 2 milliárd forint támogatást nyert az *Új Széchenyi Terv* pályázatán - közölték a résztvevők az MTI-vel.

A mezőgazdasági tudásközpont és döntéstámogató rendszer megvalósítása 2013. januárban indult, és várhatóan 2017 végéig valósul meg.

A közlemény szerint a projekt célja egy széleskörű tudásbázis létrehozása a "Big Data" elemző rendszer mezőgazdasági területre történő implementálásával, amely szolgáltatásként, illetve tudástárként a hazai és a nemzetközi piacon egyaránt segítségére lesz az agrárszféra egyes szereplőinek. A kutatás során, az ország területén elszórtan ezer szenzoregységet helyeznek ki, s az adatokból, illetve az adatbázisokból és több egyéb nemzetközi, a témához közvetve kapcsolódó nyílt tudásbázis segítségével kifejlesztenek egy komplex elemző, előrejelző és döntéstámogató megoldást. [12]

<sup>284</sup> SAS egy üzleti analitikai szoftvereket szolgáltatásokat és üzleti intelligenciát fejlesztő multinacionális cég,

<sup>285</sup> Nemzeti Adó- és Vámhivatal

<sup>286</sup> Országos Egészségbiztosítási Pénztár

## Big Data alkalmazása külföldön

### A holland vízügyi alkalmazás

Hollandiában elindult a *Digital Delta* program, ami a Big Data hatékonyabb elemzéssel járul hozzá az árvízvédelem fejlesztéséhez és a holland vízgazdálkodási rendszer optimalizálásához. A kezdeményezés célja a katasztrófák és a környezetkárosodás megakadályozása, de a rendszer *15 százalékkal* csökkentheti a vízügyi feladatok költségeit is.

A program során felálló új felügyeleti rendszer a Big Data elemzése során olyan problémákra keres megoldást, mint az ivóvíz minőségromlása, a szélsőséges időjárás, vagy az árvíz- és aszályveszély. Pontos modelljei abban segítik a holland hatóságokat, hogy optimális döntéseket hozzanak például a víz tárolásával, elvezetésével kapcsolatban, vagy megakadályozzák a sós víz és szennyezések ivóvízbe szivárgását, illetve a szennyvíz túlfolyását.

A rendszer fontos jellemzője, hogy az adatok azonnal megoszthatók a szervezetek és hivatalok között, a vizuális adatmegjelenítés és a részletes elemzések valós idejű információkat biztosítanak a vízügyi szakemberek számára. A hatóságok megakadályozhatják a katasztrófákat vagy a környezetkárosodást, koordinálhatják a szükséges lépéseket, és akár *15 százalékkal* csökkenthetik költségeiket. A *Digital Delta* hosszú távon a teljes vízgazdálkodás hatékonyabb működését teszi majd lehetővé. [13]

### Big Data alkalmazása az USA bűnüldözésében

A BBC<sup>287</sup> „*The age of Big Data*” című dokumentumfilmjében [14] a Los Angeles egyik körzetében zajló rendőri munkát mutatják be elsőként. Egy matematikus az előre jelezhetetlen földrengéseket kísérő nagyon is kiszámítható – és modellezhető – utóregések analógiájára kezdett el gondolkodni a bűnesetek előrejelzésén. A rendőrségi adatbázis évtizedekre visszamenőleg tárolja a különböző bűneseteket. Ezt az adatbázist kezdte el vizsgálni a kutatócsoport, mely végül előállt egy algoritmussal, amely a megtörtént bűntények mintázata alapján minden nap elkészít egy előrejelzést, amelyben kiemeli az előző napok eseményei és a múltbeli tapasztalatok alapján az adott napon kiemelt kockázatúnak minősülő helyeket. A járőrök kézhez kapták a kockázatos zónákat jelölő térképeket – hangsúlyozzuk: nem arról van szó, hogy bejelölték a köztudottan balhész utcasarkokat, hanem naponta frissítették az előrejelzést az események tükrében -, és napi rutinjuk csak annyival módosult a szokásoshoz képest, hogy az adott térségeket jóval gyakrabban érintették a menetrendszerű cirkálás közepette. Az eredmény megdöbbentő: a körzet egész *Los Angelesben* a legnagyobb mértékben javította a statisztikáit. *26%-kal csökkent* a rablások száma, *12%-kal* a vagyontárgy ellen elkövetett bűncselekmények mennyisége. A sikeren felbuzdulva az egész városra kiterjesztették a programot, és egyre több község veszi azt át.

### Big Data az orvostudomány szolgálatában

Noha eddig talán sokan nem gondoltak bele, könnyű belátni: az orvostudomány és a „Nagy Adat” jelenség összefonódása teljesen természetes, elkerülhetetlen jelenség. Hiszen hol lehet nagyobb sűrűségben adatokra bukkanni, mint a saját DNS<sup>288</sup>-ünkben, mely közel *3 milliárd* karakter formájában kódolja önnön tervrajzunkat és használati utasításunkat? A DNS-ben kódolt információ határozza meg, hogy sejtjeink milyen proteinek állítanak elő. Kevés dolog fejlődik olyan ütemben, mint a genomszekvenálás<sup>289</sup>. 1984-ben egyetlen nap alatt nagyjából egy DNS *300 karakterét* lehetett elolvasni. Fáradságos, lassú és költséges munka volt ez, mely még csak nem is kecsegtetett annak lehetőségével, hogy egy szép napon emberi időn belül el lehet jutni a *3 milliárd* karakteres mű végére. Új technika kidolgozására volt szükség. Ezt tűzte ki célul a mára már történelemmé vált

<sup>287</sup> BBC – British Broadcasting Corporation, állami tulajdonú, brit közszolgálati műsorszolgáltató, egyben a világ legnagyobb műsorszolgáltatója.

<sup>288</sup> DNS – deoxiribonukleinsav; angol rövidítése: DNA - deoxyribonucleic acid, örökítőanyag

<sup>289</sup> Genomszekvenencia – bioinformatikai úton történő fehérjekódoló gének azonosítása.

Human Genome Project<sup>290</sup> 1990-ben. Míg 2001-ben a költség nagyságrendileg *100 millió dollár* volt (a fentebb emlegetett *3 milliárd* a projekt teljes költsége volt, melybe beletartozott az új eljárások, berendezések kifejlesztése is), addig ez az érték 2007-re *tizedére csökkent*, jelenleg pedig valahol a *3-4000 dollár* környékén jár... Az előrejelzések szerint 3-4 éven belül egy emberi genom elolvasása összesen *10 dollárba* (!!!) fog kerülni. A költségek tehát rohamosan csökkennek, de az egészségügy szempontjából ez önmagában mit sem érne, ha éveket kellene várni az eredményre. A szekvenálás sebessége tehát legalább annyira fontos paraméter, mint az ára. Míg egy évtizeddel ezelőtt a DNS analízis költségének *80%-át* maga a szekvenálás jelentette, addig most ez csak parányi töredéke a kiadásoknak a fentebb vázolt tendenciáknak köszönhetően. A költségek *60%-át* ma az adatok feldolgozása adja. Ne felejtjük el, a szekvenáló berendezés a véletlenszerűen apró darabokra szabdaltságot DNS rövid szakaszait határozza csak meg, ezt követően hátra van még a már meghatározott szegmensek összeillesztése, a hibák kijavítása, stb. A kihívást tehát nem a nyers adatok előállítását jelenti, hanem azok feldolgozása. Most – nem a jövőben, hanem már ténylegesen ezekben a percekben is – lehetőség van arra, hogy egy vérvételt követően elolvassák a páciens DNS-ét, és abból következtessenek arra, hogy mely proteinek termelése vagy működése tér el az elvárttól. Itt arról van tehát szó, hogy a betegség tüneteinek – úgy, mint a sejtek elváltozása – megfigyelése helyett közvetlenül annak okait lehet vizsgálni a DNS szintjén, az adott páciensre szabott módon. Ez a váltás különösen gyümölcsöző lehet a rákos megbetegedések ellen folytatott küzdelemben. Másrészt a betegségek mechanizmusának megértése is könnyebbé válhat: a példának okáért rákos sejtek DNS-e eltárolható, így a világ összes páciensének leleteiből egy olyan adatbázis hozható létre, amelynek analizálásával korábban soha észre nem vett mintázatok rajzolódhatnak ki, esetleg utat mutatva a lehetséges gyógymódok felé. Ezekből sokat talán észre sem veszünk majd, hiszen a kórházak laborjaiban, a szemüktől elzárva végzik majd az elemzéseket, mi „csupán” a pontosabb diagnózisoknak és az egyre több, ma még gyógyíthatatlannak vélt betegség leküzdésének örülhetünk majd, na és persze a kevesebb mellékhatással és kellemetlenséggel járó kezeléseknél. [15]

### **Big Data az “okos városokban”**

Rio de Janeiróban<sup>291</sup> a pontos *időjárás előrejelzés* garantálja a megapolisz *közlekedésbiztonságát*, Birminghamban<sup>292</sup> mobilapplikáció segít a *parkolóhely keresésében*, míg Szingapúrban<sup>293</sup> a forgalom valós idejű követésével dolgoznak a *dugók elhárításán* a városfejlesztők. A városvezetők világszerte hasonló problémákra keresnek megoldást, amelyhez a modern technológiák és rendszerek kiépítése mellett óriási adatmennyiség feldolgozására van szükség; a szenzorokból, járművekből, az áramszolgáltatatótól vagy éppen a rendőrségtől begyűjtött nyers adatok elemzése és értelmezése pedig jól képzett szakembereket kíván. [16]

### **A brit minta az adatbázisok egyesítésére**

A brit kormány egyik legnagyobb jelenkori problémája az, hogy minden egyes részlege más és más adatbázisban dolgozik, az adatok megosztását pedig föltöbb rosszul intézik egymás között. Mivel az információhiány sokszor problémákat vagy felesleges kifizetéseket, késéseket okoz, ezért mindezt megelőzendő a kormány azon gondolkodik, hogy a létező összes adatbázist egyesíti. Mindez persze azt jelenti, hogy az emberekről minden adat egy helyen lenne, a kibányászott big data-elvű működés pedig azt eredményezi, hogy aki belép az adatbázisba, az szó szerint mindent tud az életünkről. A hibák kiküszöbölésével mindenesetre 46 milliárd eurós megtakarítást remélnek. [17]

<sup>290</sup> Humán Genom Projekt – (HGP) egy nagyszabású projekt volt 1990-től 2006 májusáig, mely feltárta a teljes emberi genomot egészen a nukleotidok (bázispárok) szintjéig, és azonosította a benne található összes gént.

<sup>291</sup> Rio de Janeiro Brazília egyik legismertebb városa

<sup>292</sup> Birmingham az Egyesült Királyság második legnagyobb saját önkormányzattal rendelkező városa.

<sup>293</sup> Szingapúr független ország, városállam Ázsia délkeleti részén.

## A Big Data és a jövő kihívása

### A “nagy adat tudós”

A Big Data módszernek köszönhetően létrejött új kutatási eszközök folyamatosan változnak, lehetőséget teremtve az előrejelző elemzésre. Ez a folyamat jelentős és sürgető igényt generált az ún. data scientist<sup>294</sup> szakember megjelenésére, aki kiaknázhajta a Big Data lényegét, a megszerzett információt pedig hatékonyan adhatja át üzleti célok megvalósítására. [8]

A változás folyamat a szervezetek minden területét érinteni fogja, elsőként természetesen az IT<sup>295</sup>-részleget. Míg eddig az informatikai vezetők némileg aggódtak, hogy a folyamatok automatizálása részlegük leépítésével is együtt jár, azonban minőségi váltás megy végbe az új munkakörök létrejöttével (tudományos adatelemző), és összességében az IT további munkahelyeket teremt. A meglévő szakembereknek új feladatai lesznek, és a cégek irányvonalát erősen befolyásolni fogja, hogy ki mennyire tudja az adatokat információvá alakítani. Mivel a Big Data kérdéskörrel minden valószínűség szerint az IT-részleg fog foglalkozni, így a szervezeteken belüli súlya növekedni fog, kiemelt stratégiai pozíciót fog kapni. Ugyan az adatok kezelését informatikusok végzik majd, viszont a speciális szakmai tudást a többi részleg (pénzügy, marketing stb.) adja hozzá, de a partnerségi viszonyokat az IT fogja majd össze.

A tényleges felhasználás még nyitott kérdés, a cégen belüli kommunikációtól, a vállalati kultúrától, a belső beszélgetésektől függ, illetve a Cisco<sup>296</sup> startup cégeket is támogat. Nagy szükség van olyan speciális gyakorlati képességekkel rendelkező kreatív munkatársakra, akik elképzelik és realizálják az adatokban rejlő lehetőségeket. Egyre nagyobb az igény az olyan informatikai szakemberek képzésére, akik jártasak ezen a speciális területen – például „adattudósokra”, akik képesek az adatokat kreatív és látványos módon szemléltetni az eredményeket, és feltárni azok lehetséges üzleti hatásait.

A trend terjedését leginkább adatbiztonsági kockázatok hátráltathatják, de valós veszélyt jelent az is, hogy ez az egész jelenség a cégek többségénél szűz terület. Adott esetben nincs rá elég pénzük, nincsenek tapasztalataik, megfelelő szakembereik, kész megoldásokat pedig nem mindig tudnak venni. Előfordulhat, hogy egyszerűen túl sok a feldolgozandó adat, azokhoz túl sokféleképpen lehet hozzáférni, a biztonság szavatolására pedig nincs elég pénz. Mivel a kockázatok fele emberi tényező, ezek elméletben kezelhetők, képzéssel megoldhatók.

Fontos kiemelni, hogy az egész akkor hatékony, ha az adatfeldolgozás intelligens eszközökkel már ott lezajlik, ahol keletkezik az adat, különben hatalmas hálózati- és tárolókapacitási igény keletkezik.

A Cisco szakemberei szerint az informatika szolgáltatássá alakul, és a szakembereknek egyre inkább nem a hardverekkel, switchek vagy routerek programozásával kell foglalkoznia, hanem azt a logikát kell értenie, hogy mire képes az informatika és ezt hogyan lehet az üzleti igényekhez igazítani. Az adatközpontoknál most ugyanaz a változás megy végbe, mint ami egy évtizeddel ezelőtt például a PC-nél lezajlott. Régen például foglalkozni kellett a különböző hálózati kártyák beszállításával, de egy integrálási folyamat után ez az alaplapra került, így ma csupán be kell dugnunk a gépbe a kábelt és az interneten vagyunk. Az adatközpontokban ugyanígy integrálódik a hálózat, a szerver és a tárolóegység, egy egységes blokká válnak, csak a nagyságrend tér el. [18]

### Big Data-diploma

A következő két évben világszerte 4,4 millió munkahelyet kell majd betölteni Big Data-elemzésben jártas szakemberekkel. Tudásuk abban a 2500 metropoliszban is nélkülözhetetlen lesz, ahol az IBM okos városfejlesztései futnak. A vállalat ezernél több felsőoktatási intézménnyel indított már képzési együttműködést, Írországból már mesterszakon is lehet *Big Data-diplomát* szerezni. Magyarországon még nincs kifejezetten Big Data-képzés. [19]

<sup>294</sup> data scientist – tudományos adatelemző

<sup>295</sup> IT – Information Technology; információs technológia

<sup>296</sup> Cisco Systems Inc.; amerikai multinacionális vállalat, amely fogyasztói elektronikai berendezéseket, hálózati eszközöket, hang-, illetve kommunikációs technológiát és szolgáltatásokat tervez és értékesít.

## Infrastruktúrafejlesztés

### A hardver és szoftver fejlesztése

A hatalmas adatbázisok és az ahhoz tartozó műveleteknek számtalan technikai feltétele van. Az adatoknak, illetve adatbázisoknak több szakaszból álló, a keletkezéstől az ismételt felhasználásig terjedő életciklusa van, amelyek mindegyikéhez megfelelő technikai háttérrel és támogatást kell biztosítani. Erre az általános asztali gépek, nagy számban használt szerverek technikai és gazdasági szempontból egyre kevésbé alkalmasak. Az adatok feldolgozása növekvő részben megosztott rendszereken, különleges, gigantikus méretű, rengeteg egymáshoz kapcsolt gépből (computer grid) álló adatközpontokban, az úgynevezett „számítástechnikai felhőkben” történik. A nyitott forráskódú szoftverek megjelenése és elterjedése, valamint a felhasználással arányosan fizetett, a felhasználóktól nagy indító beruházást nem igénylő webes szolgáltatások egyre alacsonyabbra kényszeríti a belépési korlátokat, ami olcsóbbá teszi az adatkezelési- és feldolgozási technológiák használatát. Az interneten nyújtott szolgáltatásokhoz és a gigantikus adatbázisok továbbításához megbízható széles sávú kapcsolatok kellenek.

Önmagában a hardver nem elegendő: szoftverre is szükség van, ami alatt nem csak a programokat, hanem az alkalmazott algoritmusokat, adatkezelési és –feldolgozási eljárásokat is értjük. A „Big Data” jelenség lehetőséget nyújt ahhoz, hogy a tudományágak művelői a megfogalmazott kérdésekre matematikai statisztikai, adatbányászati eszközökkel keressenek választ, az adattömegben mintázatok, sokszor nagyon bonyolult, sokváltozós összefüggések után kutassanak tanulásra is képes alkalmazásokkal. Fejlődnek a strukturálatlan adatok feldolgozására szolgáló technológiák is, mint például a szövegbányászat.

A számtalan féle forrásból kapott, sokfajta módon keletkeztetett adat feldolgozásához, kombinálásához és lekérdezéséhez szabványokra és szabványos meta adatokra<sup>297</sup> is szükség van. [7]

A „Big Data” szempontjából nyilvánvalóan fontosak a vállalati informatikai beruházások, amelyek éves nagysága manapság több trillió dollárra rúg és az elmúlt években jelentős emelkedést mutatott. Eközben az információk keletkeztetésének, menedzselésének és tárolásának fajlagos költsége az elmúlt öt évben a hatodára zsugorodott. Az egyik oldalon tehát hatalmas befektetések állnak – a kérdés az, hogy mi van a másik oldalon, ezek a befektetések miként térülnek meg, milyen eredményt hoznak.

A Financial Times a „Big Data” jelenséget *üzleti aranybányának* nevezi. Ennél a kifejezésnél maradva kijelenthető, hogy a bánya működtetéséhez beruházások kellenek. Mint az fentebb látható volt, az informatikai beruházások óriási összegekre rúgnak. [20]

### Digitális Nemzeti Fejlesztési Program

Hazánkban is, mint az *Európai Unió* többi országában, beindult a Digitális Nemzeti Fejlesztési Program, amely a *szupergyors internetből*, a *digitális közösség és gazdaság fejlesztéséből*, az *E-közigazgatási szolgáltatásokból*, és a *digitális kompetenciák fejlesztéséből* tevődik össze. A kormányzat 2018-ra szupergyors internet lefedettséget szeretne kiépíteni az egész ország területén, ami azt jelenti, hogy minden háztartás és intézmény számára legalább *30 Mb/s* sávszélességű internet elérést kíván biztosítani. A program keretében ez a *Nemzeti Digitális Közmű (NDK)* nevet viseli. A program célja a sokkal szofisztikáltabb elérhetőség, az életminőség javítása, a teljes lefedettség biztosítása. A cél tehát nem csak az infrastruktúra végleges kiépítése, hanem a hozzáadott értékkel növelni az életminőség színvonalát a magyar lakosság számára. Ehhez hozzátartozik az E-ügyintézési szabályok kidolgozása, az *„Intelligens város”* állami és önkormányzati, valamint a vállalati szolgáltatási modellek pilot rendszerű alkalmazhatóság megteremtéséhez szükséges alapok letétele. A kormányzati cél és lépések a *30 Mb/s* sávszélesség kialakítása mellett az *adójogszabályok felülvizsgálata*, a *piaci szabályozás felülvizsgálata*, valamint a vonatkozó *építésügyi jogszabályok áttekintése*. [21]

<sup>297</sup> Meta adat jelentése: strukturált adatok az adatokról. Ilyenek a könyvtári katalógusok, a tárgyszó indexek, újabb formájukban a weben található dokumentumok leíró adatai.

### A közigazgatás céljai a fejlesztéssel

Jelen társadalmunkban az információ jelentősen befolyásolja a társadalmi kapcsolatokat. Központi elem az *információ* és ennek csatolmánya az *internet*. A lakosság körében az *E-vásárlás* egyre népszerűbb, mert kényelmes és gyors! A hivatali ügyintézés is egyre népszerűbb, bár a meglévő infrastruktúra és az elérhető hivatali szolgáltatások fejlesztése, bővítése elengedhetetlen és szükséges. A közigazgatási cél, hogy megjelenjenek az interneten, minél szélesebb szolgáltatási palettával, és minél nagyobb számú felhasználót elérve. Amíg a '90-es években az *interneten való jelenlét* volt a közigazgatás célja, addig a 2000-es években az *információs tartalmak szolgáltatása*, bár ez még csak statikus jelleggel. Azonban a 2010-es években az *interaktív lehetőség* mellett az E-közigazgatásban a *fogyasztóbarát felhasználásra* kell törekedni, egyszerűsítéssel, tematikus kialakítással és a célközönség meghatározásával. Úgy, mint fórum és chat szolgáltatások kialakítása, amely elvart az E-közigazgatás oldalain. Továbbá a Facebook megjelenés és mobil applikációk is szükségesek felhasználók minél szélesebb körű eléréséhez az E-közigazgatás elérhetősége szempontjából. Mind emellett a *digitális írástudás* fejlesztése is nagyon fontos, legalább annyira, mint a „Z generáció”, mint a „*digitális bennszülöttek*” információ éhségének kielégítése és a hírközlés ellenőrzésének marginális szerepe. A kormányzat célja a Nemzeti Fogyasztóvédelmi Hatóság által egy „*minta WEB áruház*” kidolgozása, ezzel is segítve a digitális világában való aktív közreműködést.

A magyar IKT világ a nemzetközi mezőnyben elég jó helyen áll, mert számtalan törvény és kormányrendelet szabályozza már a szektort. Az IKT szektornak kulcs szerepe van az Unióban a gazdaság fellendítésében. Magyarország az infrastruktúra fejlesztésében jól áll, komoly előre lépések történtek mind a vezetékes, mind a mobil széles sávú lefedettségében. Ennek ellenére vannak olyan területek az országban, ahova nem ért el a fejlesztés. A *digitális otthon, a smart city, az E-közigazgatás* fejlesztése során minden 30 Mb/s-os sáv szélesség kiépítése kormányrendelet által *kiemelt beruházásnak* számít. Az a szolgáltató, aki ilyen innovációba fog, az kérheti a hatósági támogatást. A „*Sikerkritériumok*” a célok, keretek, szereplők és felelősség négyeseként értendő. A *célok*: az infrastruktúra, a szolgáltatások, kompetencia fejlesztése. A *szereplők*: az állam, a piac, és az EU. A *keretek*: az adók, az EU-s források, és a szabályozás. A *felelősség*: az edukáció, az összhang, és az együttműködés. Magyarországon megkezdődött a megfelelő szabályozási környezet kialakítása. Úgy, mint például az iparági igények és a vélemények figyelembe vétele. A program komplex, mert tartalmazza a hálózat, a kompetencia és a szolgáltatás fejlesztéseit. A programot továbbá, a felhatalmazás, a szoros együttműködés, a konszolidált javaslatok, képviselt egységes iparági álláspont jellemzi. Ezen program az *Európai Digitális Menetrend* és a *Nemzeti Infokommunikációs Stratégia, a Partnerség a Digitális Magyarorszáért* programok és tervek keretében valósulhat meg. Mivel, a *Next Generation Access* (NGA) lefedettség Magyarországon 2013-ban már 75,7% volt. A projekt megvalósulása során a hosszú távú előnyök messze magasabbak, mint azt gondolnánk. Az egész projekt a publikáció írásakor még a tervezési fázisban van, így a részletek kidolgozása is folyamatos. A projekt az előkészítési, a tervezési, a pályáztatási fázisok után, a hálózatok, azon belül is a helyi hálózatok, és azon túl a kimaradt területek építését tartalmazza. Ami a projekt megvalósulásához elengedhetetlen, az a humán erőforrás mellett a pénzügyi erőforrások zavartalan biztosítása. Ez az elmúlt idők legnagyobb hírközlési fejlesztési projektje. A széles sávú fejlesztések megvalósítás a korábbi projektek tapasztalataiból építkezve, a kockázatok folyamatos figyelése és elemzése és az előtervezés szerepének fontossága mellett a hálózatok nyilvántartásán fog alapulni. Nyitott kérdésként merül fel a *Hírközlési Érdekegyeztető Tanács* részéről a közműadó, a vezetékekkel gazdaságosan nem elérhető területek, az *Open Access*, valamint a *kormányzati célú hálózatok* fejlesztése. A *64/2014 EU irányelv*, melynek folyamatban van a honosítása, irányadó az országos elérhetőségű, az állam által biztosított hálózati lefedettség kiépítése szempontjából.

A projekt az *Európai Digitális Menetrend COM(2010)245*; az *Európai Parlament és Tanács 2014/61/EU irányelve* a nagy sebességű elektronikus hírközlő hálózatok kiépítési költségeinek csökkentésére irányuló intézkedésekről, a *Nemzeti Infokommunikációs Stratégia, a 1162/2014. (III. 25.) Korm. határozata* a Digitális Nemzet Fejlesztési Programról, a *1631/2014. (XI. 6.) Korm. határozata* a „Digitális Nemzet Fejlesztési Program” megvalósításáról szóló intézkedések teremtik meg a jogszabályi háttérrel annak, hogy az IKT, mint iparág 2020-ig milyen irányú fejlődésen kell, hogy

keresztülesen, és ez által hozzá járuljon az Európai gazdaság működéséhez. A *Szupergyors Internet Projekt* (SZIP) megvalósulásához feltétlenül szükséges az adminisztrációs akadályok lebontása, a beruházások ösztönzése a stratégiai célok elérése érdekében. A SZIP keretében szükséges az akadálymentesítési munkacsoport létrehozatala. Már hatályba léptek a Digitális Nemzet Fejlesztési Program végrehajtásához kapcsolódó szélessávú hálózatfejlesztési beruházások megvalósításával összefüggő közigazgatási hatósági ügyek nemzetgazdasági szempontból kiemelt jelentőségű ügyé nyilvánításáról és az eljáró hatóságok kijelöléséről szóló 392/2014. (XII. 31.) Korm. rendelet, valamint az elektronikus hírközlési építmények elhelyezéséről és az elektronikus hírközlési építményekkel kapcsolatos hatósági eljárásokról szóló 14/2013. (IX. 25.) NMHH rendelet, továbbá a Nemzeti Média- és Hírközlési Hatóság egyes eljárásainak igazgatási szolgáltatási díjairól és a díjfizetés módjáról szóló 5/2011. (X. 6.) NMHH rendelet, a Nemzeti Média- és Hírközlési Hatóság egyes eljárásainak igazgatási szolgáltatási díjairól és a díjfizetés módjáról. Ezeknek a jogszabályoknak nagy szerepe lesz abban, hogy a jelenleg 400 ezer háztartás lefedettségét meg lehessen oldani, amelyek nem rendelkeznek se vezetékes se mobil internet elérhetőséggel.[21]

### A magyar „smart city”

A következő kérdés a Digitális Menetrendben, hogy *mitől lesz egy város digitális?* Lehet sorolni, hogy az informatika, a hálózatok, a mobil penetráció. Valamint ide lehet sorolni a szolgáltatót, úgymint az állam, az önkormányzatok és a piac, továbbá a fogyasztót, mint állampolgárok, vállalatok. A pénzügyi források elengedhetetlenek a megvalósításhoz és a fenntartáshoz. A „smart city” a *közbiztonság* javítása, a *városkártya* használatának növelése, az *eszközbiztosítás* és a *képzés* keretében kezdődött meg hazánkban több vidéki nagyvárosban, pl. Miskolcon, Nyíregyházán, Salgótarjánban.

A biztonság kérdéseként a digitális városok tekintetében, meg kell említeni, hogy igen is a sok más pozitív hatás mellett a digitalizálásnak megvannak a maga pozitív hatásai a biztonság területén is. Gondoljunk csak arra, hogy az IP kamerák terjedésével, sokkal jobban lehet védeni, biztosítani a közterületeket, amelyek valósidejű képet tudnak közvetíteni a nap 24 órájában, a beépített éjjeli látó funkciójának köszönhetően az adott irányítóközpontba, ahol akár az adott képet tárolni is lehet. Ez a kép a HD felbontásnak köszönhetően, nagyon tudja segíteni a kamera képet használók elemzők munkáját, legyenek azok a világ bármely pontján, köszönhetően az Internet Technológiáknak köszönhetően.

A digitális város, mint „a már megvalósult jövő”, a Wi-Fi szolgáltatást is tartalmazza a lakosság részére. De nem szabad amellet sem elmenni, hogy a „smart city” projekt *bevételt is termel* az azt alkalmazó közigazgatási egység számára. Ami megjelenhet az interneten foglalt szállások növekedésében, amely a felmérések szerint egyes városok esetében akár *60%-ban* bővült. Ez a bővülés a foglalkoztatások számát is növeli. A szakértők egybe hangzó véleménye szerint a ennek a növekedésnek elsősorban fenntarthatónak kell maradnia és a másik cél, hogy ne járjon kiadással az adott település számára. Továbbá a bevételtermelő funkcióval olyan környezetet kell teremteni, ami a helyi vállalkozásoknak fejlődést, árbevétel növekedést, amely a helyi- és központi adóbevételt jelent.

A piaci szereplők számára árbevétel növekedést jelent, ha a felhasználók digitális írástudása növekszik. Az egészségügyi fejlesztések jó irányú fejlesztések lehetnek a smart city program keretében. Ez megjelenhet egy városkártya intézményében. De nem csak arra a plasztik lapra kell gondolni, amit eddig rendszeresen használtak az emberek, hanem ez a szolgáltatás akár a mobil eszközökre is költözhet az új mobil technológiáknak köszönhetően.

A társadalmi gátakat át kell törni a materiális dolgok és a digitális eszközök között. Ameddig ez a gát elszeparálja egymástól az embereket és a digitális világ által nyújtotta szolgáltatásokat, addig a megálmodott jövő csak egy álom marad, amelyet csak nagyon kevesen tudnak a saját jólétük megteremtésére és fenntartására fordítani. A gátakat a költséghatékonysággal lehet felszámolni.

Az állam szempontjából két fontos feladat jelenik meg. Az egyik a közösségi közlekedés elektronikus jegyrendszerének a megteremtése, a másik az állam által nyújtott megoldási csomag, úgymint a parkolási-, a személyszállítási kedvezmények biztosítása, amelyre az adott település rá teheti a saját szolgáltatásait is. Ennek eredményeképpen a „smart city” programot *70%-ban* központilag kell

kiszolgáltatni. Viszont ezekből jól látszik milyen széles a paletta a digitális város program keretén belül. Abból, hogy mi fog kialakulni, kifejlődni, az az adott település igényeinek megfelelően fog alakulni, ugyanis két egyforma „smart city” nincs, ahogyan két egyforma gondolkodású város sincs. [21]

### **A digitális írástudás fejlesztése**

A digitális írástudatlanság társadalmi csökkentése érdekében több program is elindításra került az állam által, ilyen például a *TÁMOP 2.1.2-1/1*, amely az idegen nyelvi és informatikai kompetenciák fejlesztésére irányult, valamint a *digitális tankönyv* program, valamint a *kutatás-fejlesztési* támogatások, és nem utolsósorban a *programozás* oktatásának erősítése. A digitális írástudatlanság csökkentését segíti elő a nagy számban terjedő mobil alkalmazási lehetőségek oktatási célú felhasználása. Úgy, mint a mobil platformokon futó tananyagok és tanítást-, tanulást támogató alkalmazások fejlesztése. A nagyfokú penetrációt segítheti a nyílt oktatási tananyagok alkalmazása, valamint a pedagógusok digitális kompetenciájával kapcsolatos helyzetfelmérés és átfogó vizsgálat lefolytatása, amely a pedagógusok digitális kompetenciájának bővítését szolgálná a képzés területén. Továbbá a pedagógusok továbbképzésének megújításáról sem szabad megfeledkezni, a digitális írástudatlanság csökkentése érdekében.

Sajnos a mai 15 évnél idősebb korosztály tagjai között 40,2%, azaz megközelítőleg 3 millió ember digitálisan nem kompetens ma Magyarországon. De ez a szám sajnos nem azt jelenti, hogy a maradék 7 millió ember az digitálisan kompetens lenne, mert ezek az emberek se használnak olyan tartalmakat, amelyek mindegyike kimeríti a digitális írástudás keretét. Ennek az oka lehet az, hogy az emberek félnek a felfedezni maguknak a digitális világot, nem érzik magukénak, amely azért okoz problémát, mert nem mélyednek el azokban az ismeretekben, nem alakulnak ki azok a digitális készségek, amelyek szükségesek lennének ahhoz, hogy megfelelő szinten tudják használni, kihasználni a digitális világ által nyújtott jólléti javakat. A digitális tartalmak alacsony számú alkalmazásának a társadalom körében még lehet oka az is, hogy a biztonsági kérdés folyamatosan napirenden van. Ezt megoldandóan a *biztonság kérdése* a kompetencia fejlesztésének fontos részét képezi, úgymint a *biztonságos adattárolás* és *biztonságos üzemeltetés* kérdése.

Az e-Magyarország hálózat fejlesztéséhez elengedhetetlen a különböző eszközök és a nagyszámú mentor állomány alkalmazása az informatikai felnőttképzési pontokon. A kormányzat számára a jelenlegi 16 év feletti korosztály jelent nagy kihívást ezen a területen, ugyanis az említettek részére különböző motivációs program beindítása kezdődik meg. Az egyik ilyen a „Netre kész” című program újra élesztése, mellyel 2-3 éven belül megközelítőleg 1 millió ember oktatása kezdődhet meg az idegen nyelv és informatikai ismeretek területén akár e-learning képzéssel. Az oktatásba be kell vonni azokat a személyeket, akiknek az életük része az informatika. [21]

### **A kockázatai és a biztonsági kérdések**

A kritikus infrastruktúra védelmét szabályozói szinten az állam biztosítja a *Kritikus Infrastruktúra Védelem Nemzeti Programról szóló 2080/2008. (VI.30.) Korm. határozattal*. Valamint ezen kormányhatározat alapján kiadott, a létfontosságú rendszerek és létesítmények azonosításáról, kijelöléséről és védelméről szóló 2012. évi CLXVI. Törvénnyel és annak végrehajtásáról szóló 65/2013. (II.8.) Korm. rendelet. A magyarországi kritikus infrastruktúrákat, azon belül is a kritikus információs infrastruktúrát, mely magában foglalja az informatikai közműveket, valamint azok védelmét Magyarország viszonylatában a szabályozása lefedi a védelem kialakításához szükséges területet. [22]

### **Kockázatok csökkentése**

A Big Data jelenség a fentiek alapján azt a benyomást kelti, keltheti, hogy ez az „adat-boom” mindenki számára csak pozitív hatással jár. Sajnos, mint mindennek, ennek is megvannak a maga negatív tulajdonságai.

A szakirodalom által „digitális árnyék”-nak (*digital shadow*) nevezett adatok tömege az, ami egy-egy emberrel kapcsolatban folyamatosan létrejön. Ezeknek a digitális (láb)nyomoknak egy része az „árnyék” gazdája által tudatosan keletkezik, a „digitális életünkben” adódóan. Az „árnyék” másik része azonban ismeretlen a tulajdonos számára, nem tud a létezéséről. Ennek okán a személyes

adatok védelme, napjainkban egyre nagyobb hangsúlyt kap, köszönhetően a világ, azon belül is az európai trendeknek, melynek hatására Magyarországon is elfogadásra került a 2011. évi CXII. törvény az információs önrendelkezési jogról és az információszabadságról.

Ez a törvény az információs önrendelkezési jog és az információszabadság biztosítása érdekében, a személyes adatok védelmét, valamint a közérdekű és a közérdekből nyilvános adatok megismeréséhez és terjesztéséhez való jog érvényesülését szolgáló alapvető szabályokról, valamint az ezen szabályok ellenőrzésére hivatott hatóságról az Alaptörvény végrehajtására, az Alaptörvény VI.<sup>298</sup> cikke alapján készült.

A fenti törvény és az általa életre hívott Nemzeti Adatvédelmi és Információszabadság Hatóság által magyarországi viszonylatban biztosítja a személyes adatok védelmét.

A felhalmozott adatvagyon védelme, megóvása érdekében az adatvagyon monopolizálása is szükséges. Az adat olyan erőforrás, amelynek a másolása, többszörözése minimális direkt költséggel jár, ezzel szemben egy, a hasznosíthatóság szempontjából multifunkciós adatbázis reprodukciója nagy tőkeinvestációt igényelhet. Ennek okán a magánérdek szembe kerülhet a közösségi érdekekkel. Ennek az anomáliának a megoldása az adatbázisok államosítása lehetne, de ismert a különbség vállalati szektor és az állam kreativitása és gyorsasága között. [7][23]

### Biztonsági feltételek kialakítása

A BalaBit<sup>299</sup> syslog-ng nevű megoldása az informatikai rendszerekben keletkező naplőüzenetek gyűjtését és feldolgozását valósítja meg. Jelenleg azonban a piacon található szoftverek az operációs rendszerek és alapvető rendszeralkalmazások naplőüzeneteire fókuszálnak, a felhasználók fő motivációja pedig az adatvédelmi szabályozásoknak való megfelelés, informatikai biztonsági kontroll és az üzemeltetés támogatása. Az alkalmazásokból való információgyűjtés viszont egy egészen más megközelítést igényel. Elég, ha egy ERP<sup>300</sup> rendszerben, online bankban vagy közösségi portálon végzett felhasználói aktivitásra gondolunk. [24]

### Összefoglalás

A Big Data, mint a jövő *“olaja”*, vagy *“aranybányája”* mit sem ér, hogyha azt nem tudjuk biztonságos módon elérni, tárolni. Látható, hogy ahogy maga az informatika, és ezáltal a létfontosságú információs közművek, mint a kritikus információs infrastruktúra részei, biztosítják a jólétünket, melynek a XXI. század második évtizedében mára elengedhetetlen részét képezi a digitális jólétnek. A digitális jólét a szükséges és kellő mennyiségű információ elektronikus úton történő feltétel-nélküli elérését jelenti. Az üzleti világ mellett a civil társadalomnak, benne magának az állampolgárnak is nap, mint nap döntések sorozatát kell meghoznia. A minél pontosabb döntések meghozatala érdekében a gyors, pontos, naprakész, sokféle megjelenésű és nagy mennyiségű információra van szükség. Tehát ez esetben a Big Data-t jellemző 4V modell érvényesül. A fenti publikációban bemutatott Big Data fogalmi és gyakorlati alkalmazási lehetőségei azt bizonyítják, hogy az ezt követő életünk jelentős részét fogja képezni. Gondolva itt a gazdasági növekedésről, mely a sok formában megjelenhet, úgy, mint a foglalkoztatás növekedése, új szakmák kialakulása, ezáltal az oktatási struktúra átalakulása. Valamint az egészségügy modernizációja, a betegségek gyorsabb

<sup>298</sup> Magyarország Alaptörvénye IV. cikk:

- (1) Mindenkinek joga van a szabadsághoz és a személyi biztonságához.
- (2) Senkit nem lehet szabadságától másként, mint törvényben meghatározott okokból és törvényben meghatározott eljárás alapján megfosztani. Tényleges életfogytig tartó szabadságvesztés csak szándékos, erőszakos bűncselekmény elkövetése miatt szabható ki.
- (3) A bűncselekmény elkövetésével gyanúsított és őrizetbe vett személyt a lehető legrövidebb időn belül szabadon kell bocsátani, vagy bíróság elé kell állítani. A bíróság köteles az elé állított személyt meghallgatni és írásbeli indokolással ellátott határozatban szabadlábra helyezéséről vagy letartóztatásáról haladéktalanul dönteni.
- (4) Akinek szabadságát alaptalanul vagy törvénytelenül korlátozták, kárának megtérítésére jogosult.

<sup>299</sup> BalaBit – a világ egyik vezető IT-biztonsági szoftverfejlesztő vállalata

<sup>300</sup> ERP – Enterprise Resource Planning System; integrált vállalatirányítási rendszer

pontosabb diagnosztizálása, gyógyítása, mely a gazdasági növekedés esetében nem elhanyagolandó eleme. A mezőgazdaságot is megemlítve, és ezáltal az élelmiszeripar hatékonyabb működését, mely közvetve és közvetlenül segíti a gazdasági növekedést. Az adatok gyors és biztonságos átvitelét az állam a bemutatott Digitális Nemzeti Fejlesztési Program keretében kívánja megteremteni. Jól látható, hogy ez a kritikus infrastruktúra fejlesztését jelenti. Ahhoz, hogy a digitális jólét megvalósulhasson, ki kell alakítani az elérés lehetőségét, az elérhetővé kell tenni a különböző szolgáltatásokat, megfelelő tartalommal ellátva. Továbbá, ahhoz, hogy mindez felesleges és kihasználatlan beruházás legyen, mindenképp szükséges a digitális írástudás növelése, amely a funkcionalitást sem nélkülözi. Mindez akkor hatékony, hogyha megfelelően vannak védve az adataink és az a közeg ahol az információink tárolódnak és továbbítódnak.[25]

## Referenciák

- [1] Steven Rosenbush, Michael Totty: How Big Data Is Changing the Whole Equation for Business The Wall Street Journal, March 10, 2013, <http://www.wsj.com/news/articles/SB10001424127887324178904578340071261396666?mg=reno64-wsj&url=http%3A%2F%2Fonline.wsj.com%2Farticle%2FSB10001424127887324178904578340071261396666.html>, letöltve: 2014. március 18.
- [2] IDC Nagy Adat és Üzleti Analitikai Fórum 2014, Budapest, 2014. 04. 24., <http://idchungary.hu/hun/events/54405-idc-big-data-and-business-analytics-forum-2014/7-overview>, letöltve: 2014. május 10.
- [3] prof. Patrice Koehl: Big Data Why it matters, előadás, Department of Computer Science, UC Davis Genome Center, 2013. szept. 5., [https://www.youtube.com/watch?v=\\_HI5pLCFbu0&list=PLKJRT9JmiGVcibyh1Fw7CjMdlpq2wC2uk](https://www.youtube.com/watch?v=_HI5pLCFbu0&list=PLKJRT9JmiGVcibyh1Fw7CjMdlpq2wC2uk), letöltve: 2014. június 8.
- [4] SI-prefixumok, <http://fizika.qwqw.hu/?modul=oldal&tartalom=1112407> letöltve: 2014. május 20.
- [5] Bináris prefixumok, [http://wiki.hup.hu/index.php/Bin%C3%A1ris\\_prefixumok](http://wiki.hup.hu/index.php/Bin%C3%A1ris_prefixumok) letöltve: 2014. május 20.
- [6] Bitport: Nagy problémákat hozhat a Nagy Adat, 2013.05.14., <http://bitport.hu/trendek/nagy-problemakat-hozhat-a-nagy-adat> letöltve: 2014. június 20.
- [7] Adatrobbanás a vállalatvezetésben, Computerworld, 2011. november 10., <http://computerworld.hu/cio/adatrobbanas-a-vallalatvezetesben-.html> letöltve: 2014. június 15.
- [8] Piackutatás blog: Nagy Adat vagy nagy cirkusz?, 2014.02.10., [http://piackutatas.blog.hu/2014/02/10/nagy\\_adat\\_vagy\\_nagy\\_cirkusz](http://piackutatas.blog.hu/2014/02/10/nagy_adat_vagy_nagy_cirkusz) , letöltve: 2014. június 20.
- [9] Prím Online: Hét lépés szükséges a big data kiaknázásához, 2013. április 15., [http://hirek.prim.hu/cikk/2013/04/15/het\\_lepes\\_szukseges\\_a\\_big\\_data\\_kiaknazasahoz](http://hirek.prim.hu/cikk/2013/04/15/het_lepes_szukseges_a_big_data_kiaknazasahoz) letöltve: 2014. június 15.
- [10] Piac és Profit: Nem alaptalan a „nagy adat” körüli felhajtás, 2012. október 26., <http://www.piacprofit.hu/infokom/nem-alaptalan-a-nagy-adat-koruli-felhajtas/> letöltve: 2014. június 20.
- [11] hvg.hu: Adatrobbanás a NAV-nál, 2013. március 03., [http://hvg.hu/tudomany/20130303\\_adatrobbanas\\_a\\_nav\\_nal](http://hvg.hu/tudomany/20130303_adatrobbanas_a_nav_nal) letöltve: 2014. június 20.
- [12] Computerworld: Mezőgazdasági Big Data, 2014. április 21., <http://computerworld.hu/computerworld/mezogazdasagi-big-data.html> letöltve: 2014. június 20.
- [13] hvg.hu: Így védekeznek az árvíz ellen a hollandok, 2013. július 11., [http://hvg.hu/tudomany/20130711\\_igy\\_vedekeznek\\_az\\_arviz\\_ellen\\_a\\_hollandok](http://hvg.hu/tudomany/20130711_igy_vedekeznek_az_arviz_ellen_a_hollandok) letöltve: 2014. június 16.

- [14] BBC Horizon 2013 The Age of Big Data, 2013. aug. 22.  
<http://www.youtube.com/watch?v=CO2mGny6fFs&feature=share&list=PLKJRT9JmiGVcibyh1Fw7CjMdlpq2wC2uk> letöltve: 2014. június 20.
- [15] Maverick: 83. Big Data és DNS: az orvostudomány forradalma, 2013. dec 16.,  
[http://voyager.blog.hu/2013/12/16/83\\_big\\_data\\_es\\_dns\\_az\\_orvostudomany\\_forradalma](http://voyager.blog.hu/2013/12/16/83_big_data_es_dns_az_orvostudomany_forradalma)  
letöltve: 2014. június 21.
- [16] hvg.hu: Ez lesz a legmenőbb ágazat pillanatokon belül, 2013. október 02.,  
[http://hvg.hu/kkv/20131002\\_Okos\\_varosokat\\_terveznek\\_a\\_jovo\\_szakember](http://hvg.hu/kkv/20131002_Okos_varosokat_terveznek_a_jovo_szakember) letöltve: 2014. június 20.
- [17] Harangi László: Az ország minden adatbázisát egyesítenék a britek, 2014. augusztus 5.,  
<http://pcworld.hu/iskolakezdes/az-oroszag-minden-adatbazisat-egyesitenek-a-britek.html>  
letöltve: 2014. június 15.
- [18] SG.hu: Big Data: sok a kihasználatlan lehetőség, 2013. április 21.  
[http://sg.hu/cikkek/96813/big\\_data\\_sok\\_a\\_kihasznalatlan\\_lehetoseg](http://sg.hu/cikkek/96813/big_data_sok_a_kihasznalatlan_lehetoseg) letöltve: 2014. június 16.
- [19] hvg.hu: Ez lesz a legmenőbb ágazat pillanatokon belül, 2013. október 02.,  
[http://hvg.hu/kkv/20131002\\_Okos\\_varosokat\\_terveznek\\_a\\_jovo\\_szakember](http://hvg.hu/kkv/20131002_Okos_varosokat_terveznek_a_jovo_szakember) letöltve: 2014. június 20.
- [20] Világtudomány.hu: A Nagy Adat, mint aranybánya...,  
<http://vilagtudomany.hu/index.php?data%5Bmid%5D=7&data%5Bid%5D=1627&a-nagy-adat-mint-aranybanya> letöltve: 2014. június 16.
- [21] Digitális Nemzet, digitális fogyasztók „Rajta vagyunk a bittérképen” c. konferencián elhangzottak alapján, Budapest, 2015.03.24.
- [22] Andras Kerti, Zoltan Nyikes (2015): Overview of Hungary Information Security, The Issues of the National Electronic Classified Material of Transmission, SACI 2015 10th Jubilee IEEE International Symposium on Applied Computational Intelligence and Informatics PROCEEDINGS, (pp. 327-334.), Timișoara, Romania May 21–23, 2015, ISBN: 978-1-4799-9910-1
- [23] 2011. évi CXII. törvény az információs önrendelkezési jogról és az információszabadságról
- [24] hvg.hu: 300 millió a felhasználói naplózás fejlesztésére, 2013. április 09.,  
[http://hvg.hu/tudomany/20130409\\_300\\_millio\\_a\\_felhasznaloi\\_naplozas\\_fejles](http://hvg.hu/tudomany/20130409_300_millio_a_felhasznaloi_naplozas_fejles) letöltve: 2014. június 16.
- [25] Andras Kerti, Zoltan Nyikes (2015): The Information Security Standards Dealing with Hungarian and International Organizations; Proceedings of the 3rd international conference and workshop Mechatronics in Practice and Education – MECHEDU 2015; (pp: 104-114), Subotica, Serbia May 14-15, 2015; ISBN 978-86-918815-0-4
- [26] A katasztrófavédelmi kommunikációs rendszer jellemzői II. VÉDELEM KATASZTRÓFAVÉDELMI SZEMLE (ISSN: 2064-1559) XIII: (4) pp. 17-18. (2006