

A BIG DATA NEMZETBIZTONSÁGI HASZNOSULÁSÁNAK LEHETŐSÉGEI

The potential of Big Data for National Security

Eck Gábor¹

Absztrakt: Adatok vezérelte világunkban a Big Datának mélyreható gazdasági, társadalmi és nem utolsósorban biztonsági hatásai vannak. A Big Data segítségével létrehozott új strukturális, információs és technológiai környezetet felügyelő rendszerek magas szinten képesek lehetnek ellenőrizni, irányítani és akár befolyásolni a fenti folyamatokat. Az eddig soha nem látott adatbőség, az állandó digitális jelenlét és technológia új kihívások elé állítja a nemzetbiztonságáért felelős szervezeteket, át kell gondolni az eddig bevett és elégségesnek bizonyult módszereket, eszközöket. Biztonsági szakemberek kezdik felismerni a Big Datában és a hozzákapcsolódó technológiai megoldásokban rejlő lehetőségeket, illetőleg az ezekhez kapcsolódó esetleges veszélyeket. A biztonságért felelős szervezetek feladata, hogy Big Data nyújtotta lehetőségeket úgy integrálja rendszereibe, hogy az megfeleljen a társadalmi, etikai és biztonsági elvárásoknak.

Kulcsszavak: Big Data, nemzetbiztonság, információ, technológia, infokommunikáció

Abstract: In our data-driven world, Big Data has profound economic, social and, not least, security implications. The new structural, information and technological systems created by Big Data to oversee the environment could have the potential to monitor, control and even influence these processes at a high level. The unprecedented abundance of data, the constant digital

¹ Eck Gábor, Nemzeti Közszolgálati Egyetem, Hadtudományi Doktori Iskola, doktorandusz hallgató.

ORCID azonosító: <https://orcid.org/0000-0002-4381-2135>

A szerző további munkásságát lásd a Magyar Tudományos Művek Tára oldalán: <https://m2.mtmt.hu/gui2/?type=authors&mode=browse&sel=10083198&view=pubTable>

E-mail cím: eck.gabor@icloud.com

presence and technology are posing new challenges for organisations responsible for national security, and the methods and tools that have been tried and tested and have proved sufficient up to now need to be rethought. Security experts are beginning to recognise the potential of Big Data and related technologies, and the threats they may pose. It is up to security organisations to integrate Big Data into their systems in a way that meets social, ethical and security requirements.

Keywords: Big Data, national security, information, technology, infocommunication

BEVEZETÉS

Ma egy olyan világban élünk, ahol az adat már nem csupán szűk forrásból érkezik hozzánk, hanem teljesen körbeveszi az életünket, befolyásolja azt. Sok esetben mi magunk szolgáltatunk információt, nem egyszer tudtunk nélkül, de legalább is nem kifejezetten szándékosan. Elegendő csak a magunknál tartott okoseszközeink által továbbított, a lokációkra vonatkozó adatokra, akár a közterületeken elhelyezett kamerák felvételeire, vagy a különböző pontgyűjtő-hűségprogramok keretein belül megosztott információkra gondolni. Természetesen nem szabad elfeledkeznünk a közösségi médiáról és az ott elérhető adatok jelentőségéről és mennyiségéről, hiszen az internet használók 93% használja valamelyik közösségi média felületét. 2022 áprilisa és 2023 áprilisa között 150 millió új felhasználó csatlakozott valamelyik közösségi platformhoz, a Föld lakosságának 59,9% jelen van az online közösségi felületeken.²

Az adatközpontú digitális technológia, köszönhetően a tudomány és a technológia fejlődésének, fokozatosan a gazdasági fejlődés új hajtóerejévé vált, ezzel új lendületet adva a digitális területek további fejlődésének.³ A technológiai innovációkban rejlő lehetőségek, mint például a személyre szabott távoroslás, a távmunka, vagy a mobil eszközök által nyújtott új és újabb lehetőségek egyre több ember boldogulását segíti. A növekvő adatforrások, valamint a tárolási és feldolgozási technológiák változásai folyamatos kihívások elé állítják a fejlesztőket, az elemzőket és hozzájuk

² <https://datareportal.com/global-digital-overview#:~:text=Roughly%204.66%20billion%20people%20around,over%20the%20past%20twelve%20months%20>

³ NUCCIO, GUERZONI, 2019.

csatlakozva azon nemzetbiztonsági szakembereket, aki az adatok nemzetbiztonsági érdekből való védelmén és/vagy hasznosulásán tevékenykednek.

A globális internethasználók száma 2022 elején 4,95 milliárdra nőtt, ami a világ teljes népességének 62,5 százaléka, a közösségi platformokat 2021-ben több mint 4.26 milliárd ember használta, 2026-ra ez a szám meg fogja közelíteni a 6 milliárdot.⁴ A létrehozott, rögzített, másolt és felhasznált adatok teljes mennyisége az előrejelzések szerint gyorsan növekedni fog, mérete 2022-ben 97 zettabájt, ami várhatóan 2025-re 181 zettabájtra emelkedik.⁵ A négy legnagyobb tárhely- és szolgáltató cég, mint a Microsoft, az Amazon, a Facebook és a Google 1,2 millió terrabájt mennyiségű online tevékenységből származó adatot tárol.⁶

A Big Data megjelenése jelentős változást hozott a nemzetbiztonság területén is. A nagy mennyiségű adatok elemzése lehetővé teszi a biztonsági szolgálatok számára, hogy azonosítsák azon területeket, ahol a biztonsági kockázatok a legnagyobbak, és hogy hatékonyabban kezeljék a váratlan eseményeket. Az adatelemzés továbbá segíti a nemzetbiztonsági szervek munkáját az előrejelzésben és a prevencióban is, ezzel csökkentve a biztonsági kockázatokat.⁷

A nagy mennyiségű adat gyűjtésének, tárolásának és elemzésének képességével a Big Data értékes támogatást nyújthat, amely javíthatja a nemzetbiztonsági erőfeszítéseket. Az összegyűjtött hatalmas mennyiségű adat miatt azonban aggályok merülnek fel az adatok biztonságával és titkosságával, valamint az elemzésekhez használt adatok pontosságával kapcsolatban. Ez a tanulmány megvizsgálja a Big Data és a nemzetbiztonság közötti kapcsolatot, kiemelve azokat a lehetőségeket, kihívásokat, amelyek meghatározzák egy modern és naprakész információs rendszer működését, továbbá azt, hogy a magánszektorban kifejlesztett és alkalmazott adatelemzési technológiák és adatfeldolgozó technikák milyen módon hasznosulhatnak a nemzetbiztonsági szolgálatok információ feldolgozó, elemző munkája során.

A téma kutatása során áttekintettem a Big Data fogalmi kereteit vizsgáló tudományos munkákat,⁸ figyelemmel arra, hogy a hétköznapijainkba belopózott a Big Data és az azzal kapcsolatos folyamatok, de mindenki a saját

⁴ <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>

⁵ <https://www.statista.com/statistics/871513/worldwide-data-created/>

⁶ <https://www.sciencefocus.com/future-technology/how-much-data-is-on-the-internet/>

⁷ HOLLIN, 2015.

⁸ DE MAURO, GRECO, GRIMALDI 2016; LIMAJ, BILALI 2018.

szakterületén fontos jellemzőket helyezi előtérbe, általánosan elfogadott meghatározása nem ismert. Rengeteg tanulmány fókuszál a Big Data fogalmára, ezek feldolgozása során kifejezetten támaszkodtam Maddalena Favaretto, Eva De Clercq, Christophe Olivier Schneble, Bernice Simone⁹ kutatására. A Big Data általános jellemzőinek, azok összefüggéseinek vizsgálatakor segítségül hívtam Douglas Laney¹⁰ munkásságát, aki elemzésében meghatározza a Big Data vonatkozásában az adatok keletkezésének, terjedésének, forrásainak jellemzőit. Laney ezzel megteremtett egy olyan alaprendszert, melyet a változó világhoz való igazodás jegyében kiegészítve, de a mai napig használunk.

Az cikk első szakaszában áttekintem a Big Data jellemzőit, azok fejlődését, alakulását, a benne rejlő, a döntési rendszereket támogató lehetőségeket. Ezek felhasználásával kísérletet tettem egy, a nemzetbiztonsági szféra részére is alkalmazható és elfogadható Big Data meghatározás alkotására. A második szakaszban bemutatom a Big Data és a nemzetbiztonság lehetséges kapcsolódási pontjait, ahol a nagy mennyiségű adatban rejlő lehetőségek képesek támogatni a biztonsági elvárások teljesítését. A harmadik szakaszban a Big Data nemzetbiztonsági célú alkalmazásának során megjelenő esetleges kihívásokra törekedtem rávilágítani. A tanulmányt a konklúziók kifejtése zárja.

I. BIG DATA

A bevezetésben említett hatalmas mennyiségű adatok, információk, emberi aggyal szinte felfoghatatlanok, de feldolgozásuk, értelmezésük, hasznosításuk sem tartozik a hétköznapi feladatok közé. Rengeteg információt olvashatunk a Big Data kezelésével, feldolgozásával, hasznosításával kapcsolatban ezért érdemes egy rövid áttekintést adnunk a Big Datáról.

A Big Data kifejezéssel találkozhatunk gazdasági, kereskedelmi területeken, a különböző tudomány területeken folytatott kutatások, egészségügyi szolgáltatások igénybevétele, de érezhetjük hatását egyszerű, hétköznapi tevékenységeink során is. A Big Data sokak számára az eddig megoldhatatlan kihívásokra adandó válaszok lehetőségét rejt magában, ugyanakkor kiterjedt kutatások¹¹ sem voltak képesek egy olyan több

⁹ FAVARETTO, DE CLERCQ, SCHNEBLE, SIMONE 2020.

¹⁰ LANEY 2001.

¹¹ LEONELLI 2020.

tudományra kiterjedő, közös fogalmi meghatározást találni, ami mindenki számára kielégítő volna.¹²

A pontos, tudományos igényességű definíció hiányában az alaposabb megértés érdekében segítségünkre lehet a Big Data jellemzőinek áttekintése.

Laney¹³ szerint a big data fő jellemzője a 3V, azaz Volume (mennyiség), Velocity (sebesség) és Variety (változatosság), ahol a mennyiség az adatok keletkezésére, a sebesség az adatok terjedésére, míg a változatosság az adatok sokszínű forrásaira vonatkozik. Az elmúlt évek során a klasszikus 3V további jellemzőkkel egészült ki, mint a Variability (változékonyság), a Virtual (virtualitás), a Veracity (hitelesség), végül, de semmiképpen sem utolsósorban a Value (érték). Álláspontunk szerint a hetedik V, az érték, ami igazán megmutatja a Big Data jelentőségét, hiszen a kinyerhető érték, az adatok feldolgozása, strukturálása és kezelése következtében előkerülő plusz mondanivaló, amely egy gyorsabb, egy több forrású és különböző formátumokból, architektúrákból kinyert extra tartalom. Ezek alapján kijelenthető, hogy a Big Data valódi értékét nem az azt alkotó adatok képviselik, hanem az azokból kinyert információ többlet.

A kinyert extra tartalom vagy tudás jelentős mértékben képes befolyásolni, átalakítani a hétköznapi folyamatokat, így hatással van az üzleti világban zajló események megértésében, a pénzügyi, kereskedelmi irányvonalak alakításában. Elegendő az ügyfelek szokásait, viselkedését, választási módjait összegző adatok elemzése során kinyert minták felismerésére, modellek kialakítására gondolnunk. A Big Datához kapcsolódó értékek jelentős eredményeket képesek felmutatni az egészségügyi helyzet optimalizálásában, a tudományos élet, a kutatások támogatásában.¹⁴ Gondoljunk csak az orvosi képzőképző diagnosztikai eljárások gépi elemzésére vagy a tudományos kutatásokat kísérő nagymennyiségű adatok feldolgozáshoz alkalmazott számítási kapacitásokra.

A fentiek figyelembevételével megállapíthatjuk, hogy a Big Data olyan különböző forrásokból összegyűjtött, strukturált, félig strukturált és strukturálatlan adatok, információk gyűjteménye. Ennek a hatalmas mennyiségű és folyamatosan, gyorsan bővülő adathalmaznak a kezelése hagyományos adatfeldolgozó eszközökkel nehézkes. Ezek az adatok felhasználhatók az analitikai alkalmazásokban, például a prediktív modellezésben és a gépi tanulásban, a kinyert többlet információ, tudás nagy

¹² FAVARETTO, DE CLERCQ, SCHNEBLE, SIMONE 2020.

¹³ LANEY 2001.

¹⁴ MAGYARORSZÁG MESTERSÉGES INTELLIGENCIA STRATÉGIÁJA 2020-2030.

mértékben képes támogatni a különféle döntés előkészítő, támogató folyamatokat, rendszereket.

II. A BIG DATA ÉS A NEMZETBIZTONSÁG

A nagy mennyiségű adatok kiemelt fontossággal bírnak a biztonságos környezet, a nemzetbiztonsági aspektusokkal rendelkező eljárások menetében is, átalakítják a hírszerzési ismereteket, tevékenységeket és szervezeteket, hatással vannak a nemzetbiztonsági fenyegetésekre, valamint átalakítják a jelenlegi és jövőbeli hírszerzési tevékenységet és döntéshozatalt. A jövőben sokkal nagyobb hangsúly fog kerülni az adatgazdálkodásra, az információ áramlásának és megosztásának jelentősége tovább emelkedik, valamint az információ technológiai szerepe is emelkedni fog a nemzet biztonságát garantáló tevékenységek során.

A nemzet biztonságáért felelős szervezetek az adatok növekvő száma egyre összetettebb kihívások elé állítja. Munkájuk során információkat gyűjtenek, kezelnek és elemeznek a lehető legtöbb forrás bevonásával a biztonság érdekében.

Ennek során meg kell küzdeniük a nagy mennyiségű adatok jelentette kihívásokkal. Rövid idő alatt rengeteg új típusú adat keletkezik, melyek jelentős mértékben képesek a hagyományos adatbázis-kezelő rendszereket, elemző programokat és folyamatokat túlterhelni. Ezen problémák egy részével a magánszektor a saját maga területén már megküzdött (direkt marketing, ügyfél elégedettség mérése, stb).¹⁵ Az általuk kifejlesztett, alkalmazott adatelemzési technológiák és adatfeldolgozó technikák ígéretesek lehetnek a nemzetbiztonsági szolgálatok információ kinyeréssel kapcsolatos akadályainak leküzdésében.

¹⁵ HOFACKER, MALTHOUSE, SULTAN 2016.

2.1. Big Data, mint lehetőség

A rendelkezésre álló adatokkal kapcsolatos tevékenység mindig is jelen volt a biztonsági szférában és folyamatosan fejlődött felhasználva a rendelkezésre álló technikákat és technológiákat. A papíralapú adatfeldolgozásokat felváltották a számítógépes rendszerek, a pusztán az emberi agy kapacitására hagyatkozó elemzéseket, értékeléseket, ma már különböző értékelő, vizualizáló, kapcsolatiháló elemző programok segítik. A napjainkra jellemző megnövekedett adatmennyiségekkel való célirányos és hatékony gazdálkodás érdekében kiemelt figyelmet kell fordítani a rendelkezésre álló újszerű opciók integrálására a meglévő rendszerekbe. Érdeemes áttekinteni azokat a Big Data rendszereiben rejlő lehetőségeket, melyek az adat vezérelt világunkban a nemzetbiztonságért felelős szervezetek hatékonyságának fokozása érdekében gyorsítani, pontosítani tudják az információhoz jutást. A biztonság garantálása szempontjából a Big Data eszközök egyik fő és talán legfontosabb hozzájárulása a hírszerzési adatok feldolgozása során érhető el:

1. Kiemelt jelentőséggel bírnak az osztályozási lehetőségek melynek során az adatbázisokból elérhető entitások különböző besorolási kategóriákhoz rendelnek hozzá. A besorolási kategóriák alkalmazásával jól azonosíthatóvá válik az adott objektum/entitás biztonsági szempontú osztályozása, ezen keresztül a kapcsolódó tevékenység kockázati szintje pontosabban értékelhető, azonosítható.

2. A változatos közösségi média felületeken felbukkanó adatok átvizsgálása szinte kivitelezhetetlen a hasonlósági jellemzők alapján folytatott csoportosítás, klaszterezés¹⁶ nélkül. Ezen hatalmas mennyiségű adatok elemzése, téma szerinti csoportosítása hozzá segíthet a közösségi média használóinak tipizálásához.¹⁷ Az így kinyert információk segíthetnek eldönteni, hogy a nemzetbiztonsági szempontból látókörbe került közösségi platform használó rejt e magában valós biztonsági kockázatot, vagy csak sajátos módon érvényesíti a véleménynyilvánításhoz való jogát.

3. A nagymennyiségű adatok lehetőséget biztosítanak arra, hogy adatbányász eljárások segítségével azokban minták váljanak felismerhetővé. Ilyen lehet például a különféle forrásokból származó információs jelentések összevetése.¹⁸ Ezek a látókörbe került személy múltbéli magatartásáról,

¹⁶ itt klaszterezésen adatok, entitások hasonlóságokon alapuló csoportosítását kell érteni.

¹⁷ ALSAYAT, EL-SAYED 2016.

¹⁸ SHAHRIAR HOSSAIN, BUTLER, BOEDIHARDJO, RAMAKRISHNAN, 2012.

kapcsolatairól vagy meggyőződéseiről szóló adatok, továbbá a pénzügyi, oktatási profiljainak elemzésén keresztül keletkezett információk. A minták használatára és azok felismerésére jó példa a New York állam által használt bűnüldözési és terrorizmus ellenes jelző rendszer. A rendszer a korai figyelmeztető jelzések érdekében ellenőrzi az esetleges célszemélyek utazási, szállodai, az általuk használt gépjárművekkel kapcsolatos szokásait.¹⁹

4. A mintáktól való eltérés vizsgálata választ adhat arra, hogy az azonosított adat, elem vagy tevékenység megfelel-e az eddig ismert mintáknak.

5. Az önállóan alkalmazkodni és tanulni képes, gépi tanulási eljárások a feldolgozott és az újonnan megjelenő adatok szintetizálásával támogatják a biztonságot. Különböző adatbázisokhoz hozzáféréssel rendelkező rendszerek lekérdezéseken keresztül személyeket azonosíthatnak, kapcsolatokra világíthatnak rá, új mintákon keresztül új attitűdökre hívhatják fel a figyelmet.

Kijelenthető, hogy a nemzetbiztonsági szervezetek elemzési tevékenységén a Big Data alkalmazások változtattak,²⁰ változtatni fognak. Ez nem kétséges, de azt nehéz prognosztizálni, hogy a hagyományos adatfeldolgozói, elemző rendszereket mikor és milyen határfokon fogja teljesen kiváltani. Ugyanakkor elvitathatatlan az, hogy a magánszektorban elért eredmények felhasználása nélkül a sikerekhez vezető út nemcsak jóval hosszadalmasabb lesz, de sokkal rögzesebb is.

III. KIHÍVÁSOK

A Big Data rendszerek használata lehetővé teszi, hogy a nemzetbiztonsági szervezetek sokkal hatékonyabban gyűjtsék, elemezzék és hasznosítsák a rendelkezésre álló információkat, ezzel támogatva a biztonság magasabb szintre emelését. Segíti az adatok gyorsabb feldolgozását és az azonnali értelmezését, ami fontos az olyan helyzetekben, amikor az idő kritikus szerepet játszik. Az adatelemzésnek köszönhetően a rendszerek képesek észlelni a szokatlan tevékenységeket, mint például a terrorista csoportokat, a kibertámadásokat és más biztonsági fenyegetéseket. Ugyanakkor számos olyan kihívással is szembe kell néznünk a Big Data

¹⁹ <https://publicintelligence.net/new-york-state-law-enforcement-terrorism-indicators-reference-card/>

²⁰ LYON, MURAKAMI 2020.

alkalmazása során melyek a biztonsági kérdéseken felül egyéb jogi, társadalmi, technológiai és gazdasági kérdéseket is feszegetnek.

3.1. Hamis pozitív és hamis negatív eredmények

A Big Data elemzési technológiák alkalmazása révén a nagy mennyiségű adatokból kinyert információk lehetővé teszik a nemzetbiztonsági hatóságok számára, hogy előre jelezzék a lehetséges fenyegetéseket és azonosítsák a potenciális bűnözőket. Azonban az ilyen technológiák használata során gyakran előfordulhat, hogy a rendszer tévesen azonosít valakit, ami komoly jogi és társadalmi következményekkel járhat. 2013-ban a bostoni futóversenyen bekövetkezett robbantás helyszínén és az odavezető utakon elhelyezett köztéri kamerák, ipari kamerák, valamint a nézők által rögzített felvételek elemzése során több hibás azonosítás is történt. Először egy afroamerikai férfit hoztak összefüggésbe a robbantással, később egy szaúdi férfit tartoztattak le,²¹ mire sikerült a valódi elkövetőket azonosítani.

Az adatelemzés során a rendszerek olyan mintákat és viselkedéseket keresnek, amelyeket összefüggésbe hozhatnak a biztonsági fenyegetésekkel. Azonban, ha a rendszer rossz mintát talál, az hamis pozitív, vagy negatív eredményekhez²² vezethet, ami azt jelenti, hogy egy olyan tevékenységet észlel, amely nem jelent valós fenyegetést, vagy nem ismeri fel a veszélyhelyzetet.

A hamis eredmények oka lehet a nagy mennyiségű adatok rendszerezésében és elemzésében alkalmazott algoritmusok hibája. Ez akkor tud előfordulni amikor nincs megfelelően általánosítva az algoritmus és túlilleszti az adatokat, amelyek alapján betanították.²³ Az algoritmusoknak olyan sémákat kell felismerniük, amelyek azonosak vagy hasonlóak a korábbi bűncselekmények vagy terrorista tevékenységek sémáival. Az algoritmusok azonban tévesen azonosíthatnak olyan mintákat is, amelyek csak véletlenül hasonlítanak a korábbi sémákra, ezért az algoritmusok által felismert minták és összefüggések közötti kapcsolatokat emberi elemzőknek kell értelmezniük. Azonban a Big Data elemzési technológiák használata során nem csak a hamis

²¹ <https://news.yahoo.com/saudi-national-questioned-boston-bombings-allegedly-flagged-terror-134039390.html>

²² YEE, CHU, 2015.

²³ <https://learn.microsoft.com/ru-ru/azure/machine-learning/concept-manage-ml-pitfalls?view=azureml-api-2>

pozitív eredmények jelenthetnek problémát a nemzetbiztonság területén. Az ellenkező eset, azaz a hamis negatív eredmények is előfordulhatnak, amikor az adatok elemzése és értelmezése során egyáltalán nem sikerül azonosítani egy valós fenyegetést vagy bűncselekményt. Ez szintén komoly veszélyt jelenthet a biztonságra, mivel a hatóságok képtelenek lehetnek a megfelelő lépéseket megtenni az időben történő beavatkozás érdekében. A hamis negatív eredmények születésének fő okai a hiányzó adatok és az érzékeny információk védelme miatti korlátozott hozzáférés. Az adatok védelme szempontjából szükséges korlátozások miatt az adatok meghatározott részei nem állnak rendelkezésre az elemzők számára, ami negatív hatással lehet az elemzés pontosságára és teljességére.

A hamis pozitív és a hamis negatív eredmények hatásai nemcsak a személyiségi jogok megsértésével és a bizalom elvesztésével járnak, hanem a nemzetbiztonság megtévesztésével is. A hamis eredmények miatt az erőforrások a tévesen azonosított személyekre irányulhatnak, miközben az igazi fenyegetések nem kerülnek azonosításra. A hamis eredmények hibásan aktiválhatják a rendszereket, ami idő és pénz veszteséggel járhat, nem beszélve a nemzetbiztonsági szervezetek társadalmi megítélésében keletkező károkról.

Ezért a Big Data elemzése során a hamis eredmények minimalizálása érdekében egyaránt biztosítaniuk kell a pontos és megbízható adatokat, az algoritmusok és a feldolgozási folyamatok alapos ellenőrzését, valamint az emberi értelmezés és értékelés lehetőségét.

3.2. Társadalmi és etikai kérdések

Hajlamosak vagyunk úgy gondolni a Big Data technológiára, hogy az mindig megbízható és pontos, mindent megold önállóan, pusztán kérdezni kell tőle.²⁴ Az igaz, hogy a nagy adathalmazok és azok adatelemzése, az így kinyert információ, valamint az azon keresztül keletkező többlet tudás kiterjedt felhasználása és alkalmazása rendkívül hasznos a nemzetbiztonsági célok elérésében, de az adatok átfogó elemzése önmagában nem jelent biztonságot. A helyes döntések meghozatalához az adatokat a megfelelő módon kell értelmezni és felhasználni.

A nagy adathalmazok elemzése és értelmezése bonyolult folyamat, amely magas szintű szakértelmet igényel. Megfelelő szakértői csoportokat kell

²⁴ NEEDLE, 2015.

létrehozni annak érdekében, hogy az adatokat hatékonyan lehessen felhasználni, és a szükséges személyi és pénzügyi erőforrásokat is biztosítaniuk kell a Big Data alkalmazások kezeléséhez. Az adatelemzéshez és az adatok kezeléséhez szükséges eszközök és szoftverek komplexek lehetnek, és a kezelésük is nagy kihívást jelenthet.

A Big Data nagyban hozzájárulhat ahhoz, hogy az állampolgárok biztonságban érezzék magukat, ugyanakkor a nemzetbiztonsági célzatú Big Data alkalmazások számos társadalmi és etikai kihívást hordoznak magukban. A rendszer működtetése a nagy mennyiségű adatok gyűjtése és elemzése során rengeteg technikai eszközt és szakértelmet igényel, valamint az adatok biztonságos tárolása és védelme is kiemelten fontos feladat. Egyetlen hibás beállítás vagy megfelelő védelem nélküli adatbázis könnyen célpontjává válhat a hackerek, vagy ellenérdekelt titkos szolgálatok támadásainak. Éppen ezért a Big Data rendszerek, mint minden más informatikai hálózat, állandóan fejlesztéseket, karbantartást és frissítéseket igényelnek, amik nagyon magas költségekkel járnak. Az említett költségek magukkal hozhatják a központosítás szükségességét, figyelemmel arra, hogy az adatokat gyűjtő és elemző rendszerek kialakítása, fenntartása és fejlesztése központi irányítást igényel.

Ezek a központosított rendszerek erősíthetik a nemzetbiztonsági szervek által folytatott adatgyűjtésekből származó adatok feldolgozásához, felhasználásához kapcsolódó társadalmi aggodalmakat és bizalmatlanságot.

Napjainkban kiemelt jelentőségű kérdés a személyiségi jogok, azon belül is a magánélet, a magán titok és a személyes adatok védelmének teljes körű biztosítása. Az emberek nem aggódhatnak az őket esetleg érintő adatgyűjtésekkel és az így megszerzett adatok felhasználásával kapcsolatos kérdések miatt, különösen a magánéletükre vonatkozó adatok esetében. Ha az emberek úgy érzik, hogy az adatokat visszaélésre használják fel, vagy nem megfelelően biztonságosak, akkor elveszítik a bizalmukat a rendszer iránt. 2013-ban Edward Snowden²⁵ nyilvánosságra hozta, hogy az Amerikai Egyesült Államok Nemzetbiztonsági Ügynöksége /NSA/²⁶ a Prism projekt keretein belül 2007 óta, a globális internetes, informatikai cégóriások úm. Apple, Google, Microsoft, Facebook, Youtube stb. közreműködésével, szervereikhez biztosított hozzáféréseken keresztül, valós időben tudta nyomon követni a felhasználók e-mailjeit, chat és hívás kommunikációját, videóit,

²⁵ GREEWALD 2014.

²⁶ NATIONAL SECURITY AGENCY.

fényképeit. A tömeges adatszerzések, megfigyelések nagyban befolyásolják az állampolgárok és a kormányzat közötti bizalmi viszonyt. Az emberek ráébredtek arra, hogy a terrorizmus elleni küzdelem égisze alatt a mindennapi életük került indokolatlanul nagytű alá. A bizalomvesztést jól érzékelteti, hogy 2013. nyarától Obama elnök népszerűsége folyamatosan esett vissza és az amerikai társadalom többsége úgy gondolta, hogy rosszul végzi a munkáját.

27

A döntéshozóknak egyensúlyt kell találniuk a nemzetbiztonság garantálása, az állampolgárok védelme és az egyéni jogok tiszteletben tartása között. A biztonságról való állami gondoskodáson túl, a nem megkerülhető kérdések közé tartozik az adatok védelme, az emberi jogok és méltóság, valamint a magánélet tiszteletben tartása.

Az adatelemzés és az információtechnológiai fejlesztések robbanásszerű növekedése lehetővé teszi, hogy az államok nagy mennyiségű adatot gyűjtsenek az állampolgáraikról. Az adatokat csak a megfelelő intézkedésekkel lehet hatékonyan felhasználni, és a végső döntésnek mindig figyelembe kell vennie az állampolgárok érdekeit és jogait is, és felvetnek néhány fontos kérdést a nemzetbiztonság és az állampolgári szabadságjogok közötti egyensúly kapcsán.

KONKLÚZIÓ

A nagy adathalmazok nemzetbiztonsági felhasználása az elmúlt években jelentős figyelmet kapott. A nemzetbiztonságáért felelős szervezetek analitikai megközelítéssel képesek lehetnek felismerni és megérteni azokat a mintákat és jellemzőket, amelyek segíthetnek a biztonságra veszélyt jelentő események, személyek és tevékenységek korai azonosításában.

A Big Data nemzetbiztonsági területen történő alkalmazása értékes és teljesen újszerű rálátást biztosít a potenciális biztonsági fenyegetésekre, fokozhatja a nemzetbiztonsági erőfeszítéseket. A nagy mennyiségű adatok kiemelt fontossággal bírnak a biztonságos környezet garantálásában, a nemzetbiztonsági aspektusokkal rendelkező eljárások menetében, átalakítják a hírszerzési ismereteket, tevékenységeket és szervezeteket, valamint hatással vannak a nemzetbiztonsági fenyegetések kezelésére, illetőleg formálják a jelenlegi és jövőbeli hírszerzési tevékenységet és döntéshozatalt. A nemzetbiztonságára leselkedő veszélyek elhárítása érdekében a hagyományos

²⁷ <https://news.gallup.com/poll/116479/barack-obama-presidential-job-approval.aspx>

hírszerzési területeket ki kell terjeszteni egy teljes körű, minden forrást figyelembe vevő, az új típusú adatokat is felhasználó adatfeldolgozó rendszerek kiépítésére, ezzel jelentős előrelépési lehetőséget biztosítva a korai előre jelző rendszerek működtetésének.

A Big Data lehetőséget ad a nemzetbiztonsági szervezetek számára, hogy hatékonyabban dolgozzanak a fenyegetések kezelésén és a nemzetbiztonság megőrzésén. Azonban, az adatelemzés kihívásai és a felmerülő aggályok azt mutatják, hogy az adatgyűjtés és elemzés megfelelő kezelése létfontosságú. Kiemelt figyelmet kell fordítani az adatvédelem és biztonság kihívásaira, a hamis pozitív és hamis negatív eredményekből fakadó pontatlanságok kiküszöbölésére. A nemzet biztonságáért felelős szervezeteknek az adatbiztonsági és adatvédelmi intézkedések mellett biztosítaniuk kell az elemzéshez használt algoritmusok és gépi tanulási modellek pontosságát. Csak ezekkel az óvintézkedésekkel lehet kiaknázni a Big Data előnyeit a kockázatok minimalizálása mellett.

A szervezeteknek gondoskodniuk kell a megfelelő biztonsági intézkedésekről, a társadalmi és etikai kérdések kezeléséről, valamint a technológiai függőségből fakadó kockázatok minimalizálásáról annak érdekében, hogy a Big Data elemzés az elvárt eredményeket hozza a nemzetbiztonság terén.

A nemzetbiztonsági szervezetek elemzési tevékenységén a Big Data alkalmazások változtattak, változtatni fognak, de azt megjósolni, hogy a hagyományos adatfeldolgozói, elemző rendszereket mikor és milyen határfokon fogja teljesen kiváltani nem lehet.

FELHASZNÁLT IRODALOM

- ALSAYAT, AHMED – EL-SAYED, HODA (2016). Social media analysis using optimized K-Means clustering. 2016 IEEE 14th International Conference on Software Engineering Research, Management and Applications (SERA), Towson, MD, USA, 61-66. <https://doi.org/10.1109/SERA.2016.7516129>
- DE MAURO, ANDREA – GRECO, MARCO – GRIMALDI, MICHELE (2016). A formal definition of Big Data based on its essential features. Library review, Vol. 65, Issue 3, 122-135. <https://doi.org/10.1108/LR-06-2015-0061> Letöltés helye: <https://www.emerald.com/insight/content/doi/10.1108/LR-06-2015-0061/full/html> (Letöltve: 2023.05.25.)

- FAVARETTO, MADDALENA –DE CLERCQ, EVA – SCHNEBLE, CHRISTOPHE OLIVIER – ELGER, BERNICE SIMONE (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. <https://doi.org/10.1371/journal.pone.0228987> (Letöltve: 2023.03.16.)
- GREENWALD, GLEEN (2014). A Snowden-ügy. HVG Kiadó. ISBN 9789633041833
- HOFACKER, CHARLES F. – MALTHOUSE, EDWARD CARL – SULTAN, FAREENA (2016). Big Data and consumer behavior: imminent opportunities. *Journal of Consumer Marketing*, Vol. 33 No. 2, 89-97. <https://doi.org/10.1108/JCM-04-2015-1399> (Letöltve: 2023.05.26.)
- HOLLIN, RUPERT (2015). [Drilling into the Big Data Gold Mine: Data Fusion and High-Performance Analytics for Intelligence Professionals.](https://doi.org/10.1016/B978-0-12-801967-2.00002-1) <https://doi.org/10.1016/B978-0-12-801967-2.00002-1> (Letöltve: 2023.05.25.)
- HOSSAIN, M. SHAHRIAR – BUTLER, PATRICK –BOEDIHARDJO, ARNOLD P. – RAMAKRISHNAN, NAREN (2012). Storytelling in entity networks to support intelligence analysts. *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1375–1383. <https://doi.org/10.1145/2339530.2339742> (Letöltve: 2023.03.16.)
- <https://datareportal.com/global-digital-overview#:~:text=Roughly%204.66%20billion%20people%20around,over%20the%20past%20twelve%20months%20> (Letöltve: 2023.05.26.)
- <https://learn.microsoft.com/hu-hu/azure/machine-learning/concept-manage-ml-pitfalls?view=azureml-api-2> (Letöltve: 2023.05.25.)
- <https://news.gallup.com/poll/116479/barack-obama-presidential-job-approval.aspx> (Letöltve: 2023.05.25.)
- <https://news.yahoo.com/saudi-national-questioned-boston-bombings-allegedly-flagged-terror-134039390.html> (Letöltve: 2023.05.25.)
- <https://publicintelligence.net/new-york-state-law-enforcement-terrorism-indicators-reference-card/> (Letöltve: 2023.05.25.)
- <https://www.sciencefocus.com/future-technology/how-much-data-is-on-the-internet/> (Letöltve: 2023.03.10.)
- <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/> (Letöltve: 2023.03.10.)
- <https://www.statista.com/statistics/871513/worldwide-data-created/> (Letöltve: 2023.03.10.)
- LANEY, DOUGLAS (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety. Gartner. Letöltés helye:

<https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--velocity--an> (Letöltve: 2023.03.16.)

<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-ControllingData-Volume-Velocity-and-Variety.pdf>
(Letöltve: 2023.05.24.)

LEONELLI, SABINA (2020). Scientific Research and Big Data. The Stanford Encyclopedia of Philosophy (Summer 2020 Edition). Letöltés helye: <https://plato.stanford.edu/archives/sum2020/entries/science-big-data/>
(Letöltve: 2023.05.26.)

LIMAJ, EVERIST – BILALI, EDONA (2018). Big Data Systems: A renewed definition of the concept. 11th IADIS International Conference Information Systems Letöltés helye: https://www.esearchgate.net/profile/Everist-Limaj/publication/326081575_Big_Data_Systems_A_Renewed_Definition_Of_The_Concept/links/5b372a770f7e9b0df5dc7f75/Big-Data-Systems-A-Renewed-Definition-Of-The-Concept.pdf (Letöltve: 2023.05.25.)

LYON, DAVID – WOOD, DAVID MURAKAMI (2020). Big Data Surveillance és Security Intelligence: The Canadian Case. University of British Columbia Press ISBN 9780774864190

MAGYARORSZÁG MESTERSÉGES INTELLIGENCIA STRATÉGIÁJA 2020-2030, Letöltés helye: <https://digitalisjoletprogram.hu/files/6f/3b/6f3b96c7604fd36e436a96a3a01e0b05.pdf> (Letöltve: 2023.05.25.)

NEEDLE, DAVID (2015). Big data has peaked, and that's a good thing. Datainformed. <http://datainformed.com/nate-silver-big-data-has-peaked-and-thats-a-good-thing/>
<https://www.dataversity.net/why-nate-silver-thinks-big-data-has-peaked/> (Letöltve: 2023.05.24.)

NUCCIO, MASSIMILIANO – GUERZONI, MARCO (2019). Big data: Hell or heaven? Digital platforms and market power in the data-driven economy. Competition & Change, 23(3), 312–328. <https://doi.org/10.1177/1024529418816525> (Letöltve: 2023.05.26.)

YEE, STEPHANIE – CHU, TONY (2015). A visual introduction to machine learning. R2D3. Letöltés helye: <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/> (Letöltve: 2023.03.16.)