



Ügyszám: NAIH-1938-2/2013/T
Készítette: dr. Bíró János ov.

Gyakorlati útmutató védett adatot nem tartalmazó kivonat készítéséhez

Számos jogszabály ír elő olyan adattovábbítási vagy nyilvánosságra hozatali kötelezettséget, amelynek teljesítéséhez védett adatot nem tartalmazó kivonat készítése szükséges. A Nemzeti Adatvédelmi és Információszabadság Hatóság (NAIH) által végzett vizsgálatok tapasztalatai szerint a kivonatkészítésnek olyan buktatói vannak, amellyel az adatkezelők esetenként nincsenek tisztában, ezért nem megfelelő adattartalmú kivonatot készítenek. A személyes adatok védelme és a közérdekű adatok nyilvánossága szempontjából egyaránt fontos, hogy a kivonat tartalma megfeleljen a jogszabályi előírásoknak. A NAIH e célt hét jótanáccsal kívánja elősegíteni.

A vizsgálatok során megismert, helytelenül anonimizált dokumentumokra adat- és titokvédelmi okból nem utalhatunk, ezért a kivonat elkészítésének buktatóit egy példamese keretei között mutatjuk be, amely arról szól, hogy Dzsafár vezír ellen nyomozás indult, mert segítséget nyújtott a perzsa Alinak, hogy Harún ar-Rasid kalifa közelébe jusson. A perzsa Ali a gyanú szerint azonos azzal az Ali Babával, aki a negyven rabló révén valamilyen kapcsolatban áll a szervezett bűnözéssel, és bizonyos lámpaolaj ügyekben is felmerült a neve. A politika és a szervezett bűnözés kapcsolatai feltárásának jegyében elrendelték a vonatkozó iratok nyilvánosságra hozatalát, azonban a folyamatban lévő nyomozásra tekintettel ez csak anonimizált formában történhet. Nemzetbiztonsági érdekből védendő továbbá az az információ, hogy a kalifa álmatlanságban szenved. A védett adatokat tartalmazó szövegrészlet¹ a következő:

Harún ar-Rasíd kalifa egy éjjel álmatlan volt, és elhívatta vezírjét, Dzsafárt. Mikor ez megjelent előtte, így szólt hozzá: "Dzsafár, ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet." Dzsafár így felelt: "Igazhívők fejedelme, van egy barátom, a perzsa Ali, az tud meséket, históriákat; felüdíti a lelket, minden bánatot elkerget." - "Hívd

¹ A részlet az Ezeregyéjszaka meséiből, a Harún ar-Rasid és a perzsa Ali, meg a tarisznýája című meséből származik. Fordította Honti Rezső. A Magyar Elektronikus Könyvtárból a <http://mek.oszk.hu/03200/03210/index.phtml> URL-ről tölthető le.

ide” - szólt a kalifa. Dzsafár azt felelte: „Hallom és engedelmeskedem” - és eltávozott megkérteni a perzsát;

A kivonat készítését az írnokra bízta, aki a következőképp alakította át a szöveget:

██████████ kalifa egy éjjel ██████████, és elhívatta vezírjét, ██████████t. Mikor ez megjelent előtte, így szólt hozzá: „██████████; ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet.” ██████████ így felelt: „Igazhívők fejedelme, van egy barátom, a perzsa ██████████, az tud meséket, történeteket; felvidíti a lelkét, minden bánatot elkerget.” - „Hívd ide” - szólt a kalifa. ██████████ azt felelte: „Hallom és engedelmeskedem” - és eltávozott megkérteni a perzsát;

Az írnok szó szerint vette a feladatot (az „anonimizálás” latin szó névtelenné tételt jelent), és csak a neveket, valamint az álmatlanságra való utalást törölte. Ezzel hibát követett el, mert a szöveg összefüggései, és közismert információk alapján – mint például az, hogy a kalifát igazhívők fejedelmének nevezik, vagy a vezírt Dzsafárnak hívják – , rekonstruálható, hogy kiktől van szó a szövegben.

1. tanács: az anonimizált kivonat elkészítése nem pusztán ügyviteli feladat. Az adatkezelőnek felelősségteljes döntést kell hoznia arról, hogy melyek a védendő adatok, amelyeket a kivonat nem tartalmazhat. A védendő adatok meghatározására csak olyasvalaki alkalmas, aki jól ismeri a kivonat alapjául szolgáló irat tartalmát, valamint az iratban lévő adatok védelmére, illetve nyilvánosságára vonatkozó jogszabályi előírásokat.

Ezek után a bölcs és óvatos Hasszán al-Fotira bízta az irat anonimizálást, aki a következő szövegváltozatot készítette:

[...] egy éjjel elhívatta [...]. Mikor ez megjelent előtte, így szólt hozzá: “[...]” [...] így felelt: “[...]” - “[...]” - szólt [...]. [...] azt felelte: “[...]” - és eltávozott megkérteni a [...];

Első pillantásra látható, hogy Hasszán túlbuzgó volt, mert a védett adatokon kívül nyilvános közérdekű adatokat is törölt a szövegből, továbbá az anonimizálás olyan módját választotta, amely megsemmisítette a szöveg belső összefüggéseit és eltorzította a

szövegrészletek arányait. A kalifa álmatlanságára vonatkozó részt pedig nyomtalanul eltüntette a szövegből.

2. tanács: a kivonat készítése során csak a törvény alapján védendő információkat szabad a szövegből törölni. Az adatkezelő nem zárhat el önkényesen olyan információkat a nyilvánosság elől, amelyek bizalmasságát, védelmét jogszabály nem írja elő.

3. tanács: a kivonatot úgy kell elkészíteni, hogy a szöveg belső összefüggései lehetőleg fennmaradjanak, mert egyébként az eredmény használhatatlan szóhalmaz lesz. A bemutatott szöveg elektronikus formában rendelkezésre áll, anonimizálása helyesen úgy történhet, hogy a szereplők azonosító adatait pszeudonim kódokra cserélik².

4. tanács: Egyértelműen jelezni kell, hogy honnan töröltek adatot, mert a védett adat észrevétlen, nyomtalan törlése éppúgy meghamisíthatja a szöveg értelmét, mint egy adat meghamisítása, vagy valótlan adat betoldása.

5. tanács: a szöveg arányai is információt hordoznak. A kivonat felhasználói nem jogosultak a védett adatokat megismerni, azonban az nem titkolható el előlük, hogy valahonnan egy szót, egy sort, egy bekezdést, vagy egy oldalt töröltek. Ezért törekedni kell arra, hogy a kivonat megőrizze az eredeti szöveg arányait.

Hasszán a tanácsokat megfogadva gyorsan készített egy olyan szövegváltozatot, amelyben a szereplők neveit pszeudonim kódokra cserélte.

Harún ar-Rasid kalifa pszeudonim kódja (X) lett, Dzsfár vezír (Y), a perzsa Ali pedig (Z). Ügyelt arra is, hogy a pszeudonim kódokat ne csak a nevek helyére illessze be, de oda is, ahol név nélkül történik hivatkozás valamely szereplőre. A kalifa álmatlanságára

² Előfordul, hogy a közzétételt elrendelő jogszabály meghatározza az anonimizálás módját. Például a bíróságok szervezetéről és igazgatásáról szóló 2011. évi CLXI. törvény 166. § (1) bekezdése szerint a Bírósági Határozatok Gyűjteményében közzétett határozatban szereplő személyek azonosítását lehetővé tevő adatokat olyan módon kell törölni, hogy az ne járjon a megállapított tényállás sérelmével. Egyebekben a határozatban szereplő egyes személyeket az eljárásban betöltött szerepüknek megfelelően kell megjelölni.

vonatkozó adatot pedig a törlés tényét jelző, és a szöveg arányait nem módosító módon törölte a szövegből. Íme:

(X) egy éjjel [REDACTED], és elhívatta (Y)-t. Mikor ez megjelent előtte, így szólt hozzá: "(Y), ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet." (Y) így felelt: "(X), van egy barátom, (Z), az tud meséket, históriákat; felüdíti a lelkét, minden bánatot elkerget." - "Hívd ide" - szólt (X). (Y) azt felelte: "Hallom és engedelmesködöm" - és eltávozott megkérteni (Z)-t;

Ezek után már készülődtek a kivonatot tartalmazó adatállomány elektronikus közzétételére, amikor a kalifa udvari informatikusa észrevette, hogy a közzétételre szánt szövegállományban nem kívánatos információk maradtak. Ugyanis a megjegyzések megjelenítését bekapcsolva így festett a szöveg:

(X) egy éjjel [REDACTED], és elhívatta (Y)-t. Mikor ez megjelent előtte, így szólt hozzá: "(Y), ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet." (Y) így felelt: "(X), van egy barátom, (Z), az tud meséket, históriákat; felüdíti a lelkét, minden bánatot elkerget." - "Hívd ide" - szólt (X). (Y) azt felelte: "Hallom és engedelmesködöm" - és eltávozott megkérteni (Z)-t;

Megjegyzés [U4]: Harún ar-Rasíd kalifa hívatta

Sőt, az elektronikus dokumentumhoz tartozó adatlapban is maradt védendő információ, amelynek nem lett volna szabad a kivonatba bekerülnie:

Harún ar-Rasíd és a perzsa Ali meg a tarisznnyája.doc tulajdonságok

Általános	Adatlap	Statistika	Tartalom	Egyéni
Cím:	Harún ar-Rasíd és a perzsa Ali meg a tarisznnyája			
Tárgy:	Az Ezeregyéjszaka meséi			
Szerző:	Seherezáde			
Felelős:				
Cégnév:	-			
Kategória:				
Keresőszó:				
Megjegyzések:				
Hivatkozás alapja:				
Sablon:	Normal.dot			
☐ Mintakép mentése				
OK Mégse				

Hasszán nem lévén informatikus, nem tudta, hogy az irodai szövegszerkesztővel használható elektronikus dokumentumok nem szembetűnő módon különféle kísérőinformációkat tartalmaznak. Ilyen lehet például a szöveghez tartozó megjegyzés, fájlnev, statisztikai adatlap, normál nézetben nem látható fejléc, lábléc, lábjegyzet. Az elektronikus dokumentumba beágyazott objektumok (képek, táblázatok stb.) szintén tartalmazhatnak rejtett kísérőinformációkat, amiként a hipertext leíró nyelven készített dokumentumok nem megjelenített kódrészletei³.

6. tanács: Elektronikus közzététel esetén célszerű informatikus szakértővel előzetesen megvizsgáltatni, hogy tartalmaz-e az elektronikus dokumentum védett adatot nem megjelenített kísérőinformációként. Kétség esetén az is megoldás lehet, ha a kivonatot tartalmazó elektronikus dokumentumot kinyomtatják. Ha a kivonatra elektronikus formában van szükség, akkor a nemkívánatos kísérőinformációktól megtisztított, kinyomtatott lapok utólag beszkennelehetők.

Hasszán annyira komolyan vette a kísérőinformációkkal kapcsolatos intelmet, hogy „biztos ami biztos”, nem is a kivonatot nyomtatta ki, hanem az eredeti dokumentum másolatát, majd filctollal kisatírozta rajta a védendő információkat⁴.

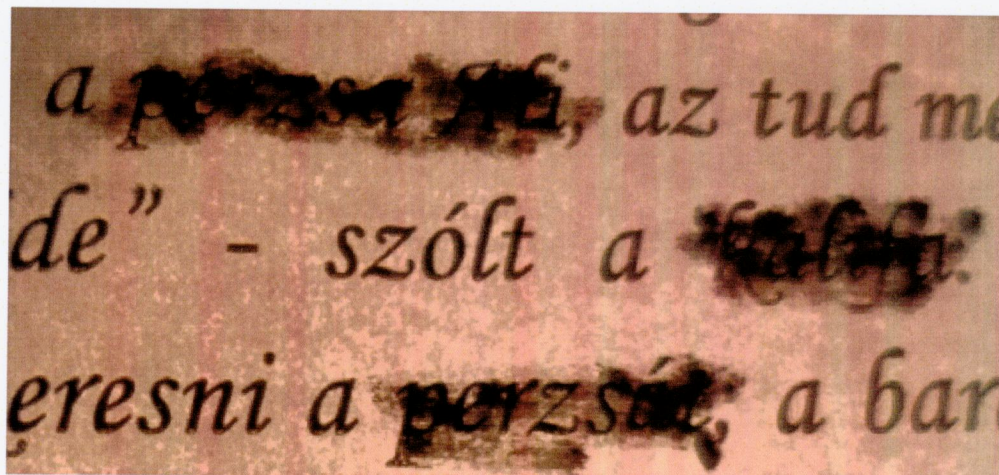
³ Hangsúlyozzuk, ez az útmutató nem tárgyalja, hogy miként lehet szándékosan kódolt, illetve rejtett információkat elhelyezni egy dokumentumban, és miképpen fedhető fel ez. E tárgykörökkel a rejtjelezés, a szteganográfia és a rejtjelfejtés foglalkozik. Kísérőinformáció alatt itt csak olyan információkat értünk, amelyek az elektronikus dokumentumokban rendszerint előfordulhatnak.

⁴ A bölcs Hasszán természetesen tudta, hogy a köziratokról, a közlevéltárakról és a magánlevéltári anyag védelméről szóló 1995. évi LXVI. törvény szerint az irattári anyaggal rendelkező szervek kötelesek a szervesen összetartozó irataik egységének, illetve eredeti rendjének megőrzéséről, valamint a tulajdonukban vagy birtokukban lévő maradandó értékű iratok megóvásáról gondoskodni, ezért a kivonat készítése során az eredeti irat épségét meg kell őrizni. Ha kivonatkészítés során az irat épségét sértő változtatás – például kivágás, befestés, satírozás stb. – szükséges, akkor ezt nem az eredeti iraton, hanem annak másolatán kell elvégezni.

Az eredmény nem lett valami szép, de biztos volt abban, hogy az irat immár nem tartalmaz rejtett kísérőinformációkat:

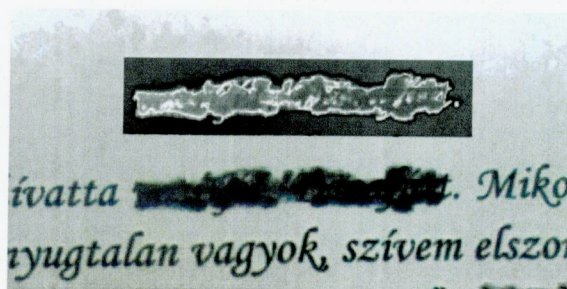
„...egy éjjel ... és elhívatta Mikor ez megjelent előtte, így szólt hozzá: „... ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet.” ... így felelt: „... van egy barátom, a ... az tud meséket, históriákat; felüdíti a lelket, minden bánatot elkerget.” - „Hívá ide” - szólt a ... azt felelte: „Hallom és engedelmeskedem” - és eltávozott megkeresni a ... a barátjáért¹ küldött, és mikor ... eljött, azt mondta neki: „Engedelmeskedj ... hívásának.” - „Hallom és engedelmeskedem” - felelte az, és elment

Azonban Hasszán ezúttal a 3. tanács megszegésén túl elkövetett egy másik hibát is: olyan módszert választott, amely nem garantálja a védett adatok teljes és visszafordíthatatlan törlését. Ugyanis a kisatírozással, lefestéssel lefedett adatok változatlanul megmaradnak a papíron, és később fizikai, vegyi, illetve informatikai eljárásokkal általában visszanyerhetők. Esetünkben az egyik kisatírozott szövegrészlet így fest, ha a papírt erős fénnel átvilágítják, és a szövegről nagy felbontású digitális fényképet készítenek:



A védett adatok („perzsa Ali”, „kalifa”, „perzsát”) jól olvashatók. A filctollal történő satírozás információvédelmi szempontból azért sem megfelelő, mert az idő múlásával fakul, és az alatta lévő szöveg magától is előtűnik.

Az adatok kisatírozása olyankor sem teljesen biztonságos, ha nem a satírozott szöveget, hanem annak szkennelt vagy fényképezett másolatát kívánják kivonatként felhasználni. Ilyen esetben is előfordulhat, hogy a látszatra törölt adatok képszerkesztő programmal, az iratról készült kép szabad szemmel nem látható fényesség- és színkülönbségeit felerősítve, kontrasztosabbá téve mégis visszanyerhetők, mint az alábbi példánkban a „Dzsafárt” szó:



Akkor is érheti kellemetlen meglepetés az adatkezelőt, ha az elektronikus dokumentumban elektronikus úton fedi el a védendő információt. Például a már bemutatott, pszeudonimizált szövegben Hasszán a kalifa álmatlanságára vonatkozó adatot a háttérszín feketére állításával „tüntette el”. Ám a dokumentumot szövegszerkesztővel megnyitva a háttérszín átállításával az információ visszanyerhető:

(X) egy éjjel álmatlan volt, és elhívatta (U) (Y)-t. Mikor ez megjelent előtte, így szólt hozzá: "(Y) ma éjjel szörnyen nyugtalan vagyok, szívem elszorul, szeretném, ha valamivel felvidítanád a lelkemet, jókedvre hangolnád a szívemet." (Y) így felelt: "(X), van egy barátom, (Z), az tud meséket, történeteket, felvidíti a lelkét, minden bánatot elkerget." - "Hívá ide" - szólt (X). (Y) azt felelte: "Hallom és engedelmeskedem" - és eltávozott megkeresni (Z)-t.

Az elektronikus dokumentum utólagos módosítása és az elfedett adatokhoz történő hozzáférés egyes dokumentumformátumok esetén DRM (Digital Right Management – digitális jogkezelés) alkalmazásával korlátozható, azonban célszerű abból kiindulni, hogy a legtöbb DRM feltörhető vagy megkerülhető⁵.

⁵ Közérdekű adatokat tartalmazó kivonat esetében a DRM alkalmazása azért sem ajánlott, mert a digitális korlátozások – például a nem védendő szövegrészlet kimásolásának, vagy a szöveg kinyomtatásának lehetetlenné tétele – akadályozhatja a közérdekű adat felhasználását és terjesztését.

7. tanács: a teljes és végleges adattörlést akár papír alapú iraton, akár elektronikus dokumentumban minden esetben úgy kell végrehajtani, hogy a kivonat még elfedve, vagy DRM segítségével hozzáférhetetlenné téve se tartalmazza a védett adatokat.

Végül a kalifa úgy rendelkezett, hogy a pszeudonimizált szövegű elektronikus dokumentum oldalainak képét pixelgrafikus formátumban tegyék közzé a palotája honlapján.

Idáig tartott a tanmese.

Bízom abban, hogy a mesés kerettörténetben elhelyezett komoly tanácsok segítenek az adatkezelőknek az adat- és információvédelemre, valamint a közérdekű adatok nyilvánosságára vonatkozó jogszabályi előírásoknak megfelelő kivonatok készítésében.

Budapest, 2014. január „ 31 ”.

Dr. Péterfalvi Attila
elnök
c. egyetemi tanár

