

Munkadokumentum és ajánlások a személyes adatok Interneten való közzétételére, a honlap-tartalmak indexelésére és az adatvédelemre vonatkozóan*

53. Ülés – Prága, 2013. április 15-16.

1. Háttér

Az adatvédelem fő pillérjeinek egyike mindig is az érintettnek az a joga volt, hogy ellenőrzést gyakorolhat saját adatai fölött. Ennek az ellenőrzésnek egy lényeges eleme az érintettnek az a joga, hogy töröltheti adatait, ha azokat jogszabályi rendelkezésekkel ellentétesen dolgozzák fel, vagy az adatalany már nem járul hozzá feldolgozásukhoz. Az Európai Bizottság egy új szabályozási keretre tett újkeletű javaslata azáltal szigorítja ezt a jogot, hogy rendelkezik a mások által vagy az Interneten való „felejtés jogáról”. Ez nem vonatkozik azokra az esetekre, amikor törvényes és jogilag indokolt érdek fűződik az adatok nyilvánosságához és láthatóságához, például médiaarchívumok esetében, vagy történelmi feljegyzések céljára, és egyértelmű, hogy a felejtés joga *a priori* a véleménynyilvánítás szabadságát vagy a média szabadságát nem előzheti meg.¹

Az Internet jelenlegi struktúráját illetően a „felejtés joga” ugyan többféle módon megvalósítható, ám ezek többsége mind technikai, mind jogi szempontból még megoldásra vár. Nagyon valószínű, hogy személyes adatok (és bármely más információ), ha már online nyilvánosságra hozták őket, a nagyközönség számára hozzáférhetőek maradnak. Még ha törölték is azokat az eredeti honlapról, az a kapcsolat vagy tükrözött tartalom, melyet a törlést megelőzően más honlapokkal létesítettek, megmarad. Az Internet nem tudja, „hogyan kell felejteni”, és jelenleg nem áll rendelkezésre olyan technikai eszköz, amely biztosítaná az Interneten közzétett adatok szisztematikus törlését (vagyis megtanítaná az Internetet arra, hogyan kell felejteni). Röviden: nincs „törlés gomb”, és kétséges, lesz-e valaha.

Mindazonáltal már jelenleg is léteznek olyan módszerek, amelyek bizonyos mértékig védik az egyén felejtéshez fűződő jogát, melyek a honlap adminisztrátorai számára - mint az Internetre feltett személyes információk felfedésének korlátozására alkalmas eszközök² –

* A Munkacsoport „Working Paper and Recommendations on the Publication of Personal Data on the Web, Website Contents Indexing and the Protection of Privacy” című dokumentumának fordítása (Dr. Könyves-Tóth Pál munkája) figyelemmel német nyelvű változatára is. Letölthető: <http://www.datenschutz-berlin.de/content/europa-international/international-working-group-on-data-protection-in-telecommunications-iwgdp/working-papers-and-common-positions-adopted-by-the-working-group>

¹ EU adatvédelmi reform, 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age (*Európa mint a korszerű adatvédelem alapvető szabályainak megállapítója*), Viviane Reding előadása Münchenben, 2012 január 22-én, Innovation Conference Digital, Life, Design (*innováció, digitális, élet, tervezés*) konferencián: <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/12/26>; e megközelítés kritikáját lásd: Rosen, The Right to Be Forgotten, 64 Stan. L. Rev. Online 88.

² Ezeknek az eszközöknek az egyike a Google Webmaster Tools, amelyek lehetővé teszik, hogy a webmester-ek lássák, a Google hogyan kutatja és indexeli honlapjukat, és hogy befolyásolja azt a módot, ahogyan az URL-

rendelkezésre állnak, ugyanakkor az adminisztrátorok a keresőmotorok lehetőségeit is kihasználhatják. Az Internet jelenleg a felejtéshez fűződő jogot³ inkább úgy érthetjük vagy fordíthatjuk, mint a „meg nem találáshoz való jogot”.

2. Az Interneten megjelenő személyes adatok feletti ellenőrzés visszanyerésének kilátásai

A személyes adatok szaporodó közzététele az Interneten az utóbbi években új kihívásokkal és kockázatokkal jár a polgárok magánszférájára nézve, és egyúttal súlyosbítja a már létező kockázatokat is. A közösségi hálózatok térhódítása különösen döntő szerepet játszott ebben az összefüggésben⁴.

Míg az adatok – többek között személyes adatok – az Interneten való közzétételét támogató és hozzáférhetővé tevő technológiák ebben az összefüggésben drámaian fejlődtek, ezeknek az adatoknak a Interneten való hozzáférését ellenőrző technológiák fejlődése még gyerekcipőben jár. Miközben az utóbbi évtizedben egy “policy-aware Internet” létrehozásán fáradoztunk, még mindig várat magára egy hatékony, könnyen használható és széles körben hozzáférhető eszköz létrehozása, amely képessé teszi a polgárokat arra, hogy adataik felett – ha már egyszer közzétették őket az Interneten – ellenőrzésüket, még ha korlátozott mértékben is, (vissza-) nyerjék.⁵

Ilyen technológiák egy lehetséges tervezési célja lehet az adatok valamennyi másolata törlésének elősegítése minden olyan eszközről vagy tároló területről, ahol még megmaradtak. Ez jelenleg skálázhatósági problémákkal járhat (még automatizált megközelítés esetén is), különösen, ha specifikus adatokat terjesztettek az Interneten, vagy kiváltképp akkor, ha az internet használóinak közössége az idők folyamán az adatokat tovább pontosítják vagy szövegvagy környezetét átalakítják. Mostanában nincs olyan technikai módszer, amellyel egy adat minden másolatát vagy azzal összefüggő információ másolatát azonosítani vagy helyüket megállapítani lehetne. Mindez azonban lehetővé válhat egy jövőbeli “policy-aware Internet”-en.

Újonnan generált adatok esetén megjelenésük az Interneten határidejük (lejárati dátumuk) megadásával korlátozható. Ezt többféle módon is megtehetjük. Az adatokat elláthatjuk például egy olyan “aktív” (végrehajtható) szoftverrel, amely a lejárat határidő bekövetkeztekor közbelép, és megakadályozza az adatok megjelenítését a képernyőn, vagy lehetetlenné teszi, hogy a képről felvétel készüljön, vagy végérvényesen törli vagy titkosítja az eredeti tartalmat. Vagylagos megoldásként az adatok lejárat idejükkal „megjelölhetők”, s így minden olyan szerver, amely ezeket kezeli, e dátumra való tekintettel a lejárat dátumának beálltával eltávolíthatja az adatokat.

eket indexelik és megtalálják. Ezek a rendelkezésre álló eszközök hozzáférhetők:

<http://www.google.ca/webmasters/>.

³ Megjegyezzük, hogy a „felejtés jogát” ebben a dokumentumban szélesebben értelmezzük, mint az EU Adatvédelmi rendelete, és ez a dokumentum semmiképpen nem állítja, hogy a „felejtés jogát” ez a rendelet definiálja-e vagy sem.

⁴ Vö.: E Csoport jelentése és ajánlása a közösségi hálózati szolgáltatásokban megjelenő adatok védelmére: „Rom Memorandum (Róma, Olaszország), 2008. március 3-4.

<http://www.datenschutzberlin.de/attachments/897/675.36.5.pdf>

⁵ Néhány javaslatot egy “policy-aware Web” létrehozására a „Rome Memorandum” (lásd fenti, 4. lábjegyzet) tartalmaz a 10. oldalon a 27. lábjegyzetben. A policy-aware Web koncepciója több létező technológiát ölel fel, például a strukturált adatokat, az azonosság-kezelést, a hozzáférés ellenőrzését és a sticky policies-t (vagyis azokat a használati szabályokat, amelyek az adatokkal együtt utaznak).

Az Interneten újonnan generált adatok élettartamának az egyedi igényekhez való igazítására további érdekes példákkal szolgál néhány más, újabb alkalmazás. A felhasználó például egy biztonságos overlay hálózatot használhat, amely egy, ugyanahhoz az overlay hálózathoz tartozó közösség számára korlátozza a tartalom – például egy hír vagy egy kép – láthatóságát azáltal, hogy végponttól végpontig terjedő biztonságot vagy hozzáférhetőséget ellenőrző szabályokat alkalmaz. További egyéb alkalmazások esetén a mobil szöveges üzenet egy konfigurálható lejáratú dátumig a felhasználó számára hozzáférhető marad. Végül megemlíthetjük azokat a „felhasználó-centrikus” megoldásokat, melyek esetében az adatok legitim tulajdonosa azokhoz szelektív hozzáférést biztosít, megadva a kapcsolatot ahhoz a helyhez, ahol az adatokat ténylegesen egy meghatározott időtartamig tárolják.

Ezek a példák egy jövőbeli „policy-aware Internet” építőkövei lehetnek. Mindazonáltal alapos kutatás és fejlesztés szükséges ahhoz, hogy ezeknek az elemeknek a további fejlesztése a polgárok magánszférájának jobb védelmére szolgáljon. A Munkacsoport felhívja e terület illetékes szereplőit (ipar, tudomány, kormányzat), hogy e terület fejlesztése érdekében tovább fokozzák erőfeszítéseiket.

3. Személyes adatok hozzáférhetőségének korlátozása az Interneten keresőmotorok által való indexelhetősége ellenőrzésével

További építőelem a jelenlegi Interneten megjelenő adatok elérhetősége korlátozásának és törölhetőségének elősegítésére az adatok elérhetőségének korlátozása a keresőmotorok kérdéseire adott eredmények rendelkezésre állásának behatárolásával⁶. Technikailag ez már lehetséges, és a honlap-adminisztrátorok ezt a lehetőséget kihasználhatják. Ez lényegében két alternatívára támaszkodik: az egyik a „robots.txt-Protokoll”⁷, a másik, amikor egy objektumot jelzésekkel látnak el, melyek arra utalnak, hogy egy meghatározott tartalmat vagy egy meghatározott honlapot egy keresőmotor nem indexelhet.

A robots.txt protokoll egy olyan egyszerű utasításkészlettel működik, amely egy szöveges fájlba (a robots.txt fájlba) kódolnak, amelyet egy domén gyökérkönyvtárban helyeznek el (pl.: <http://example.com/robots.txt>). Ezt a fájlt, ha létezik, egy pásztázó (keresőmotorok által használt szoftverprogram, amely egy honlap „pillanatfelvételének” elkészítésére szolgál) olvassa a vonatkozó honlap indexelését megelőzően. A tárgyhoz tartozó utasítások lehetővé teszik, hogy egy meghatározott pásztázót arra késztesse, hogy a honlapon meghatározott fájlokat és/vagy feljegyzéseket ne vegyen figyelembe. Az utasításokat a pásztázó úgy hajtja végre, hogy az alfanumerikus jelsorozatokat a szöveggel veti egybe annak a sorrendnek megfelelően, amelyet a robots.txt fájl tartalmaz. A protokoll alkalmazását korlátozza, ha egy kielégítő skálázhatóság hiányzik, ftp-szerverekkel nem működik, és az információ elvesz, ha a tartalmat a honlapról kimásolják⁸.

Alternatívaként a jelek különféle kategóriái, mint egy specifikus honlap jellemzői, használhatók (de egy adott oldal individuális elemével, például egy benne megjelenő képpel

⁶ Lásd továbbá: Recommendation CM/Rec(2012)3 of the Council of Europe on the protection of human rights with regard to search engines. (Az Európa Tanács CM/Rec(2012)3 ajánlása az emberi jogok védelmére a keresőgépek használata során.)

⁷ A robots.txt protocol-ra mint Robots Exclusion Protocolra (robotkizárási protokoll) és Robots Exclusion Standard-ra (robotkizárási szabvány) is hivatkoznak. A protokollt egy lejárt IETF Internet Tervezet definiálja. Hozzáférhető online: <http://www.robotstxt.org/norobots-rfc.txt>.

⁸ Az Internet tartalmának változásait és/vagy az indexelés preferenciáit a keresés eredményei olykor nem tükrözik. Jelentős problémának mutatkozik a keresőgépeket indexeik aktualizálására készíteni.

vagy fájlal összefüggésben is) annak jelzésére, hogy az adatot/oldalt a keresés eredménye nem tartalmazhatja.

Hangsúlyozni kell, hogy ezek a megközelítések egyaránt teljesen a hálózati etiketten (vagyis az érintett felek együttműködésén) alapulnak, ezért nagyon nehéz a kikényszerítésük. Végrehajtásuk a honlapon, és betartásuk a keresőmotorok által teljesen önkéntes. Míg ezek az indexelésnek azokat a kockázatait, melyeket harmadik személyek lapjaival létező kapcsolatok határoznak meg, csökkentik, nem képesek *per se* biztosítani, hogy egy meghatározott információt egy keresőmotor soha ne indexelhessen, különösen akkor, ha ez az információ nyilvánosan hozzáférhető, és azt más honlapok – eltérő hozzáférési szabályok szerint működő pásztázókkal – feldolgozhatják.⁹

4. Ajánlások honlap-adminisztrátoroknak

A honlap-adminisztrátorok döntő szerepet játszanak a törlés mindkét, fent körülírt kategóriája tekintetében, éspedig azért, hogy lehetőségük van az adatok rendelkezésre állásának és az adatok indexelhetőségének korlátozására. A fentebb megnevezett célok elérésére a Munkacsoport a következő ajánlásokat teszi:

- A honlapok működtetőinek tájékoztatniuk kell felhasználóikat arról, milyen adatokat tárolnak és milyen célból. Egyszerű mechanizmust kell felhasználok rendelkezésére bocsátaniuk, mellyel hozzáférhetnek személyes adataikhoz, és lehetővé teszik számukra, hogy azokat helyesbíthessék és/vagy véglegesen törölhessék, ahogyan arról az érvényes adatvédelmi jogszabályok rendelkeznek. E hozzáférési mechanizmusoknak felhasználóbarátnak kell lenniük, és a felhasználónak nem jelenthetnek semmiféle további költséget, nem járhatnak indokolatlan késéssel vagy gyakorlati teherrel.
- Az adatalany kifejezett kérelmére, ha egyéb jogszerű érdekek vagy jogerős korlátozások nem állnak fenn, a webmestereknek azonnal el kell távolítaniuk weboldalukról a vonatkozó információt. Ezen túlmenően jelezniük kell a keresőmotor-szolgáltatóknak, hogy újból indexelni kell a honlap érintett részét, hogy az adatokat a keresőindexből is, és a keresőmotorok valamennyi másolatát a gyorsítótárban törölje.
- A webmestereknek különleges eszközöket kell felhasználok rendelkezésére bocsátani, melyek lehetővé teszik, hogy a keresésre szolgáló indexelési preferenciáikat¹⁰ egyéni igényeinek megfelelően testre szabhassák. Vagylagos megoldásként mérlegelhető a „noindex“ meta-jel – melyet a szóban forgó oldal HTML kódjának vagy a http fejlécnek kell tartalmaznia –, vagy a sitemap.xml fájljának a használata, hogy a meghatározott adatokkal összefüggésben lévő szóban forgó keresési preferenciákat jelezze¹¹.

⁹ E tekintetben lásd az IWGDPT (Nemzetközi Távközlési Adatvédelmi Munkacsoport) ajánlásait is: “Common Position on Privacy Protection and Search Engines” (*Közös álláspont a magánszféra védelmére és a keresőgépekre vonatkozóan*); elfogadva 1999-ban, átdolgozva 2006-ban.

¹⁰ Vö.: a “blogger.com” platform által rendelkezésre bocsátott mechanizmus, amely lehetővé teszi a felhasználó számára, hogy indexelési preferenciáit a blog-szolgáltatásra való feliratkozásakor egy olyan adatlapon adja meg, amely tudatja a webmesterrel, hogyan konfigurálja az ő saját robots.txt fájlját (<http://buzz.blogger.com/2012/03/customize-your-search-preferences.html>)

¹¹ Ez a javaslat különösen jelentős dinamikus környezetek vagy bonyolult honlapok esetében, ahol a robots.txt megoldás esetleg nem eléggé illeszkedik a honlap méretéhez. Egy példa arra, hogyan használható a robots.txt parancs, hogy jelezze a keresőmotoroknak az oldal lejáratát dátumát:

<http://googleblog.blogspot.fr/2007/07/robots-exclusion-protocol-now-with-even.html>. Ehhez hasonlóan a sitemap.XML fájl jelzi, milyen gyakran változhat egy honlap, és azt a prioritási szintet, melyet a webmester tulajdonít egy URL-nek, lehetővé téve, hogy a keresőgép a potenciálisan legmegfelelőbb frissítési gyakoriságot válassza ki. Vö. a következővel is: <http://lists.w3.org/Archives/Public/public-privacy/2012OctDec/0224.html>

- Különös figyelmet kell fordítani a robots.txt fájl írásakor az utasítások lexikai és szemantikai helyességére, valamint inherens logikai konzisztenciájukra (hogy elkerüljük az ellentmondó és/vagy átfedő utasításokat). Hangsúlyozni kell, hogy ha a robots.txt fájlban *nincsenek specifikus kizárási utasítások*, a pártázó feltételezi, hogy az adminisztrátor megengedi a hely indexelését vagy specifikus alkönyvtárak indexelését (egy pártázó pl. feltételezi, hogy a honlap tartalmát hozzáférhetővé kell tennie a keresőmotorok számára).
- Meg kell azt is jegyeznünk, hogy a robots.txt protokoll nem alkalmas a különösen „kockázatos” tartalmakhoz – például elektronikus hírközlési szolgáltatások által generált forgalmi adatokhoz, SMS üzenetek tartalmához, hangposta-tárolóhoz, helymeghatározó adatokhoz, pénzügyi adatokhoz, stb.) – való hozzáférés szabályozására, és nem is célja egy honlap specifikus adminisztratív területeihez való hozzáférés elhárítása.¹² A robots.txt protokoll nem helyettesíti a kriptográfiát vagy a hozzáférés-ellenőrzés mechanizmusait.

Ha egy webmester jelezni akarja, hogy specifikus oldalakat és/vagy fájlokat egy keresőmotor nem indexelhet, különös gonddal kell az URL-eket kiválasztani. Valóban, mivel a robots.txt fájl a nagyközönség számára látható, „önmagát megmagyarázó” URL-ekre hagyatkozni azzal jár, hogy a vonatkozó tartalmak felfedésének lehetősége esetleg végérvényesen megnövekszik, miáltal a protokoll előnyei sem érvényesülhetnek. A robots.txt fájl tartalma különösen értékes a hackerek és bárki más számára, akik személyes adatok terjesztésére/ megszerzésére specializálódnak.

5. Ajánlások keresőmotorok számára

A keresőmotor-szolgáltatók – mivel ez fő tevékenységeik egyike – főként információ brókereként/közvetítőkként dolgoznak¹³. Mindazonáltal vannak bizonyos feldolgozás-típusok, melyek esetében önálló adatkezelőként cselekszenek.

Némely keresőmotor sok, különféle tevékenységet végez a honlapok indexelésétől a vonatkozó tartalmak átmeneti tárolásáig, hogy lehetővé tegye a felhasználónak az információ visszanyerését, ha egy szerver és/vagy egy kapcsolat nem működik/nem elérhető. Ez a gyorsítótárban megőrzött adat egy ismételt közzétételt jelent, ezért a keresőmotor-szolgáltató adatkezelőnek minősül¹⁴.

Ennek megfelelően a következő, keresőmotor-szolgáltatóknak tett ajánlásokat az általuk játszott szerepek szerint különböztettük meg.

Csupán indexelés

¹² 2011 júliusában mintegy 8000 SMS üzenet érkezett a MegaFon mobil hálózatra, melyeket a Yandex, egy orosz keresőmotor indexelt, s így a tartalom adatait és a címzettek mobil telefonszámait a nagyközönség számára hozzáférhetővé tette. Hasonlóképpen a Yale Egyetem 43 000 hallgatójának TAJ-száma került nyilvánosságra, miután azokat a Google 2011 augusztusában indexelte, mert egy ftp szerver nyilvános alkönyvtárában tárolták őket.

¹³ Vö.: az Európai Adatvédelmi Biztosok 29. cikk alapján létrehozott Munkacsoportjának 1/2008 véleménye a keresőmotorokkal kapcsolatos adatvédelmi kérdésekről (WP 148): http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2008/wp148_de.pdf. Vegyük figyelembe, hogy ezt az ügyet jelenleg az Európai Bíróságán tárgyalják.

¹⁴ Mint azt az Európai Adatvédelmi Biztosok 29. cikk alapján létrehozott Munkacsoportjának véleménye (WP 148) hangsúlyozza, ha „az indexált honlapokon található személyes adatok gyorsítótárazási időtartama meghaladja a műszaki elérhetőséghez szükséges időt, az önálló ismételt közzétételnek minősül. A Munkacsoport véleménye szerint az adatvédelmi jogszabályok betartása az említett gyorsítótárazási funkciók szolgáltatójának mint a tárolt közlésekben szereplő személyes adatok kezelőjének feladata.”

- A keresőmotoroknak mindig figyelembe kell venniük a honlapokon – az általuk befogadott tartalomra való tekintettel – a robots.text fájl vagy más „noindex” jelölési mechanizmus által kifejezett indexelési preferenciákat, ide értve a lejáratí dátum parancsokat is. Ezek az indexelési preferenciák a honlap első pásztázását megelőzően vagy akkor, ha már pásztázásukra sor került, kell kifejezni. Ez utóbbi esetben egy keresőmotor által végrehajtott indexelést e lehető leghamarabb aktualizálni kell.
- A keresőmotoroknak fokozniuk kell a webmesterekkel kommunikáló csatornáik hatékonyságát, hogy gyorsan értesüljenek bármely, az indexelési preferenciákban bekövetkezett változásról, melyet a webmesterek a robots.txt protokoll megfelelő parancsaival fejeznek ki, továbbá a honlap adatainak bármely egyéb módosításáról. Az aktualizálási/helyesbítési eljárásoknak a lehető legnagyobb mértékben magánszféra-barátnak kell lenniük, és különösen nem kérhetők további személyes adatok a azoktól a felhasználóktól, akik személyes információ bizonyos tételeinek aktualizálását/helyesbítését igénylik.
- Pásztázási sebességüket a keresőmotoroknak a webmester által kifejezett keresési preferenciákhoz kell igazítaniuk. A webmester bármely, honlapjaik vagy azok részeinek újra indexelési igényét, személyes adatok törlését vagy helyesbítését követően, haladéktalanul teljesíteniük kell.
- Minthogy jelenleg nincs konzisztens értelmezése a robots.txt fájlban írt utasításoknak vagy egyéb indexelési preferenciákban megjelenő jelző mechanizmusoknak (pl. metajelek, sitemap.xml fájlok), nehéz előre látni, hogy ilyen mechanizmusok milyen hatást fognak gyakorolni egy honlapnak a különböző pásztázók által való indexelésére. Kívánatos, hogy a keresőmotorok e tekintetben egy “modus operandi”-ban megegyezzenek. Az egyes utasításokra vonatkozó mechanizmusokat egy, a felhasználók számára egyszerűen (pl. a keresőmotor-portálokhoz a fő oldalairól) hozzáférhető oldalon világosan körül kell írni.
- A keresőmotoroknak nagyobb mértékben kell támogatniuk a honlap-adminisztrátorokat azáltal, hogy útmutatókat és/vagy eszközöket biztosítanak az indexelési preferenciák automatikus elemzésére. Ez képessé teszi az adminisztrátorokat arra, hogy ellenőrizzék, milyen hatást váltanak ki az általuk adott utasítások az indexelés körülményeiben.
- A keresőmotoroknak egyértelműbben kell meghatározniuk egy adott honlapon általuk végrehajtott „pásztázás” időzítését és kritériumait, hogy az adminisztrátorok és felhasználók egyszerűen eldönthessék, hogy egy keresés eredményeként nyert információ mennyi ideig maradjon hozzáférhető.

A pásztázott információ ideiglenes tárolása

- A keresőmotoroknak specifikus pásztázókat kell alkalmazniuk, ha adatokat különféle kategóriák és különféle célok (pl. általános indexelés, hírek, képek stb.) szerint akarják csoportosítani, hogy lehetővé tegyék a honlap-adminisztrátorok számára annak a szövegkörnyezetnek az ellenőrzését, amelyben az információ közzétételre kerül.
- Egy honlap indexelésekor a keresőmotoroknak összetettebb és részletesebb információt kell elfogadniuk pásztázók számára, például a következőket:
 - Specifikus célú információ indexelésének engedélyezése (pl. általános célú keresőmotorok számára hírek keresőmotorjaival szemben, stb.)¹⁵;

¹⁵ Vö. pl. azokkal a tényállásokkal, melyekről az Olasz Antitröszt Hatóság számolt be, miután az Újság Kiadók Olasz Szövetsége panasszal élt a Google ellen. Ezt követően a Google nyilvánosan kötelezettséget vállalt arra,

- Specifikus célú információ ideiglenes tárolásának engedélyezése, ide értve az időbeli korlátokat is (pl. gyorsítótárzás, töredékek);
- Információ közlésének engedélyezése specifikus célokra harmadik feleknek;
- Specifikus, például földrajzi területet vagy IP címtartományt felhasználó esetekre¹⁶ visszanyert adatok feldolgozásának engedélyezése.

Ha a pásztázást a hely tartalmának ideiglenes tárolása követi más célból, mint hogy a felhasználót, ha egy adott szerver/hálózat nem működik/nem elérhető, képessé tegye ezeknek a tartalmaknak a hozzáférésére, a keresőmotoroknak világos, specifikus információval kell ellátniuk a hely adminisztrátorait az említett tár időszerűségéről és technikai mechanizmusairól.

A keresőmotoroknak, a webmesterek specifikus, keresési preferenciáikat illető kérelmére, haladéktalanul törölniük kell a honlapokból visszanyert adatok valamennyi, gyorsítótárás másolatát, tartózkodniuk kell ezen adatok további feldolgozásától, csökkentve ezáltal az adatok terjesztésének és túlzott feltárásának kockázatát.

6. Végző fenntartások

Ebben a dokumentumban a munkacsoport a (személyes) adatok Interneten való hozzáférését ellenőrző eszközöket kutatta fel, amelyek manapság a felhasználók, webmesterek és keresőmotorok rendelkezésére állnak, és többnyire azon alapulnak, hogy korlátozzák egy honlap tartalmának felfedését vagy (automatikus) törlési mechanizmust alkalmazva¹⁷, vagy keresési mechanizmust jelző protokollok által. Emlékeztetnünk kell arra, hogy ez utóbbi mindazonáltal a pásztázókra érvényes egyszerű be/ki (bináris) szabályokon nyugszik, melyeket több, mint 15 évvel ezelőtt alkottak meg. Ennek megfelelően a keresőmotorok az évek során növekvő mértékben bonyolultabbá váltak, és a szóban forgó protokoll alapját képező, meglehetősen egyszerű beleértés/kizárás mechanizmus már nem teljesen képes megbirkózni az adatok visszanyerésének és tárolásának egyre növekvő lehetőségével. Hangsúlyozni kell például, hogy az adatok hozzáférhetősége (ideértve azokat az adatokat, melyeket saját magukról a felhasználók tesznek közzé) kombinálva az arcfelismerési technikákkal és helymeghatározó adatokkal, végtére is jobban lehetővé teszi az egyének indexelését, mint egyszerűen a tartalmakét vagy információkét. Ezekre a szempontokra ezért sürgősen összpontosítani kell.

A személyes adatok Interneten való jobb védelmét szolgáló újabb várható technológiai áttörést hozhat a „policy-aware, szemantikus Internet” megjelenése, ahol az adatokat jellemzőkkel (pl. egy „jelentéssel”) és hozzáférési szabályokkal lehet elválaszthatatlanul összekapcsolni. Ez egyrészt lehetővé teszi új összefüggések létesítését az adatok között és erősíti egy összekapcsolt világ koncepcióját, másrészt hatékonyabb mechanizmust nyújt a tartalom azonosítására és helyének meghatározására, és potenciálisan másolja azt az információt, amely korrelál a jellemző-összehasonlításon alapuló tétellel (még hozzá jobban, mint egy egyszerű szöveg-összehasonlításon, mint az manapság történik). Így elképzelhetővé

hogy a kiadókat olyan eszközökkel látja el, amelyek segítik megkülönböztetni az általános keresőmotorok tartalmának indexelését a hírek keresőmotorjainak indexelésétől.

¹⁶ Minthogy a keresőmotorok által pásztázott információra vonatkozó felhasználói esetek bonyolultsága növekszik, helyénvaló lehetne az aktuális séma ellentétének figyelembe vétele, amely megengedi, hogy a pásztázók információt olvassanak, ha egy utasítás formálisan helytelen vagy a pásztázó által nem értelmezhető.

¹⁷ Érdemes rávilágítani, hogy az Internet publikus természetének köszönhetően, ha egy honlap-adminisztrátor tartalmat akar eltávolítani a „nyilvános területről”, más hozzáférés-ellenőrző mechanizmusokat kell alkalmazni, például a felhasználó azonosítását és/vagy az adatok titkosítását.

válik az információ eltávolítása a legkülönbébb honlapokról, és a keresés eredményei honlapokkal való kapcsolatának megszüntetése, mely által elkerülhető az adatok bármiféle, akaratlan terjesztése¹⁸.

Az Internet használhatóságát természetesen nem szabad aláásni, és egyensúlyt kell teremteni az innováció és az egyének magánszféráját és adatai védelmét szolgáló alapvető jogai között. További megfontolást igényel a részletesebb mechanizmusok alkalmazásának lehetősége, amelyek nem az egyszerű beleértés/kizárás szabályon alapulnak, hanem inkább megkísérlik az adatalanyokat képessé tenni arra, hogy jobban kifejezzék saját keresési preferenciáikat, és az információt a megfelelő szövegkörnyezethez kapcsolják (megengedve például, hogy az adatalanyok jelezzék, egy adott információ még mindig aktuális vagy releváns, vagy jelezzenek bármely eseményt, amely esetleg hatással volt erre az információra). Ez az adatalanyoknak több lehetőséget kínál, mint az egyszerű választás a honlapon való túlzott feltárás és az új technológiáktól való teljes tartózkodás között.

Gazdasági természetű, fontos, erősödő érdekek fűződnek mind a keresőmotorokhoz, mind a honlap-adminisztrátorokhoz, arra készítetve őket, hogy az adatoknak az adatok és az információ indexelésével a lehető legszélesebb hozzáférhetőségét biztosítsák. A honlapok ilyen indexelése egyes piaci szereplők gazdasági érdekét szolgálja, és eltávolítja a nagyközönség rendelkezésére álló honlap-tartalmat vagy jelzi, hogy efféle tartalmakat a keresőmotoroknak nem szabad indexelni és visszanyerni, mert az hatást gyakorol a piac dinamikájára és az üzleti modellekre. A szóban forgó érdekek és az adatvédelmi igények összehangolásához a különféle befektetők együttműködésére van szükség.

¹⁸ A Google nemrégiben bejelentette, hogy aktualizálja keresési algoritmusait, ami csökkenti az eltávolításra utaló, nagy számú figyelmeztetéssel ellátott honlap rangsorolását, amit csak a szerző jogok megsértése eseteiben alkalmaz. (<http://insidesearch.blogspot.fr/2012/08/an-update-to-our-search-algorithms.html> vagy <http://www.google.com/insidesearch/howsearchworks/>).