Работен лист за Big Data и приватност

Принципи на приватност под притисок во ера на аналитика на Big Data

55 Состанок,5-6 Мај 2014,Скопје

Вовед[1]

1. Big Data е поим кој го објаснува огромниот раст во пристапот и автоматската употреба на информации[2]. Тоа се однесува на огромни количини на дигитални податоци контролирани од компаниите, органите и други големи организации што се предмет на опширни анализи базирани на употреба на алгоритми.[3]
2. Big Data сами по себе претставуваат предизвик за основните принципи на приватност. Некои тврдат дека би било невозможно да се применуваат овие принципи во ера карактеризирана со Big Data.[4] Од аспект на ова гледиште, заштита на пrivатноста мора да биде загарантирана преку обезбедување јасни и разбирливи информации за одржување на личните податоци. Мислењето на работната група е дека заштитата на приватноста денес е поважна од кога и да било , од причина што се собираат се повеќе информации за поединци.[5] Принципите на приватност ни гарантираат дека нема да бидеме предмет на опширни профилирања со растечки потенцијал. Разводнување на овој принцип во комбинација со уште поголема употреба на Big Data може да има последици врз заштита на приватноста и други општествени вредности како слобода на изразување и размена на идеи.
3. Некои носивни принципи пропишани од страна на OECD и Европската Директива за заштита на лични податоци се однесуваат на тоа како личните податоци можат да се обработуваат на разумен, коректен и легален начин. Следните принципи се од особена важност за Big Data: намерни ограничувања, важност и минимизација на податоците, целост и квалитет, транспарентност и право на пристап до информации.

---

[1] Овој документ содржи референци на правните барања кои не мора да бидат релевантни на сите овластуваа пресентирани во Работната Група.

[2] Cf.hite(2012), Big Data е е поим кој се однесува на збирки на податоци кои се големи и комплексни, кои се тешки и комплексни за обработка преку рачни алатки за управsвање со датабази или традиционални програми за обработка на податоци.

[3] Член 29 Работно тело, Став 03/2013 за намерни ограничувања, стр.35

[4] На пример, For example in Tene, Omer and Jules Polonetsky (2012), "Big Data for All: Privacy and User Control in the Age of Analytics", Northwestern Journal of Technology and Intellectual Property, Forthcoming, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2149364 , in World Economic Forum (2013), "Unlocking the Value of Personal Data: From Collection to Usage",http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf, and in Cate, Fred H. and Viktor Mayer-Schönberger (2013), "Tomorrow's privacy. Notice and consent in a world of Big Data", International Data Privacy Law, 2013, Vol. 3, No. 2

[5]

Досег

4. Целта на овој работен лист е да ги потенцира предизвиците на приватност поврзани со Big Data,во прв план во телекомуникацискиот сектор, да помогне во осигурување дека овие прашања се вклучени во агендата на органите за заштита на податоци и други стејкхолдери. Документот е наменет за донесувачи на одлуки, јавни органи, индустриско и цивилно општество.

5. Концептот на Big Data опфаќа широк спектар на предизвици. Голем број од нив, како што се ризикот од повторна идентификација, лесно може да биде предмет на опширни извештаи. Целта на овој работе лист е да се се прикажат клучните предизвици на приватност и да не се насочуваат поединечните проблеми од техничка природа во детали.

Позадина

6. Податоците се насекаде. Количеството на податоци на глобално ниво расте со 50% годишно. 90% од светските податоци се собрани само во последните 2 години. Поголем дел од овие податоци е собран директно од потрошувачите преку интернет-базирани услуги. Со подемот на Интернетот на нештата ќе се введат нови податочни струи. Се предвидува дека до 2015 ќе постојат повеќе од 50 билиони сензори, кои ќе прикачуваат информации во компјутерски облаци, а кои ќе се однесуваат на интерацијата на луѓето со околината. Ова може да предизвика промени во пазарните и бизнис моделите.

7. Постои сомнеж дека можноста за складирање и анализа на големи количини на податоци ќе донесе бенефиции за општеството на различни начини. Big Data до одреден степен веќе се користеше за анализирање на податоци за идентификување и предвидување на трендови и корелации. Во основа, ваквите анализи не претставуваат предизвик во однос на приватноста, со оглед на тоа што податоците се целосно анонимни (кконцептот на анонимност се разгледува подетално подолу во овој документ). Посотојат и анализи кои не вклучуваат употреба на лични податоци во целина, како на пример анализи на подтоци за временската состојба или податоци од опремата на нафтените платформи.)

8. Big data може да се употребува и на начини кои директно влијаат врз поединците. Постојат техники кои се употребуваат за изготвување профили и предвидување на однесувaето на поединци и групи преку собирање и анализа на податоци собрани од различни извори. Иако информациите мже да се собрани и неидентификувани, резултатот од анализите сеуште може да предизвика последици за поединците.

9. “Лични податоци „ се секоја информација поврзана со идентификувана личност или личност која може да се идентификува. IP-адреси, броеви на мобилни телефони, RFID маркери и Единствени идентификациони броеви (UDID numbers) се само дел од

примерите за единствени идентификатори кои се сметаат за лични податоци. Податоците кои откриваат информации за навиките и интересите на уникатни индивидуи се бараат од страна на компаниите и владините органи. Индустријата развива нови техники за овие цели, како на пример уреди за отпечатоци од прсти. Како резултат на тоа, листата на единствени идентификатори дефинирани како лични податоци постојано се шири.

10. Синџирот на вредности на Big Data вклучува неколку чекори почнувајќи од собирање на податоци, складирање, анализа и употреба на резултатите од анализата (види дијаграм на вредносен синџир на крајот од документот). Ќе се осврнеме кон поединечни чекори подолу во документот.

11. Првиот чекор во синџирот на вредност е собирање на податоци. Примери за потенцијални извори на лични податоци вклучуваат апликации за мобилен телефон, компјутеризирани мрежи, транспондери во возила за детектирање патарини, евиденција за пациенти, податоците за локација, социјални веб-сајтови, податоци за патници во воздушниот сообраќај, јавни регистри, програми за лојалноста на клиентите, историјат на продажба итн. Подемот на технологијата на сензори,ќе овозможи да се собираат информации од широк опсег на паметни уреди, како што се паметни четки за заби, паметни чадори, паметни фрижидери, паметни чевли, паметни телевизори и др. Овие извори на податоци можат да обезбедат информации кои потенцијално би можеле да откријат многу за начинот на живот на секој поединец.

12. На пример, личните податоци може да се прибираат на следниов начин:

i. Личните податоци можат да се објавуваат своеволно од страна на поединци (на пример, преју објавување лични податоци на социјалните мрежи).
ii. Лични податоци можат да се собираат како услов за услуга.
iii. Лични податоци можат да се собираат врз основа на законските барања.
iv. Личните податоци може автоматски да се собираат со користење на одредени услуги (на пример, трансакција податоци на патарина и податоците за локација). Ваквите збирки на податоци, може да се спроведуваат без знаење на субјектот на податоци.
v. Личните податоци може да бидат изведени со обработка и анализа на собраните податоци за претходни и други цели. Личните податоци, исто така, може да потекнуваат од различнигрупи на информации кои се анонимни.
vi. Личните податоци (на пример, податоците за клиентите (CRM)) може да се земат од надворешни извори за да се збогатат (претходно собраните) податоци.
vii. Личните податоци (на пример, (детална евиденција за клиентите) може бидат споделени со надворешни извори ,со цел збогатување на (лични) податоци од партнерски фирми.

13. Во контекст на Big Data, податоците собрани за веб корисниците се мошне атрактивни бидејќи можат да содржат детални информации за интереси, мрежи, навики и

однесувања на поединци. Таквите информации можат да се собираат експлицитно (на пример, при регистрација на социјален профил на интернет) или во поприкриена форма, со користење на различни видови на технологии за следење.

14. Вториот чекор е агрегирање и чување на податоците откако ќе се соберат. Некои организации ги собираат и анонимизираат податоците пред да ги складираат, додека други складираат податоци кои содржат лични идентификатори. Брзиот раст во чување и аналитички моќ, по се пониска трошоци, значи дека принципот на Big Data повеќе не е привилегија само на неколку гигантски компании, туку претставува алатка достапна и за мали и големи претпријатија, во ситесектори на економијата. Концептот на Big Data претставува префрлување од традиционалнитеразмислување во однос на чување и обработка на податоци со помош на супер-компјутери. Новите технологии овозможуваат обработка и генерирање вредност од нови и неструктурирани извори на податоци.

15. Третиот чекор од синџирот на вредност вклучува корелација и анализа на собраните и складирани податоци. Централен елемент во создавање на вредност при овој чекор е спојување на податоците од различни извори со цел генерирање профили и користење алатки за анализа со чија помош се изведуваат информации кои инаку не би биле достапни. Корисниците на Big Data можат да генерираат податоци интерно од своите податочни капацитети или тие можат да купат податоците од трети страни (или добијат податоци од отворени извори), и притоа ги комбинираат овие со своите податоци. Примери за техники за анализа поврзани со Big Dat вклучуваат: Data Mining, машинско учење, анализа на социјална мрежа, предвидливи анализи "Sensemaking", Обработка на природни јазици и визуелизација.

16. Четвртиот чекор од синџирот на вредност вклучува користење на резултатите од анализата. Big Data може се користат на различни начини. Зголемен број корисници, на пример банки, осигурителни компании, агенции за кредитен рејтинг, работодавците и полицијата сакаат да ги искористат резултатите добиени преку анализа на Big Data со цел овозможување на подобри, полни со информации одлуки.

17. Голем број на заинтересирани страни се вклучени низ целиот синџир на Big Data (види Слика 1 во прилог). Некои стејкхолдери се инволвирани само во одделни делови од синџирот на вредност. На пример, брокерите со податоци обично не го користат податоците за лични цели, туку ги обработуваат и го продадаваат. Други субјекти можат да бидат вклучени во сите чекори низ синџирот на вредност. На пример, трговците на мало прибираат лични податоци преку програмите за лојалност за клиентите, потоа ги чуваат и на крај се процесираат и ги користат во својот бизнис модел.

18. Личните податоци веќе долго време се мошне привлечна стока и влегуваат во развој на нови Интернет-базирани услуги. Интернет корисниците обично добиваат пристап до бесплатни услуги плаќајќи за нив во форма на лични податоци. Поради Big Data и експанзивното ширење на Интернетот на нештата, пазарот за продажба на личните

податоци постојано расте и веројатно ќе се прошири и на нови сектори во економијата: Паметни чевли со сензори ќе се нудат бесплатно, во замена за дозвола од корисникот за собирање и анализа на податоци од неговите пешачки активности. Стоматологот може да му понуди на пациентот паметен четка за заби (која бесплатно ќе биде обезбедена од страна на производителот) во замена за споделување на информациите собрани од четката за заби со различни заинтересирани стејкхолдери. Нови бизниси и бизнис модели ќе произлезат за максимизирање на додадена вредност од личните податоци кои се генерирани во сè поширок спектар на ситуации.

Последици врз приватноста

19 .Врз основа на горенаведените осврти, следните клучни предизвици за приватноста се јавуваат како резултат на употреба на Big Data за нови цели:

20. Во голема мера, концептот на Big Data вклучува повторна употреба на податоците. Ова претставува предизвик за принципот на приватност дека собраните податоци може да бидат искористени за цели што се некомпатибилни со првобитната намена за собирање. Потенцијалот на Big Data за собирање на вредни знаења преку збирка на сé поголеми сетови на податоци го води принципот на целни ограничувања под притисок. Според овој принцип, претпријатија кои ги користат личните податоци како основа за предвидувачка анализа мора да гарантираат дека анализата е компатибилна со првичната цел за собирање на податоците. Кога поединец разменува податоци со други, тие имаат одредени очекување за целите за кои податоците ќе се користат. Субјетите на податоци не ги даваат информациите на компаниита или на Владата за тие да прават што сакаат со нив. Ова може да претставува значителен предизвик за комерцијалната употреба на големи податоци.

Максимизирање на податоците

21. Концептот на Big Data е и за максимизација на вредноста на податоците. Во суштина, Big Datae спротивно на принципот на приватноста во поглед на важноста и минимизирањето на податоците. Овие принципи се наменети за осигурување дека нема да се прибираат повеќе лични информации од она што е потребно да се исполнат јасно дефинираните цели. Податоците мора да бидат избришани кога повеќе не се употребнуваат за нивната првична намена. Концептот на Big Data подразбира нов начин на разгледување на податоци, каде што податоците сами по себе генерираат вредност. Вредноста на податоците е поврзана со потенцијалните идни користи. Ваквите ставови за податоците можат да претставуваат предизвик на приципот за приватност кој утврдува дека податоците кои се обработуваат мора да бидат соодветни, релевантни за целите што се утврдени за време на собирањето. Тоа може да влијае и врз желбата на конторлорите и нивната мотивација за бришење на податоците. Компаниите и јавните органи не би сакале да ги избришат податоците кои во иднина би им биле извор на нови идеи и приходи. Распоространување на употребата на Big Data ќе претставува огромен предизвик за

оганите за заштита на податоци во правец на присилување на субјектите да ги бришат податоците.

Недостаток на транспарентност

22. Правото на пристап до вистинските информации во однос на обработка на нечии лични податоци конституира важни принципи на приватност. Недостаток на отвореност и на информации за начините на кои податоците се прибираат и се употребуваат може да преставува жртва која не се разбира и врз која се нема конторла. На пример, просечен Интернет корисник знае многу малку за тоа како пазарот на онлајн рекламирање функционире и како неговите лични податоци можат да се соберат и искористат за различни комерцијални цели. Голем број луѓе не се запознаени со "операторите" на пазарот, посебно со брокерите на податоци и компаниите за анализа. Ова предизвикуа потешкотии во примената на правотот на поединецот за пристап до информации.

Собирањето податоци може да открие чувствителни информации

23. Предизвикувачки аспект поврзан со анализите на Big Data е можноста прибраните информации , кои сами по себе не мора да бидат од чувствителен карактер, да генерираат чувствителни информации како резултат. Со користење на алатките на Big Data, постои можност за утврдување шеми кои би предвиделе бројни предиспозиции на луѓето: како на пример здравствени информации, политички гледишта или сексуална ориентираност. Ова претставува посебен информационен субјект за кој контролорите на податоци треба да бидат свесни дека носи ризик со собирање и анализирање на податоците.

Ризик од повторна идентификација

24. Еден од главните ризици поврзани со анализите за Big Data е оној од повторна идентификација. Преку собирање на информации од неклоку извори, постои ризик личните податоци на поеднци да да се издвојуат од збирките на податоци кои треба да бидат анонимни на прв поглед. Ова ја прави анонимизацијата како метод помалку ефективна во превенција на проблемите со приватност поврзани со правење профили и други анализи на податоци. Ризикот на повторна идентификација може да се намали преку осигурување дека само анонимни податоци се користат во анализите. Сепак, не е секогаш лесно да се одреди дали збирките на податоци се доволно анонимизирани. Ова може да се покаже како тешко поради две причини:

- Првата е дека поимот „да се идентификува" , па и да се „ анонимизира " е комплициран процес бидејќи поединците може да бидат идентификувани на различни начини. Ова вклучува директна идентификација , каде личноста ќе може целосно да се идентификува од еден извор на информации (на пример листа на цели имиња), и индиректна идентификација каде со комбинација на податоци од два или повеќе извори ќе се изврши идентификација.

- Втората причина е што претптијатијата кои користат збирки од анонимизирани податоци не знаат дали постојат и други податоци кои би им овозможиле на трети лица да ги ре-идентификуваат поединците од збирките на анонимизирани податоци. И по бришењето на идентификационите податоци, сеуште би било можно поврзување на одреден информации со поединци врз основа на врски помеѓу збирки на Big Data. Практичен пример за овој проблем е "Како да се разбие анонимноста на збирките на податоци на Netflix Prize„.

Безбедносни импликаци

25. Концептот на Big Data вклучува и предизвици во однос на информационата безбедност што може да предизвика последици врз заштита на приватноста. Примери на вакви сигурносни предизвици вклучуваат употреба на неколку инфраструктурни слоеви за обработка на Big Data , нов вид на инфраструктура за управување со огромните текови на податоци како и неопфатно шифрирање на големи сетови на податоци. Понатаму, пробивање на овие податоци може да има сериозни последици кога се складираат големи количини на податоци. Компаниите кои поседуваат и одржуваат големи збирки на лични податоци мора да сносат одговорност за истите.

Неточни податоци

26. Одлуките кои предизвикуваат последици врз поединците мора да бидат засновани на точни информации. Употребата на моќни технологии за „копање низ податоците" е мошне популарно во областа на осигурувањето и кредитниот рејтинг. Со помош на Big Data опсегот на употреба е проширен, и постојат нови видови на извори на податоци при подготовка на кредитни резултати профили на ризик. Денес постојат и нови кредитни агенции специјализиран за концептот на Big Data, кои изготвуваат профили на поединци засновани на информации исклучиво од онлајн извори.

27. Засновување на одлуките на информации добиени од социјалните медиуми, не вклучува ризик дека одлуките ќе се базираат на неточни информации. Одлуките засновани на вакви информации нема да бидат во потполност транспарентни како одлуките кои се базираат ан информации од официјални регистри. Слабост поврзана со алализите на Big Data е тоа што контекстот во кој се употребуваат податоците не се зема во предвид. Дури и во случаи кога податоците се точни , може да постојат проблеми со приватноста поврзани со контекстот на употреба. Засновање на одлуките на информации прибрани за други цели може да доведе до резултати кои не ја рефлектираат вистинската ситуација. Мошне важно е да се напомене дека користење на податоци за други намени од првичната е противзаконски од аспект на заштитата на податоци, освен ако другите цели се компатибилни со првичните или кога податоците се анонимизирани.

28. Транспарентноста во форма на правото за запознавање на поединецот со содржината на обработените информации е предуслов за субјектите на податоци да ги заштитат своите интереси. Клучен принцип на приватност е дека поединци можат да бараат податоци, проценки и наоди кои не се точни да бидат поправени или избришани.

Дизбаланс на моќта

29. Поединците, како генерално правило , имаат ограничена моќ да влијаат врз начинот на однесување на големите корпорации. Преголема употреба на анализите на Big Data може да го зголеми дизбалансот помеѓу големите претпријатија од една страна, и потрошувачите од друга. Компаниите кои собираат лични податоци се тие кои бенефицираат огромната вредност стекната како резултат на анализи и обработка на податоци, а не поединците кои ги остават своите лични податоци. Дури, тоа може да се препише и како слабост на поединците кои можат да трпат негативни последици во иднина (на пример, при можности за вработување, банкарски заеми и можност за здравствено осигурување).

Одредување на податоци и дискриминација

30. Концептот на Big Data се базира на претпоставката дека колку повеќе податоци се прибрани и до кои се има пристап, толку поразумни и правилни одлуки ќе се носат. Но собирањето на повеќе податоци не мора да значи и добивање повеќе знаење. Тоа може да предизвика конфузија и лажни реазултати.
Големата употреба на автоматски одлуки и предвидуваќки анализи може да има последици по поединците. Алгоритмите, иако неприродни, пренесуваат избори, меѓудругото, за податоци, врски,заклучоци, толкувања и инклузивни прагови кои унапредуваат специфична цел.
Big Data во одредени случаи може да вклучува и предрасуди и стереотипи, како и да предизвика социјално исклучување и стратификација. Употребата на коралциони анализи може да доведе до целосно неточни резултати за поединците. Корелацијата најчесто се меша со каузалност. Ако анализите покажуваат дека ако веројатноста поединец X да биде изложен на Y е 80%, не може да се заклучи дека ситуацијава ќе се појави во 100% од случаите. Притоа, дискриминацијата заснована на статистички анализи може да причини проблем на приватноста. Развој, каде се повеќе одлуки во општеството се базираат на употреба на алгоритми може да резултира во “Диктатура на податоци„ каде одлуките и пресудите не би се носеле врз основа на вистинска ситуација туку истите ќе се основаат на податоци кои прдвидуваат што би можело да се случи.

Ефект на застрашување

31. Ако развојот се темели на кредитни резултати или осигурителни прмиии засновани примерно врз информации собрани од различни контексти на Интернет и други секојдневни ситуации, тоа може да влијае врз заштита на приватноста и начинот на однесување. За десет години на пример, нашите деца би можеле да не добијат осигурување зошто претходно на социјалните мрежи имаат објавено дека имаат генетски предиспозиции кон некоја болест. Ова може да резултира во воздржаност на поединците на социјалните мрежи во голема мера или приспособување на однесувањето онлајн и воопшто. Постои страв дека трагите кои ги оставаме во различни ситуации може да влијаат врз нашите идни одлуки, како што се можноста за наоѓање работа, добивање заем,осигурување и.т.н.
Тоа може да ги одврати корисниците од барање алтернативни гледишта онлајн во односна стравот од страв да бидат идентификувани, профилирани или откриени. Во однос на употребата на Big Data од страна на органите, неизвесноста која произлегува од тоа кои извори на податоцисе употребуваат за прибирање на податоци и како тие се употребуваат може да претставува закана врз нашата сигурност во органите. Ова може да има негативно влијание врз основите за отворена и здрава демократија. Ниското ниво на заштита на нашата пруватност може да ја загрози деократијата бидејќи граѓаните ги ограничуваат своите учестав во презентирање на своите гледишта. Во најлош можен случај, големата употреба на Big Data може да има застрашуваќки ефект врз слободата на изрзаување ако претпоставките за таквата употреба не се откриваат и не може да се верификуваат независно.

Ехо комори

32. Персонализацијата на веб, со кастамизирани медиуми и и сервиси за вести засновани врз однесувањето на поединци навеб, имаат влијание и врз зацртаните услови за јавни дебати размена на иде- битни премиси за здрава демократија. Во прв план, ова не е предизвик само на приватноста, туку и за општеството во целина. Опасноста поврзана со таканаречените “ехо комори “ или “филтер балони„ е дека населениет ќе биде изложено само на содржини кои ги потврдуваат нивните ставови и вредности. Размената ан идеи и ставови може да се стави под контрола кога поединците се поретко изложени на ставови различни од оние кои ги поседуваат.

Препораки

33. Наспроти фактот дека Big Data повлекува неколку предизвици врз приватноста, постои можност за употреба на анализите без загрозување на клучните принципи на приватност. Работна група ги дава следниве препораки кои се однесуваат на тоа како Big Data би се употребувале на начин кој ја почитува приватноста на секој поединец.

Согласност

34. Се дискутира дека согласноста како легална основа за обработка на лични информации нема да функционира во ера на Big Data. Постојат тврдења дека барањето согласност на Интернет може да резултира во посиромашна заштита на поединците. Веќе се гледа напредокот кај фирми кои бараат согласност од нивните потрошувачи, шпекулирајќи дека исказите за согласност често не се разгледуваат во детали и се користат на други цели во иднина. Ваквата употреба е нелегална.

35. Несомнено иако постојат предизвици при добивање значајна согласност, таа останува основата на модерното законодавство за приватност. Намаленото користење на согласност се заканува да ја намали контролата ан поединците врз употребата на нивните податоци. Согласноста е една од неколкуте правни основи врс кои се темели обработката на податоци. Таа има важна улога, но тоа не ја исклучува можноста , во зависност од ситуацијата, некоја од другите правни основи за обработка на податоци да биде посоодветна од аспекта на контролорот и субјектот на податоци.

36.Валиднасогласност треба да се обезбеди од субјектот на податоци во склоп со личните податоци за целите на анализите и профилирањето.

37. Во случаи кога не е возможно да се бара согласност, обработката на податоци ќе биде возможна во склоп на внимателно избалансирани лимити. На пример, конторлорите на податоци може да обработуваат податоци ако обработката е неопходна за цели на правните интереси на конторлорите се додека истите не се надминати од страна на оние на поединците. Конторлорите на податоци мора да ги избалансираат двата спротивни интереси-правниот интерес и интересот на поединецот-еден наспроти друг. Резултатот од балансирање на интересите ќе се разликува од случај до случај, во зависност од кои интереси поврзани со приватноста на поединци се во прашање, како и правните интереси на контролорот. Колку позначајно е влијанието врз субјектите на податоци, толку поголемо влијание треба да се обрне кон релевантните заштитни мерки.

38. Контролорите на податоци кои сакаат да ги користат собраните податоци за цели различни од првичните мора да ја проценат компатибилноста помеѓу првичните од новите намери на основа случај по случај. Се додека компатибилност на тестот не е задоволена, личните идентификациони податоци не можат да се обработуваат.

Процедури за стабила анонимизација

39. Конторлорите на податоци мора да одлучат дали личните податоци употребени во Big Data анализите ќе бидат анонимизирани, прикажани со псевдоними или неидентификувани. Овој избор би го одредил начинот на кој легислативата која се

однесува на заштита на податоците ќе влијае врз компаниите во идните обработки на податоци. Ананонимните податоци отпаѓаат од спректарот на легислативата за заштита на податоци.

40. Анонимизацијата може да помогне при намалување или елиминирање на рзиците по приватноста поврзани со Big Data анализите, само во случаи кога анонимизацијата е правилно спроведена. Анонимизацијата настанува како резултат на обработката на личните податоци со цел да се спречине неповратност на идентификацијата. Во тој смисол, неколку елементи треба да се земат во предвид од страна на контролорите на податоци, имајќи ги во предвид сите средства за "разумна„ идентификација (од страна на контролорите и од трети лица). Мошне важно е да се тестира анонимизацијата во однос на прифатливите нивоа на ризик. Ова треба да биде документирано како дел од Оценките за влијани на приватноста.

41. Оптималното решение за анонимиозирање на податоците треба да се носи поединечно за секој случај, притоа користејќи комбинации од техники. Неколку техники за анонимизација може да пресдвидат, воглавно составени од рандомизавија и генерализација на податоците. Имајќи ги во предвид главните предности и слабости на секоја од техниките може да помогне во одредување на начинот на осмислување на одреден процес за анонимизација. Стабилноста на секоја техника треба да се заснова врз 3 критериуми:

I. Дали сеуште посто можност за издвојување на поединец.
II. Дали сеуште постои можност заповрзување на податоци со поединци
III. Дали информациите можат да бидат изведени за поединци?

42. Псевдонимизираните податоци не се еквивалентни на анонимизираните податоци. Контролорите на податоци кои одбираат да ги прикажат податоците со псевдоними, наместо со анонимизирање, мора да бидат свесни дека информациите сеуште се дефинираат како лични податоци и мора да бидат заштитени.

43. Треба да постои високо ниво на грижа кога се споделува или објавува псевдонимизирани или податоци кои може да се идентификуваат. Ако податоците се прикажани детално, може да се поврзат со други збирки на податоци, како и ако содржат лични податоци , пристапот до нив треба да биде ограничен и внимателно контролиран. Ако податоците се собираат и постои помалку ризик од поврзување со други збирки на податоци, постои поголема веројатност дека податоците може да се направат достапни без некој позначаен ризик.

44. Ако контролорот на податоци овозможува псевдонимизираните или податоци кои можат да се идентификуваат на друг начин да бидат достапни за други организации, треба со договор да им се забрани да ги ре-идентификуваат податоците. Ова исто така треба да вклучува отворени податоци.

45. Работната група дава препораки во насока на тоа дека мрежа или тело треба да биде основано кога некој кому му треба анонимини или псевдонимни податоци може да расправа за предизвиците поврзани со анонимизација . Вакви мрежи постојат во Велика Британија (theUK Anonymisation Network (UKAN))кој е координираба од универзитетите во Манчестер и Саутхемптон, Институтот на отворени податоци и Агенцијата за Национална Статистика.

Поголема транспарентност и контрола од собирање то употреба на податоците

46. Секој поединец треба да биде информирана за тоа кои податоци се собираат и како истите се употребуваат,за кои цели ќе се користат и дали ќе бидат дистрибуирани до трети страни.

47. Секој поединец треба да има пристап до неговиот профил и до податоците кои контролорите ги имаат за него. Секој поединец треба да биде информиран за изворите на лични податоци. Понатаму, согласно Законот, тие треба да бидат во можност да ги корегираат нивните податоци и да ја отселектираат можноста за понатамошно собирање на податоци.

.

48. Системот на класификација може да има несакани последици за поединците. Секој поединец треба да има пристап до информации за тоа кои алгоритми се употребени како основа за профилирање на процесот на донесување одлуки. Информациите треба да се претстават во читлив и јасен формат. Ова е важно од аспект на заштита од нефер дискриминација, како и да се избегне ситуација кога значајни одлуки за поединци би се базирале на неточни факти.

49. Секој поединец кој бара информации за себе, треба да ги добие сите податоци кои котролорите ги поседуваат во форма на лесен, преносliв и машински-читлив формат. Ова би го олеснило процесот на менување на снабдувачот на инфомрации со некој кој нуди најдобри услови , вклучитекно и заштита на приватноста. Преносливоста на податоците ќе ги заштити купувачите од нивно заробување во услуги со неприфатливи услови. Со тек на време, ваквите барања би резултирале со развој на услуги во корист на корисниците. Ова може да им помогне и на субјектите на податоци да ги подобрат нивните сфаќања за податоците поврзани со нив.

Приватност низ целиот процес и отчетност

50. Поцврсти техники на анонимизација, сами по себе, не би ги решиле предизвиците на Big Data кон приватност. Постои потреба за дополнителни решенија. Приватност низ целиот процес и отчетност се принципи кои помагаат во олеснување на предизвиците по приватноста.

51.Употребата на технологиите на Big Data треба да се засноваат на седумте принципи на Приватност низ целиот процес. Приватноста низ целиот процес налага да се зема во предвид заштитата на приватноста низ развојни фази, во процедурите и бизнис практиките.

52. Со цел зашувување на довербата на оние чиишто подароци се собираат, обработуваат и анализираат, мошне важно е да се проценат предизвиците во рамки на заштита на приватноста, колку што е можно порано, во голем број случаи пред обработка на Big Data. Ова може да се направи во форма на Проценка на влијанието врз приватноста (ПВП). ПВП треба да склучува евалуација на било каква правна основа за дистрибуција и о повторна употреба на личните податоци, како и евалуација на приватноста и заштитни мерки за безбедност. Ваква проценка треба внимателно да ги евалуира сите потенцијални последици врз субјектите на податоци.

53. Отчетноста е важен принцип за приватноста. Отчетноста гради доверба помеѓу субјектите на податоци и контролорите на податоци.Контролорите на податоцитреба да покажат дека тие се отчетливи и можат да носат одговорни и етички одлуки околу употребата на Big Data. На пример, контролорите на податоци треба да бидат свесни дека анонимизираните сеуште може да влијаат врз поединците, Анонимизираните збирки ан податоци може да се користат за збогатување на постојните профили на поединци, преку креирање проблеми со заштита ан податоците. Профилите и алгоритмите бараа континуирани проценки . Ова бара редовни контроли кои би осигуриле дека одлуките кои произлегуваат од профилирањето се одговорни, фер, етички и компатибни со целите за кои се употребуваат профилите. Неправдите кон поединците како резултат на целосно автоматизирани лажни позитивни и негативни резултати треба да бидат избегнати.

Зголемување на нивото на знаење и свест

54. Знаењето и свеста за предизвиците на приватноста поврзани со Big Data се важни за котролорите на податоци кои ја употребуваат оваа технологија. Стопанството мора да ги стави овие предизвици во агендата и да обезбеди тренинг за решавање на овие проблеми, на пример со ползување на принципот Приватност низ целиот процес.

55. Субјектите на заштита на приватноста и предизвиците поврзани со приватноста во контекст на употреба на Big Data треба да да се изучуваат на универзитетите за информатички науки.

56. Јавните органи треба да поседуваат соодветно ниво на знаење и свест во врска со потенцијалот на Big Data. Ова е особено важно кога се работи за пропишување на нови акти и регулативи. Свеста кон предизвиците е важна со цел јавните органи

да бидат во можност да ги осигурат нивните функции како заштитници на некои клучни општествени вредности.

Синџир на вредности на Big Data