Document sobre Bioètica i Big Data de salut: explotació i comercialització de les dades dels usuaris de la sanitat pública

Documento sobre Bioética y Big Data de salud: explotación y comercialización de los datos de los usuarios de la sanidad pública

Document on Bioethics and Big Data: exploitation and commercialisation of user data in public health care

M.R. Llàcer, M. Casado i L. Buisan (coords.)

Organització de les Nacions Unides per a l'Educació, la Ciència i la Cultura

Càtedra UNESCO de Bioètica de la Universitat de Barcelona

Observatori de Bioètica i Dret

www.bioeticaidret.cat
www.bioeticayderecho.ub.edu
www.bioethicsandlaw.es

Document sobre Bioètica i Big Data de salut: explotació i comercialització de les dades dels usuaris de la sanitat pública

Documento sobre Bioética y Big Data de salud: explotación y comercialización de los datos de los usuarios de la sanidad pública

Document on Bioethics and Big Data: exploitation and commercialisation of user data in public health care

M.R. Llàcer, M. Casado i L. Buisan (coords.)

Barcelona, gener de 2015

Organització de les Nacions Unides per a l'Educació, la Ciència i la Cultura

Càtedra UNESCO de Bioètica de la Universitat de Barcelona

Observatori de Bioètica i Dret UB

Universitat de Barcelona

Grup de Recerca Consolidat "Bioètica, Dret i Societat" de la Generalitat de Catalunya.

I. Llàcer, Maria Rosa. II. Casado, María. III. Buisan, Lidia.

1. Dades sanitàries. 2. Big Data. 3. Bioètica. 4. Privacitat.
5. Anonimització.

Observatori de Bioètica i Dret

Facultat de Dret, Universitat de Barcelona

Avda. Diagonal, 684, Mòdul 22

08034 Barcelona

Tel./Fax: (+34) 93 403 45 46

www.bioeticayderecho.ub.edu

www.bioeticaidret.cat

obd@ub.edu

DOCUMENT SOBRE BIOÈTICA I BIG DATA DE SALUT: EXPLOTACIÓ I COMERCIALITZACIÓ DE LES DADES DELS USUARIS DE LA SANITAT PÚBLICA

PRESENTACIÓ

El Grup d'Opinió de *l'Observatori de Bioètica i Dret,* de la Universitat de Barcelona, es va constituir el 1996 dins *l'Observatori de Bioètica i Dret,* que té, entre altres objectius, analitzar amb uns base científica i amb una metodologia interdisciplinària, les implicacions ètiques, socials i jurídiques de les noves tecnologies i els problemes biotecnològics i biomèdics, a fi d'intervenir en el diàleg entre la universitat i la societat mitjançant la transmissió del coneixement científic i tècnic, i aportant els arguments necessaris per a contribuir al debat social informat. Amb aquesta finalitat, el Grup d'Opinió ha elaborat ja vint-i-dos Documents[1] sobre temes d'actualitat i sobre els quals no hi ha una opinió unànime, ni en la societat ni en les diverses comunitats científiques implicades; això ha fet necessari identificar els problemes, contrastar els arguments i proposar recomanacions de consens.

En aquesta ocasió, el Grup fa públic el Document d'Opinió *Bioètica i Big Data de salut: explotació i comercialització de les dades dels usuaris de la sanitat pública,* a fi de cridar l'atenció sobre la necessitat de crear una cultura de la privacitat respecte de les dades personals, que han esdevingut elements o mecanismes de control en una societat informatitzada, per la qual cosa cal ser conscients de perquè i per a què aquestes dades han de ser protegides. Aquest Document analitza, des de la perspectiva bioètica, els problemes de l'explotació i la comercialització de dades dels usuaris de la sanitat pública. Prenent com a punt de partida el reconeixement del principi d'autonomia de les persones, el Document posa de manifest que la implementació de les tecnologies *big data* en l'àmbit sanitari, associada a una eventual comercialització d'aquestes dades, impacta directament en el nostre sistema sanitari i investigador —fonamentat en els principis d'igualtat i no discriminació— i afecta de ple a l'àmbit privat dels ciutadans.

El motiu immediat d'aquest Document han estat els problemes detectats en el projecte VISC+ (formalment anomenat *Més Valor a la Informació de Salut de Catalunya*) tant en relació a possibles vulneracions dels drets dels ciutadans com a la manca de transparència i de debat públic informat en una qüestió de tanta importància com és el tràfic de dades personals, reutilitzades amb finalitats diferents de l'atenció mèdica que puguin rebre directament els ciutadans. Els arguments que aquí oferim no solament fan referència a l'esmentat projecte sinó que tenen un abast major, perquè tenen a veure amb: a) la validesa de les tècniques d'anonimització en els conjunt de dades (*datasets*); b) la necessitat de redefinir el concepte de dades personals, tenint en compte la possibilitat actual de reidentificar les persones i c) l'impacte d'aquests dos aspectes en els mercats emergents de *big data, data marketplaces i digital marketing.*

[1] Tots els Documents del Grup d'Opinió de l'*Observatori de Bioètica i Dret* són accessibles en format PDF i en obert a: http://www.bioeticaidret.cat/documents (versió en català, espanyol i anglès). Consultat el 27 de gener de 2015.

Creiem que cal prendre mesures que garanteixin l'exercici dels drets i les decisions lliures i informades de totes les persones implicades. Pretenem obrir el debat sobre aquestes qüestions fent propostes que permetin afrontar el canvi de paradigma que impliquen aquestes noves tecnologies de la informació, car en una societat democràtica les decisions de l'Administració no han de ser imposades al ciutadà sense una informació prèvia, veraç i transparent sobre l'abast de les mateixes.

Aquest Document ha estat coordinat per les Dres. Maria Rosa Llàcer, María Casado i Lídia Buisan i ha estat elaborat pel *Grup d'Opinió de l'Observatori de Bioètica i Dret de la Universitat de Barcelona*, el *Grup de Recerca Consolidat "Bioètica, Dret i Societat" de la Generalitat de Catalunya*, amb la col·laboració del *Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya*. En la seva elaboració hi han participat, a més, les persones els noms i perfils professionals de les quals s'inclouen al final del document.

CONSIDERACIONS GENERALS

Els reptes del big data i l'anonimització

L'expressió *big data* és un terme que fa referència al tractament massiu de dades per mitjà d'algoritmes matemàtics a fi de generar correlacions entre elles, predir tendències i prendre decisions. Les tecnologies *big data* constitueixen un paradigma nou que, a més, implica canvis organitzatius importants tant en les empreses com en l'Administració. En l'actualitat, els objectius empresarials no són ja només la millora dels processos sinó la gestió de les dades. Estem assistint a una fase de transició cap a la *datificació* i la *monetització*, fet que comporta extreure un valor nou de les dades i rendibilitzar-les, tant en l'àmbit privat com en el públic o bé en una combinació de tots dos. Es tracta d'una tendència que s'insereix en el marc d'una indústria creixent basada en el coneixement adquirit mitjançant la reutilització de les dades i l'explotació d'aquestes, qüestió que cal tenir en compte a fi de contextualitzar el debat i entendre millor aquest canvi de model. Això no obstant, l'aposta per la innovació no ha de fer oblidar els aspectes ètics i els drets fonamentals de les persones, ni la protecció dels ciutadans en el context d'aquests nous avenços de les tecnologies. Es tracta de plantejar l'anàlisi d'aquesta situació amb la finalitat de proposar un nivell de protecció fort que suposi, per això mateix, un nivell més avançat d'innovació en aquest àmbit.

És molt important assenyalar que, fins ara, la premissa de *l'anonimització* de les dades ha estat la garantia que permetia respectar les regulacions de protecció de dades personals existents, en el benentès que en ser anonimitzada la dada personal passa a ser simplement una dada, perdent així la protecció de la normativa de protecció de dades personals, normativa que pretén ser rigorosa, tant a la Unió Europea com a l'Estat espanyol, però que amb els avenços de les tecnologies informàtiques després de gairebé vint anys ha esdevingut en bona part obsoleta. El problema rau en que, actualment, aquesta anonimització és palesament il·lusòria, perquè mitjançant tècniques d'enginyeria informàtica hom pot tornar a connectar les dades amb la persona font.[2] Tant la desanonimització de les dades com la reidentificació de persones són

[2] Art. 29 Working Party Opinion 06/2013 on open data and public sector information ('PSI') reuse (1021/00/EN WP207); Art.29 Working Party Opinion 05/2014, on Anonymisation Techniques (0829/14/EN WP216). Vegi's: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

NARAYANAN, Arvin; FELTEN, Edward W. "No silver bullet: De-identification still doesn't work", 2014. Vegi's: http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf

NARAYANAN, Arvin; SHMATIKOV, Vitaly. "Robust de-anonymization of large sparse datasets". *Security and Privacy*. IEEE Symposium on IEEE, 2008, pp. 111-125.

DE MONTJOYE, Yves-Alexandre, *et al.* "Unique in the Crowd: The privacy bounds of human mobility". *Scientific reports*, 2013, vol. 3.

OHM, Paul: "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization". *UCLA L. Rev.*

possibles si es disposa de la competència tècnica i dels mitjans necessaris per a fer-ho, per la qual cosa el debat es trasllada a un territori més tècnic i objectiu que proporciona informació i arguments que afecten directament a la cada vegada més estesa indústria de venda de dades. N'hi ha prou en saber que la reidentificació es pot fer tenint en compte el valor especial que poden adquirir determinades dades que fins ara s'ha considerat com no personals; per exemple, avui dia és prou evident que amb el codi postal, la data de naixement i el sexe ja és possible reidentificar la major part de les persones d'un *dataset*[3]. De manera semblant a com les nostres empremtes digitals ens identifiquen de manera unívoca, el mateix passa amb determinades tipologies de dades. La polèmica que hi ha darrere no és gens banal: què és una dada personal i com en podem garantir la protecció?[4] Com es pot evitar que a partir d'un conjunt de dades no personals es pugui identificar una persona?

Volem remarcar especialment aquest punt perquè el negoci de "posar en valor" les dades depèn precisament del concepte d'anonimització esmentat, ja que seria precisament aquest el que permetria complir amb les regulacions de protecció de dades personals. El debat sobre l'anonimització, tot i que ja té una certa història, no ha fet més que començar i és ben lluny d'estar acabat. En la nostra opinió, aquesta discussió és crucial en el segle XXI i no està tenint, ni de bon tros, la presència que li escauria en els diferents fòrums que hi tenen a veure (legals, ètics, tècnics, empresarials, governamentals) i en els quals caldria endegar el debat oportú perquè sigui compresa, primer, i resolta o si més no gestionada, després.

Com s'ha dit, actualment les evidències tècniques ja ens mostren que és possible reidentificar persones concretes a partir de les dades d'un *dataset* al qual s'hagin aplicat prèviament tècniques d'anonimització (o de desidentificació). Una persona, o bé una empresa, poden aconseguir la reidentificació esmentada si hi tenen interès (per motius econòmics, empresarials, delictius...), i tenen també els coneixements i els mitjans tecnològics per a fer-ho (per exemple, si disposen de les dades sanitàries d'un hospital —encara que no continguin dades personals— i d'accés a les dades personals d'un altre *dataset*, com ara un cens). Resulta evident que, en el cas de les dades de salut, no és difícil trobar un suposat "adversari" amb la motivació i els recursos per a fer-ho, i és escaient, per tant, qüestionar la validesa de les iniciatives de bescanvi de dades sensibles fonamentades en tècniques d'anonimització. En l'àmbit jurídic, l'incert recurs a "l'anonimització", entesa com una solució definitiva però inevitablement en crisi, es recolza en la normativa actual de protecció de dades, que prové d'una Directiva europea de l'any 1995 —per tant, molt anterior al fenomen del *big data*— i que es recull en la Llei 37/2007, de 16 de novembre, sobre la reutilització de la informació del sector públic. Però si el concepte mateix

2009-2010, p. 731.

[3] SWEENEY L. Simple demographics often identify people uniquely. Carnegie Mellon University, editor. Data Privacy Working Paper 3. 2000. http://dataprivacylab.org/projects/identifiability/paper1.pdf. Consultat el 27 gener 2015.

[4] SCHWARTZ, Paul M.; SOLOVE. D. "The PII Problem: Privacy and a New Concept of Personally Identifiable Information", *N.Y.U. L. Rev*. 1814 2011.

d'anonimització esdevé incert, cal trobar un fonament que legitimi l'anàlisi de dades personals de salut a gran escala. Si no és així, aleshores s'està obrint la porta a usos no desitjats d'aquestes dades, car en haver donat anteriorment el titular de les dades el seu consentiment per a determinades accions en l'àmbit sanitari i de recerca, en realitat en perd el control i queda desprotegit sense que sàpiga —perquè té una concepció equivocada de la protecció de dades i del secret professional— que les seves dades poden haver sigut utilitzades o cedides per a altres fins que no són ni desitjats ni efectivament consentits.

Aquest Document no pretén rebutjar, sense més, aquest nou model de negoci que centra l'atenció del mercat i en el qual, a més, ja estem immersos, sinó que vol alertar, tant als ciutadans com als poders públics que regulen i controlen l'activitat en l'àmbit sanitari i de recerca, dels riscos que comporta. En el nostre entorn no està prou consolidada una consciència social de la importància de protegir les dades en relació amb el dret fonamental a la intimitat i a la no discriminació. No tenim una cultura de la privacitat que ens permeti comprendre de quina manera ens pot afectar que una empresa acumuli i faci rendible la nostra informació, i que disposi d'un instrument de poder en base al qual pugui prendre decisions que ens afectin.[5] N'és un exemple el fet que l'anàlisi massiva de dades es pot fer servir per a descobrir efectes secundaris de medicaments, però també per a generar perfils de risc —i que els propis afectats poden no conèixer— que es podrien utilitzar per a "justificar" la denegació d'una assegurança.

Es fa evident, en conseqüència, la urgència d'un debat que posi de relleu la vulnerabilitat de les persones davant el risc de discriminació generat per perfils i patrons de conducta generats amb finalitats que la persona afectada no pot controlar, i també sobre l'adaptació de les lleis als reptes ètics i socials que les *big data* plantegen. Aquest és, precisament, el nucli del conflicte que cal obrir al debat públic amb implicació de la ciutadania a fi de crear una cultura de la privacitat que vagi d'acord amb l'actualitat i amb les noves realitats[6].

Sobre el Projecte VISC+

Un exemple de reutilització de dades que, des del nostre punt de vista, resulta molt qüestionable és el projecte VISC+, impulsat per l'Agència de Qualitat i Avaluació Sanitàries de Catalunya (AQuAS), que té com a objectiu —segons esmenten els seus promotors— posar la informació sanitària a disposició dels ciutadans, les empreses i la recerca per a millorar els serveis de salut i la investigació, i per a "posar en valor" el coneixement.

L'esmentat projecte es nodreix de les diferents bases de dades que ja existeixen en el sistema sanitari: el Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària

[5] COHEN, Julie: "What Privacy is for" *Harv. L. Rev.* 1904, 2012-2013.

[6] RICHARDS NM, King JH. "Big data ethics". *Wake Forest Law Review* 2014;(49):393-432.

(SIDIAB) i, de manera especial, la Història Clínica Compartida a Catalunya (HC3), que recull les dades assistencials i de consum farmacèutic, juntament amb altres informacions rellevants com són la identificació i la situació sociosanitària de cada ciutadà atès per la sanitat pública; a més, la HC3 conté informació de les proves analítiques i diagnòstiques que inclouen paràmetres metabòlics i bioquímics, així com dades de diagnòstic genètic que identifiquen les persones portadores de malalties genètiques hereditàries o bé que mostren riscos i susceptibilitats de patir malalties més complexes. Aquestes bases de dades fa que existeixin "fitxers d'usuaris" dels quals n'és responsable el Departament de Salut de la Generalitat de Catalunya. Tot i que els macro fitxers de dades estan protegits per la normativa ja existent —en especial per la Llei Orgànica de Protecció de Dades (LOPD) i el Reglament que la desenvolupa— aquesta regulació ha esdevingut del tot inadequada en el marc de la nova realitat dels *big data*, com s'ha esmentat abans, i no garanteix de cap manera que no es produeixin usos indeguts i discriminatoris.[7]

La HC3 té els objectius explícits següents: a) millorar l'atenció de la salut dels ciutadans mitjançant una eina que faciliti la feina dels professionals sanitaris respecte als malalts als quals han d'atendre; b) propiciar un nou model assistencial en permetre l'accés i la consulta de forma immediata, segura i confidencial, de la informació rellevant disponible sobre els usuaris. Com és obvi, les dades que conté la HC3 són extremadament sensibles, la recollida i el tractament de les quals es justifica en la seva eficàcia per a proporcionar una assistència de qualitat, no sols en el centre que habitualment atén l'usuari sinó en tota la xarxa assistencial pública de Catalunya[8], perquè la HC3 permet l'accés de manera organitzada, i atenent a criteris de seguretat i confidencialitat, a les històries clíniques de la xarxa assistencial. Aquesta eina ha d'oferir beneficis tant a la ciutadania, com als professionals sanitaris, com al propi sistema de salut. Per aquesta raó, el ciutadà té dret a saber qui pot accedir a les seves dades personals i amb quina finalitat; i també té dret a exigir responsabilitats si creu que se n'està fent un ús indegut o distint d'aquell al qual en el seu moment va consentir. Quan es va dur a la pràctica la HC3, ni es va informar suficientment els ciutadans d'aquesta recollida massiva de dades amb finalitat assistencial, ni es va indicar en cap moment que aquestes mateixes dades podrien ser reutilitzades amb altres finalitats, fins i tot comercials. De cap manera es pot considerar que la cessió de les dades per a finalitats no assistencials sigui el "preu" de la gratuïtat de l'assistència sanitària, ja que si s'exigís algun tipus de contraprestació l'assistència deixaria de ser gratuïta.

La informació que conté la HC3, tot i ser recollida i estructurada pels professionals assistencials, fa referència a les dades de salut del malalt i, per tant, aquestes li pertanyen, per la

---

[7] Com el Grup d'Opinió ja va advertir en el *Document* s*obre proves genètiques de filiació*, Barcelona: Signo, 2006. Disponible en format PDF a: http://www.publicacions.ub.edu/refs/observatoriBioEticaDret/documents/07899.pdf

[8] Existeix un projecte d'un abast territorial major, desenvolupat en 12 països de la Unió Europea, en el marc de www.epSOS.eu i del que formen part Alemanya, Àustria, Txèquia, Dinamarca, Eslovàquia, Espanya, França, Grècia, Holanda, Itàlia, Regne Unit i Suïssa. De l'Estat espanyol hi participen tres CCAA: Andalusia, Castilla-La Manxa i Catalunya, dins del PLAN AVANZA per a la modernització dels serveis de les Administracions públiques.

qual cosa les entitats assistencials reben peticions relacionades amb l'exercici dels drets que la normativa sobre protecció de dades vigent reconeix a la ciutadania: drets d'accés, de rectificació, de cancel·lació i d'oposició (els anomenats drets ARCO). La finalitat d'aquest conjunt de drets és impedir un tractament il·lícit i lesiu per a la dignitat i el dret de l'afectat (*habeas data*), alhora que per a garantir l'exercici del dret més general a la intimitat. Els usos de les dades recollides en la HC3 s'han de limitar a l'assistència (juntament amb les finalitats científiques —en epidemiologia, investigació i docència, o bé dirigides a la millora dels serveis públics— que la normativa actual ja autoritza) i és absolutament necessari establir garanties reals que evitin el tràfic de dades i qualsevol ús indegut per part d'empreses de l'àmbit de la salut (assegurances mèdiques, corporacions farmacèutiques, entitats financeres, i altres). Per això, el projecte VISC +, tal com en aquests moments està previst que es dugui a terme, genera dubtes importants, tant de caràcter bioètic com estrictament jurídic, que convé analitzar amb detall i debatre, a fi de prevenir-ne possibles usos discriminatoris.

Problemes rellevants del projecte VISC+

1. Denominació equívoca del projecte

La mateixa denominació del projecte és equívoca i no s'adequa al principi general de lleialtat en la recollida i tractament de dades, perquè indueix a pensar que el projecte ajuda a millorar les condicions de vida i la salut dels ciutadans. Els usuaris a qui es demani el consentiment perquè les seves dades personals de salut siguin tractades en el marc d'un projecte anomenat VISC+ poden pensar, erròniament, que col·laboren en un programa que l'única cosa que pot aportar-los són beneficis. Això pot ser fàcilment relacionat amb una pràctica deslleial, entesa com una conducta contrària a la bona fe objectiva que distorsiona la capacitat d'escollir amb ple coneixement de causa i que indueix a facilitar unes dades que altrament no s'haurien proporcionat. La lleialtat és un valor fonamental en el marc de la LOPD, ja que la recollida i el tractament de dades per mitjans fraudulents o deslleials està prohibida expressament i afecta de manera directa el principi de qualitat de les dades.

2. Limitació de les finalitats en el tractament de les dades

Com bé es diu en l'informe elaborat pel *Grup europeu d'ètica de les ciències i de les noves tecnologies* (GEE), de la Comissió Europea, en la recollida i el tractament de dades de caràcter personal les entitats públiques i les privades han de fonamentar la seva activitat en el principi de "limitació de la finalitat"[9]; és a dir, que aquest tipus de dades no haurien de ser recollides ni tractades per a qualsevol ús, sinó només amb objectius específics i legítims. Cal, a més, que les dades no estiguin

[9] EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES: Ethics of Security and Surveillance Technologies. Opinion n. 28, 20 de maig de 2014. Vegi's: http://ec.europa.eu/archives/bepa/european-group-ethics/docs/publications/ege_opinion_28_ethics_security_surveillance_technologies.pdf

per defecte a disposició de "qui les vulgui utilitzar" i que els ciutadans tinguin mecanismes efectius per a controlar i modificar les informacions que els concerneixin i que estiguin dipositades en les esmentades entitats. Insisteix també l'informe en que la possible cessió de dades amb finalitats comercials ha de fer-se només amb el consentiment explícit de les persones afectades, i que les entitats privades han d'indicar el tipus de dades que preveuen tractar i amb quin objectiu, durant quant de temps i si tenen intenció de relacionar o connectar aquestes dades amb altres procedents de diferents fonts.

Resulta especialment important, en el marc del projecte que aquí s'analitza, la gradualitat de la protecció de les dades en funció de la finalitat de l'ús, distingint acuradament les finalitats sanitària, epidemiològica i d'investigació i docència —que ja estan emparades per la legislació— de les finalitats privades, a les quals cal exigir el nivell de protecció més elevat. Ara bé, el projecte VISC+ equipara tractaments que tenen finalitats totalment diferents i aquesta confusió afecta la legitimació per tractar dades sanitàries personals, que són dades especialment sensibles i que, en conseqüència, estan sotmeses a una protecció especial.

Com ja s'ha dit, la legislació empara l'ús de dades dels usuaris de salut per a dur a terme l'assistència sanitària i per a la recerca i la millora dels serveis públics. Però si es vol anar més enllà i facilitar la utilització d'aquestes dades amb finalitats ni previstes ni autoritzades —com són els interessos comercials d'empreses privades els productes de les quals depenen de la recerca, a més d'altres factors— cal un debat social previ sobre la concurrència dels interessos públics i privats en l'àmbit de la recerca a fi definir els límits ètics i el nivell de protecció de què disposarà el ciutadà quan empreses amb interessos privats facin el tractament de dades de salut. L'*empowerment* del ciutadà es basteix amb informació adequada, clara i veraç sobre l'ús de les seves dades, a més de reconèixer-li la facultat de controlar el tractament de les mateixes, ja sigui consentint-hi o bé oposant-s'hi.

3. Habilitació legal *vs* consentiment

Comptar amb legitimació suficient és el requisit bàsic per a permetre l'accés de terceres persones a informacions o dades que pertanyen a l'àmbit personal dels afectats. Cal diferenciar entre la legitimació legal i la voluntària, basada aquesta segona en el consentiment lliurement atorgat. La primera fa referència a finalitats relacionades amb l'atenció sanitària, la qualitat i gestió del servei i a finalitats científiques (epidemiològiques, de recerca i docència). Aquesta legitimació, ja reconeguda legalment, es justifica en l'interès públic, respectant sempre escrupolosament la confidencialitat de les dades així recollides i obligant a motivar la sol·licitud d'ús de les mateixes. Pensem que el projecte VISC+, en haver estat aprovat per un mer "Acord de Govern", no té l'habilitació legal suficient per a la reutilització de dades sanitàries, perquè la legislació sanitària només legitima per a tractar dades d'usuaris amb finalitats directament assistencials, de recerca o bé administratives. El segon tipus de legitimació, la voluntària, prové sempre del consentiment explícit del pacient. Aquest és el cas del tractament de dades amb finalitats estrictament privades,

és a dir, sense interès públic evident i dirigides al desenvolupament de les indústries sanitàries, farmacèutiques i de biotecnologia, o per a la promoció i comercialització dels productes que generin. Menció a part mereixen les dades genètiques a causa de la complexitat que suposa la titularitat de les mateixes quan es compartida per un nucli familiar.

Tenint en compte l'existència del binomi públic-privat en el sistema sanitari i de recerca, el problema cal centrar-lo en com convé articular la legitimació abans esmentada a fi de poder emprar la informació de salut i reutilitzar-la. Segons el nostre parer, cal que això quedi establert en una llei, tenint en compte, a més, l'enorme asimetria existent —d'informació, però també de poder— entre el ciutadà, que pateix una malaltia i que el que vol és curar-se, i el professional que li demana el consentiment, tant per a l'assistència mèdica més escaient en el seu cas com per al tractament de les dades personals de salut en general. Però ha d'estar clar que es tracta de dos consentiments diferents, i que la prestació sanitària pública no es pot condicionar al consentiment per a tractar dades amb altres finalitats, ni pot justificar la sol·licitud de dades addicionals. La conclusió que cal extreure del que s'acaba de dir és que l'obtenció del consentiment s'ha de sotmetre a pautes rigorosíssimes a fi de compensar la situació de desequilibri en què es troba l'usuari dels serveis sanitaris públics, i més en un moment en què sol estar especialment preocupat per la seva salut, la qual cosa li pot generar una situació de vulnerabilitat que el dugui a pensar *a priori* que totes les dades que se li demanen són imprescindibles per al seu tractament i per a rebre l'assistència que necessita i que és la raó per la qual ha acudit al sistema sanitari.

4. Valor i risc

És en aquest context que el projecte VISC+ es presenta amb la finalitat de "posar en valor" l'enorme quantitat de dades de què disposa el Departament de Salut, reutilitzant aquestes dades per a finalitats no previstes inicialment i que, per tant, l'usuari no coneix. L'expressió "posar en valor" té interpretacions diferents: d'una banda, es tractaria de posar a disposició de centres de recerca, centres d'estudis epidemiològics i de salut pública, i centres de docència que així ho sol·licitin, dades de salut dels ciutadans amb l'objectiu de contribuir al progrés del coneixement, en els àmbits específics esmentats, i també, en darrer terme, a la millora de l'atenció sanitària, i la prevenció, qüestions perfectament legítimes i que, insistim, la legislació actual ja permet. Però, d'altra banda, també se'n fa una interpretació molt més laxa pel fet de posar aquestes *big data* sanitàries a disposició d'empreses que difícilment entrarien en els àmbits que s'acaben d'esmentar. Si tenim en compte el principi econòmic ben conegut segons el qual *big data is big business,* l'escenari més probable seria el d'una pura i simple venda de les dades de salut dels ciutadans en benefici de l'empresa que estigui interessada en fer rendible aquesta informació i que disposi dels mitjans per a fer-ho. Convé remarcar que el problema no és la tecnologia, sinó el sentit i la direcció que li doni qui la utilitzi i qui la financi. El marc que proposa el projecte VISC+ permetria a les empreses interessades extreure un valor merament comercial de les dades.

Els riscos potencials a què fa referència aquest Document no són ni hipotètics ni remots: n'hi ha prou amb analitzar la possibilitat de construir perfils de conducta en base a dades anònimes que en qualsevol moment es poden utilitzar per a prendre decisions automatitzades que afectin persones. N'hi ha prou de fer una passejada breu per internet per veure una bona quantitat d'empreses dedicades a la compravenda de dades i com les que ja les posseeixen —perquè són dades que ja s'han generat per a altres fins i serveis— creen, al seu torn, altres empreses i línies de negoci dedicades a la reutilització d'aquestes dades, aconseguint perfils molt precisos mitjançant encreuaments successius d'informació i altres processos d' enriquiment de dades.

5. Control de les dades

Precisament per a evitar la pèrdua de control sobre les dades i per la possibilitat d' abús, té sentit establir funcions de *Data Governance* (és a dir, el control del tractament i la gestió de les dades) que han de correspondre a les entitats públiques, garants del respecte escrupolós dels drets fonamentals dels ciutadans, amb independència de com se subministrin els serveis professionals per part de les empreses adjudicatàries. Aquestes funcions de governança han d'incloure aspectes com: la seguretat i qualitat de les dades, la privacitat, els processos d'anonimització, la traçabilitat, les polítiques de permanència de les dades i l'enriquiment de dades, posant limitacions a les fonts o bases de dades amb què es poden relacionar o connectar. Aquesta preferència per les garanties que ofereix l'àmbit públic prové de les obligacions específiques de transparència i de rendició de comptes que té l'Administració, encara que el criteri de responsabilitat compartida exigiria aquesta mateixa transparència i rendició de comptes també al sector privat. La *Declaració Universal sobre Bioètica i Drets Humans* de la UNESCO (2005) estableix, en l'art. 14, el principi innovador de responsabilitat social en salut, principi que escau perfectament en el context que aquí s'analitza i que al mateix temps alerta sobre la necessitat d'evitar els conflictes d'interessos en un àmbit tan delicat com aquest.

El projecte VISC+ parla de governança, però merament de caràcter intern; en aquest Document, en canvi, es proposa un control extern i representatiu de la societat. Cal remarcar que l'informe de l'Autoritat Catalana de Protecció de Dades (ACPD) ja assenyala que el projecte VISC+ mereix un règim específic de seguretat encara més estricte. Segons aquest informe, el departament de Salut és a efectes legals el "responsable" del tractament de les dades, mentre que "l'entitat" (l'Agència de Qualitat i Avaluació Sanitàries de Catalunya -AQuAS) és "l'encarregada" del tractament de les dades. També la mecànica prevista en el projecte VISC+ per a l'explotació de les dades és motiu de preocupació, perquè l'empresa adjudicatària rebria les dades suposadament anonimitzades a canvi d'un preu o taxa, tot i que també s'hi diu que l'adjudicatari participaria en el procés de verificació de l'anonimització i en la materialització d'un "codi anònim de la persona" abans de transferir les dades als usuaris finals. No es diu enlloc si seria aquesta mateixa empresa adjudicatària qui es responsabilitzaria de "definir, construir i posar en marxa un catàleg de serveis útil, eficient, competitiu i innovador, i contrastar les necessitats del

mercat i els clients finals del projecte, així com de definir un pla de difusió i de comercialització, canalitzant de manera adequada la demanda del mercat nacional i internacional".

En el model VISC+ no està gens clar si serien les empreses adjudicatàries qui decidirien a qui traspassen les dades de salut, presumptament a canvi de contrapartides econòmiques que tampoc no es precisen enlloc ni s'esmenta com retornarien als ciutadans. Per exemple, se cedirien bases de dades de malalts de l'hepatitis C per a desenvolupar fàrmacs que després es voldrien vendre a 60.000 € cada tractament? Precisament aquest punt seria clau, perquè en dependria de fins quin punt els clients o usuaris finals del projecte estarien disposats a contribuir-hi en funció de les expectatives de negoci que prevegin. L'ús abusiu i opac de les dades personals relacionats amb la salut genera desconfiança en la població que impedeix que aquestes dades s'utilitzin de manera legítima per a fins epidemiològics, d'investigació o docència.

6. Avaluació de l'impacte

El projecte VISC+ no inclou cap avaluació de l'impacte que, un cop engegat, pugui tenir sobre el dret a la intimitat dels usuaris, qüestió cabdal tenint en compte que estem parlant de dades tan sensibles com les de salut, en la línia del que proposa el projecte de Reglament europeu de protecció de dades. Aquesta valoració de l'impacte —també des de les perspectives ètica i social— s'ha de considerar un requisit de qualsevol llei que empari projectes com el VISC+. Ni tampoc s'especifica enlloc si els usuaris han de donar prèviament el seu consentiment per al transvasament de dades que suposa el projecte, o si s'entén que aquest consentiment no seria necessari pel fet de la interpretació laxa del concepte de recerca a què s'ha fet referència més amunt i que no diferencia entre l'interès públic i l'interès privat.

Una darrera qüestió extremadament important és que en cap moment s'explica clarament de quina manera el benefici econòmic que s'obtingués del projecte repercutiria favorablement als ciutadans i al sistema sanitari públic. Amb dades tan valuoses, i quan hi ha expectatives de beneficis i de negoci, sembla raonable que els ciutadans en rebin contrapartides clares. Si bé hom pot considerar que el lucre és un objectiu lícit, no és cert que sigui el bé primordial al qual tots els altres valors i drets hagin de sotmetre's. El Conveni sobre Drets Humans i Biomedicina, del Consell d'Europa, vigent des de l'any 2000, estableix —en l'article 2— que els interessos de la ciència o de la societat mai no han de prevaldre sobre els de les persones, i en aquesta premissa es basa tot el sistema de ciència i tecnologia, especialment el sistema sanitari i de recerca.

7. L'experiència del *National Health Service* britànic

A banda del que ja s'ha dit, val la pena remarcar que les experiències en la línia del projecte VISC+ que s'han dut a terme en països del nostre entorn han generat conflictes molt significatius que porten a extremar les mesures de prudència abans d'endegar projectes d'aquest tipus. En efecte, la suposada anonimització de dades de salut i d'atenció sanitària recol·lectades —des de 2005 fins a 2013— pel *NHS Information Centre* (NHS-IC) anglès no ha impedit que diferents empreses hagin pogut identificar les persones a qui feien referència aquestes dades, causant-los

diferents perjudicis, per exemple l'increment del preu de les primes de risc de les assegurances. La conseqüència ha estat una moratòria, a fi de reorganitzar el procediment de cessió de les dades de manera que sigui més transparent, que es garanteixi de forma més eficaç el dret a la intimitat i a la confidencialitat dels ciutadans. En l'actualitat, tot el procediment de tractament i cessió de les dades està sotmès en tot moment a auditoria i control públic; a més, s'ha creat el *Care Data Advisory Group*, grup consultor la missió del qual és enfortir la protecció dels drets dels ciutadans en l'àmbit sanitari, i el *National Data Guardian*, a fi de vetllar per la seguretat de les dades de salut.

RECOMANACIONS

1. *Generar i potenciar una cultura ciutadana de la privacitat en matèria de dades personals*. Acumular informació sobre una persona equival a adquirir poder de decisió sobre ella; en conseqüència, els mitjans per a controlar qui tracta les nostres dades, com les recapta i amb quina finalitat les utilitza es converteixen en eines tant de la llibertat personal com de la col·lectiva.
2. *Informar i formar la ciutadania sobre l'abast real del fet que els processos d'anonimització de les dades ja no garanteixen per ells mateixos la irreversibilitat*. Actualment són possibles la desanonimització, la reidentificació o la revelació de dades personals de conjunts d'usuaris o d'usuaris individuals, car les eines informàtiques emprades poden servir tant per a aquesta finalitat com per a la contrària.
3. *Alertar sobre la necessitat de redefinir el concepte mateix de "dades personals" en què es fonamenta la legislació actual.* És important tenir present que el problema no és únicament la transformació de les dades que es consideren personals en un conjunt de dades, perquè fins i tot eliminant-ne aquestes dades personals hom pot reidentificar persones concretes.
4. *Aplicar escrupolosament el principi que exigeix que les dades que es recol·lectin siguin adequades a la finalitat que motiva la seva recollida, exigint que se sol·liciti i que s'obtingui el consentiment explícit dels usuaris per a la utilització de les seves dades de salut amb finalitats distintes d'aquelles per a les quals es varen obtenir en el seu moment, implementant mecanismes eficaços per a atorgar o, si és el cas, per a denegar el dit consentiment.* En el cas de les dades genètiques, el consentiment ha de ser especialment exigent a causa de la possible afectació d'altres membres del nucli familiar.
5. *Establir excepcions a l'exigència del consentiment informat en estudis epidemiològics, com ara*: estudis de vigilància epidemiològica promoguts per les autoritats; estudis de registres sanitaris creats o reconeguts oficialment per les autoritats sanitàries; estudis en què les dades estiguin adequadament anonimitzades sense que hi hagi cap possibilitat de reidentificació de les persones; estudis de salut pública de tipus observacional, amb riscos mínims, i per als quals l'obtenció del consentiment informat sigui impossible o bé extremadament difícil, sempre que disposin del vist-i-plau d'un comitè d'ètica en investigació.
6. *Potenciar un procés d'informació i de debat entre la ciutadania —basat en una valoració honesta dels riscos i beneficis, dels avantatges i perjudicis— abans de posar en marxa un projecte com el VISC+,* dirigit a l'explotació de dades de salut que pertanyen als ciutadans, encara que estiguin en mans de l'Administració, a fi que aquests puguin expressar el seu parer sobre la conveniència de dur a la pràctica o no un projecte com aquest.
7. *Establir mecanismes de control i concretar les funcions de Data Governance, que són responsabilitat dels organismes públics*. Els processos d'anonimització de les dades s'han de dur a

terme dins del perímetre intern de l'Administració, en raó de les garanties que ofereixen les obligacions de transparència i de rendició de comptes que té l'Administració, tenint present, però, que la responsabilitat compartida exigeix també transparència i rendició de comptes al sector privat, en especial en el cas de xarxes mixtes público-privades.

8. *Controlar de manera específica la seguretat en el tractament de les dades sanitàries a fi de garantir en tot moment el seu ús correcte i evitar-ne la comercialització que no compti amb el consentiment explícit dels usuaris i que no prevegi de forma clara com el benefici que se'n pugui obtenir revertirà als ciutadans.*

   A tal fi, hi ha d'haver un òrgan de governança extern, independent dels agents implicats i representatiu de la societat catalana. Les seves funcions han de ser supervisar la gestió dels fitxers al llarg de totes les etapes del procés, verificar-ne la traçabilitat, detectar possibles conflictes d'interessos i males pràctiques, així com identificar possibles infraccions i responsabilitats, i haurà d'informar públicament de la seva actuació i resultats. D'aquesta manera, el ciutadà podria dirigir-se a aquest òrgan per a saber qui disposa de les seves dades personals i amb quina finalitat.

   Per a dur a terme aquestes funcions, l'esmentat òrgan ha de promoure i aplicar un *Codi Ètic per a la Reutilització de Dades de Salut*.

9. *Enfortir la formació en deontologia i ètica professional dels implicats en la custòdia de les dades*, insistint en que qualsevol persona que tingui accés o tracti les dades té el deure de protegir i promoure els drets i llibertats fonamentals dels ciutadans afectats. El dret a la intimitat, la confidencialitat i la no discriminació són pilars bàsics del sistema de salut i de recerca.

10. *Recordar que els diferents Comitès d'Ètica han de contribuir al desenvolupament de la cultura de respecte de la intimitat i la confidencialitat de les dades personals, perquè en el àmbit d'actuació que els és propi són la primera línia de defensa d'aquests drets.* Els membres d'aquests Comitès han de tenir formació específica en els aspectes ètics, tècnics, jurídics i socials de les noves tecnologies a fi de poder col·laborar en la presa de decisions ponderades i proporcionals —tenint sempre present que les dades sanitàries són dades sensibles que requereixen una protecció especial— i posant atenció, a més, a no actuar de manera que puguin esdevenir mers mecanismes de cobertura d'interessos aliens.

NORMATIVA DE REFERÈNCIA

Internacional

- Consell d'Europa: Conveni per a la protecció dels Drets Humans i la dignitat del ser humà pel que fa a les aplicacions de la biologia i la medicina (Conveni sobre Drets Humans i biomedicina), fet a Oviedo el 4 d'abril de 1997.
- UNESCO: Declaració Universal sobre Bioètica i Drets Humans, de 19 de octubre de 2005.

Europea

- Carta dels Drets Fonamentals de la Unió Europea (DOUE núm. 83, de 30 de març de 2010).
- Directiva 1995/46/CE, de 24 d'octubre, del Parlament Europeu i del Consell, relativa a la protecció de les persones físiques relativa al tractament de les dades personals i a la lliure circulació d'aquestes dades.
- Directiva 2002/58/CE del Parlament Europeu i del Consell, de 12 de juliol de 2002, relativa al tractament de les dades personals i a la protecció de la intimitat en el sector de les comunicacions electròniques (Directiva sobre la privacitat i les comunicacions electròniques).

Espanyola

Sistema sanitari

- Ley 14/1986, de 25 de abril, General de Sanidad.
- Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.
- Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud.
- Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud.
- Ley 33/2011, de 4 de octubre, General de Salud Pública.
- Llei 21/2000 (del Parlament de Catalunya), de 29 de desembre, sobre els drets d' informació concernent la salut i l'autonomia del pacient, i la documentació clínica.

- Llei 16/2010 (del Parlament de Catalunya), de 3 de juny, de modificació de la Llei 21/2000, de 29 de desembre, sobre els drets d'informació concernent la salut i l'autonomia del pacient, i la documentació clínica.

Sistema d'investigació

- Real Decreto 223/2004, de 6 de febrero, por el que se regulan los ensayos clínicos con medicamentos.
- Ley 29/2006, de 26 de julio, de garantías y uso racional de los medicamentos y productos sanitarios.
- Ley 14/2007, de 3 de julio, de Investigación Biomédica.
- Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.

Protecció de dades de caràcter personal

- Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.
- Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Reutilització de la informació del sector públic

- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.
- Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, sobre reutilización de la información del sector público.

DOCUMENTO SOBRE BIOÉTICA Y BIG DATA DE SALUD: EXPLOTACIÓN Y COMERCIALIZACIÓN DE LOS DATOS DE LOS USUARIOS DE LA SANIDAD PÚBLICA

PRESENTACIÓN

El *Grupo de Opinión del Observatori de Bioètica i Dret de la Universitat de Barcelona* surgió en 1996 en el seno del *Observatori de Bioètica i Dret*. Éste tiene entre sus objetivos analizar científica e interdisciplinarmente las implicaciones éticas, sociales y jurídicas de las nuevas tecnologías y los problemas biotecnológicos y biomédicos e incidir en el dialogo entre la universidad y la sociedad, mediante la transmisión del conocimiento científico-técnico y los argumentos necesarios para participar en un debate social verdaderamente informado. Con este fin, el *Grupo de Opinión* ha elaborado ya veintidós Documentos[10] sobre temas de actualidad sobre los que no existe una opinión unánime, ni en la sociedad ni en las diversas comunidades científicas implicadas; ello ha requerido identificar los problemas, contrastar los argumentos y proponer recomendaciones de consenso.

Ahora, el *Grupo* presenta el *Documento de Opinión* sobre *"Bioética y Big Data de salud: explotación y comercialización de los datos de los usuarios de la sanidad pública",* que pretende alertar sobre la necesidad de crear una cultura de la privacidad en materia de datos personales ya que estos se han convertido en elementos o dispositivos de control en una sociedad informatizada, y es preciso ser conscientes de por qué y para qué deben protegerse. Analizamos, desde una perspectiva bioética, los problemas de la explotación y comercialización de los datos de los usuarios de la sanidad pública. Partiendo del reconocimiento del principio de autonomía de las personas, el Documento se encamina a poner de manifiesto que la implementación de las tecnologías Big Data en salud, asociada a una eventual comercialización de dichos datos, produce un impacto en nuestro sistema sanitario e investigador —asentado en los principios de igualdad y no discriminación— y afecta directamente a la esfera privada de los ciudadanos.

El detonante inmediato de Documento han sido los problemas detectados en el proyecto VISC+ —*Más Valor a la Información de Salud en Cataluña*— y que son de dos tipos: 1) las posibles vulneraciones de los derechos de los ciudadanos y 2) la falta de transparencia y debate público informado en una cuestión en la que subyace el tráfico de datos personales reutilizados con fines distintos al tratamiento médico directamente recibido por el paciente. Así, planteamos argumentos que son relativos al mencionado proyecto, pero que tienen un alcance mayor: a) la validez de las técnicas de anonimización sobre conjuntos de datos (*datasets*), b) la necesidad de redefinir el concepto de datos personales, dada la actual capacidad de re-identificación de personas y c) el impacto de las dos cuestiones anteriores en los mercados emergentes de *big data, data marketplaces y digital marketing*.

[10] Todos los Documentos del Grupo de Opinión del *Observatori de Bioètica i Dret* son accesibles en formato PDF y en abierto en: http://www.bioeticayderecho.ub.edu/documentos (versión en catalán, español e inglés). Consultado el 27 de enero de 2015.

Consideramos que es preciso tomar medidas que permitan asegurar el ejercicio de los derechos y la toma de decisiones libres e informadas de todas las personas implicadas. Pretendemos abrir el debate y formular propuestas para afrontar el cambio de paradigma que implican dichas nuevas tecnologías de la información ya que, en una sociedad democrática, las medidas de la Administración no deben ser impuestas al ciudadano sin una información previa, veraz y transparente.

Este nuevo Documento ha sido coordinado por las Dras. Maria Rosa Llàcer, María Casado y Lidia Buisan, y ha sido elaborado por el *Grupo de Opinión del Observatori de Bioètica i Dret de la Universitat de Barcelona*, el *Grup de Recerca Consolidat "Bioética, Dret i Societat" de la Generalitat de Catalunya*, con la colaboración del *Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya*. Han participado en su preparación todas las personas cuyos nombres y perfiles profesionales se incluyen al final del mismo.

CONSIDERACIONES GENERALES

Los retos del big data y la anonimización

La expresión *big data* es un término que designa el tratamiento de grandes volúmenes de datos mediante algoritmos matemáticos con el fin de establecer correlaciones entre ellos, predecir tendencias y tomar decisiones. Las tecnologías *big data* constituyen un nuevo paradigma e implican cambios organizativos en las empresas y en la propia Administración. Actualmente, las empresas ya no se organizan tanto a través de la mejora de los procesos como en torno a la gestión del dato. Asistimos a una transición hacia la *datificación* y la *monetización* que comporta extraer un nuevo valor de los datos y rentabilizarlos, tanto desde el interés privado como el público, o una combinación de ambos. Es una tendencia inserta en una floreciente industria basada en el conocimiento adquirido a través de la reutilización de los datos y su explotación, que conviene tener en cuenta para contextualizar el debate y entender este cambio de modelo. No obstante, la apuesta por la innovación no puede olvidar los aspectos éticos y los derechos fundamentales de las personas, ni la protección de los ciudadanos en el contexto de estos nuevos avances de las tecnologías. Se trata de examinar esta situación para proponer un nivel de protección firme que suponga en sí mismo una mejora y, por ello, un nivel más avanzado de innovación en este ámbito.

Es crucial señalar que, hasta ahora, la premisa de la *anonimización* del dato ha representado la garantía que permitía cumplir con las regulaciones de protección de datos personales existentes. Se ha venido entendiendo que un conjunto de datos personales, al ser anonimizados, dejan de contener datos de carácter personal, perdiendo así el amparo de la normativa de protección de datos personales, que se pretende rigurosa tanto en la UE como en España pero que, con el avance de las tecnologías informáticas, tras casi veinte años ha devenido en buena medida obsoleta. El problema radica en que, actualmente, está acreditado que la anonimización no garantiza la privacidad de los datos personales, puesto que mediante técnicas de ingeniería informática es posible volver a conectar los datos con la persona a quien pertenecen[11]. La des-

[11] Art. 29 Working Party Opinion 06/2013 on open data and public sector information ('PSI') reuse (1021/00/EN WP207); Art.29 Working Party Opinion 05/2014, on Anonymisation Techniques (0829/14/EN WP216). Véase: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

NARAYANAN, Arvin; FELTEN, Edward W. "No silver bullet: De-identification still doesn't work", 2014. Véase: http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf

NARAYANAN, Arvin; SHMATIKOV, Vitaly. "Robust de-anonymization of large sparse datasets". *Security and Privacy*. IEEE Symposium on IEEE, 2008, pp. 111-125.

DE MONTJOYE, Yves-Alexandre, *et al*. "Unique in the Crowd: The privacy bounds of human mobility". *Scientific reports*, 2013, vol. 3.

anonimización y la re-identificación subsiguiente, resultan posibles cuando se dispone de la competencia técnica y los medios necesarios; así, el debate se apoya en un campo más técnico, del que se obtienen informaciones y argumentos que afectan directamente a la base sobre la que se apoya la industria de la venta de datos. Baste saber que la re-identificación puede hacerse por los valores particulares que pueden tomar ciertos datos, hasta ahora considerados como no personales; por ejemplo, se ha demostrado que con un código postal, la fecha de nacimiento y el sexo, es posible re-identificar a la gran mayoría de personas de un *dataset*[12]. De la misma manera que nuestras huellas dactilares nos identifican unívocamente, también ocurre lo mismo con ciertas tipologías de datos. La polémica que subyace es profunda: ¿qué es un dato personal y cómo podemos garantizar su protección?[13], ¿cómo podemos evitar que un conjunto de datos no personales permitan identificar a una persona?

Insistimos en este punto ya que el negocio de la "puesta en valor" depende precisamente de este concepto, ya que la anonimización sería la clave que descartaría la vulneración de las normas de protección de datos personales. Existe un debate abierto sobre la anonimización que, pese a remontarse ya algunos años, podría decirse que no ha hecho más que empezar y está todavía lejos de su fase de resolución. Es una discusión que, a nuestro juicio, resulta crucial para la sociedad del siglo XXI y que aún no ha alcanzado la suficiente presencia en los diversos foros concernidos (legales, éticos, técnicos, empresariales, gubernamentales) y en los que se debería generar el diálogo adecuado para ser, en primer lugar, comprendida, y, en segundo lugar, resuelta o al menos gestionada.

Como se ha dicho, en estos momentos, las evidencias técnicas ya nos muestran que es posible re-identificar a personas concretas a partir de los datos de un *dataset* sobre el cual se han aplicado técnicas de anonimización (o de-identificación). Una persona, o una empresa, puede conseguir la re-identificación si tiene la voluntad (por razones económicas, empresariales, delictivas…), los conocimientos y los medios técnicos para ello (por ejemplo, con los datos sanitarios de un hospital —sin datos personales— y acceso a los datos personales de otro *dataset* —digamos, un censo—). Parece evidente que, en el caso de los datos de salud, es fácil encontrar ese hipotético "*adversario*" con la motivación y los recursos para poder hacerlo y, por lo tanto, es acertado cuestionar la validez de las iniciativas de intercambio de datos sensibles que estén basadas en técnicas de anonimización. En el ámbito jurídico, el incierto recurso a la "anonimización", entendida como una solución definitiva pero irremediablemente en crisis, viene propiciado por la

---

OHM, Paul: "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization". *UCLA L. Rev*. 2009-2010, p. 731.

[12] SWEENEY L. Simple demographics often identify people uniquely. Carnegie Mellon University, editor. Data Privacy Working Paper 3. 2000. http://dataprivacylab.org/projects/identifiability/paper1.pdf. Consultado el 27 de enero de 2015.

[13] PAUL M. SCHWARTZ & DANIEL SOLOVE: "The PII Problem: Privacy and a New Concept of Personally Identifiable Information", *86 N.Y.U. L. Rev*. 1814 2011.

actual normativa de protección de datos cuyo origen se halla en una Directiva europea del año 1995, muy anterior al fenómeno del *big data*, y subyace en la Ley estatal 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. Sin embargo, desde el momento en que el propio anonimato deviene incierto es perentorio encontrar una base que legitime el análisis de datos personales de salud a gran escala. De no ser así, se abre la puerta a usos no deseados de esos datos ya que su titular, habiendo dado su consentimiento para determinadas acciones en el ámbito sanitario y de investigación, en realidad pierde el control y queda desprotegido pues —con una falsa concepción de la protección de datos y del secreto profesional— desconoce que sus datos pueden haber sido utilizados o cedidos para otros fines, ni deseados ni efectivamente consentidos.

Este Documento no pretende rechazar, sin más, este nuevo modelo de negocio que centra la atención del mercado y en el que ya estamos inmersos, sino que aspira a alertar sobre sus riesgos a los ciudadanos y a los poderes públicos que regulan y controlan la actividad en el ámbito sanitario y de investigación. Entre nosotros no se ha consolidado la conciencia social de la importancia de proteger los datos y su relación con el derecho fundamental a la intimidad y a la no discriminación. Carecemos de una cultura de la privacidad en materia de datos personales que nos permita comprender cómo puede afectarnos el hecho de que una empresa acumule y rentabilice nuestra información y tenga en sus manos un instrumento de poder para tomar decisiones que pueden afectarnos[14]. Por ejemplo, el análisis masivo de datos puede utilizarse para descubrir efectos secundarios de medicamentos, pero también permite crear perfiles de riesgo —incluso desconocidos por los afectados— que podrían utilizarse para "justificar" la denegación de una póliza de seguro.

Es evidente la urgencia de un debate que señale la vulnerabilidad de las personas ante los riesgos de discriminación basados en perfiles y patrones de comportamiento creados con fines que el afectado no puede controlar, y sobre la adaptación de las normas a los retos éticos y sociales planteados por las posibilidades que proporciona el despliegue de las tecnologías *big data*. Justamente es este el núcleo del conflicto que hay que abrir al debate público con la implicación de la ciudadanía, creando una cultura de privacidad acorde con los nuevos tiempos y las nuevas realidades[15].

Sobre el proyecto VISC +

Un ejemplo de reutilización de datos que a nuestro entender resulta cuestionable es el proyecto VISC+, impulsado por la Agència de Qualitat i Avaluació Sanitària de Catalunya (AQuAS) de la

[14] COHEN, JULIE: "What Privacy is for" 126 *Harv. L. Rev*. 1904 2012-2013.

[15] RICHARDS NM, KING JH. "Big data ethics". *Wake Forest Law Review* 2014;(49):393-432.

Generalitat de Catalunya, que tiene como objetivo —según describen sus promotores— poner la información sanitaria a la disposición de los ciudadanos, las empresas y la investigación, para mejorar los servicios de salud y la investigación y para 'poner en valor' el conocimiento.

El mencionado proyecto se nutre de las distintas bases de datos sanitarios existentes en el sistema, entre otros el SIDIAP (Sistema de información para el desarrollo de la investigación en la atención primaria) y, especialmente, de la HC3 (Historia clínica compartida de Cataluña) que recoge los datos asistenciales y de consumo farmacéutico, junto con otras informaciones relevantes referentes a la identificación y situación socio-sanitaria de cada ciudadano atendido en el sistema público; la HC3 contiene, también, la información de las pruebas analíticas y diagnósticas que incluyen datos metabólicos y bioquímicos, así como datos de diagnóstico genético que identifican a portadores de enfermedades genéticas hereditarias, o determinan riesgo o susceptibilidad de padecer enfermedades más complejas. Estas bases de datos conllevan la existencia de "ficheros de usuarios", de los cuales es responsable el Departament de Salut de la Generalitat de Catalunya. Si bien los macro ficheros de datos están protegidos por la normativa ya existente —en especial por la Ley Orgánica de Protección de Datos (LOPD) y el Reglamento que la desarrolla—, esta regulación resulta insuficiente ya que su aplicación ha sido superada por la nueva tecnología *big data* —como se ha indicado— y no evita los usos indebidos y discriminatorios[16].

La HC3 tiene los objetivos explícitos de: a) mejorar la atención de la salud de los ciudadanos mediante una herramienta que facilite el trabajo de los profesionales sanitarios; b) propiciar un nuevo modelo asistencial al permitir a los centros sanitarios de la red de asistencia pública el acceso y la consulta de forma inmediata, segura y confidencial de la información relevante disponible sobre los pacientes. Como es evidente, los datos que contiene son sensibles en extremo, y su recogida y tratamiento se justifican en la eficacia a la hora de prestar una buena asistencia, no solo en el centro habitual sino en toda la red pública catalana[17] ya que la HC3 permite el acceso de manera organizada, y bajo parámetros de seguridad y confidencialidad, a las historias clínicas de los centros sanitarios de la red asistencial. Esta herramienta debe ofrecer beneficios tanto a los profesionales sanitarios como a la ciudadanía y al propio sistema de salud. Por ello, el ciudadano tiene derecho a conocer quién dispone de sus datos personales y con qué finalidad; y también a exigir responsabilidades si considera que se está haciendo de los mismos un uso indebido o distinto de aquellos para los que hubiera otorgado el consentimiento en su momento. Cuando la HC3 se implantó, ni se informó suficientemente a los ciudadanos de dicha

[16] Como el Grupo de Opinión ya advirtió en el *Documento sobre pruebas genéticas de filiación*, Barcelona: Signo, 2006. Disponible en formato PDF en: http://www.publicacions.ub.edu/refs/observatoriBioEticaDret/documents/07899.pdf

[17] Existe un proyecto, de mayor alcance territorial, desarrollado en 12 países de la UE, en el marco de www.epSOS.eu, del que forman parte: Alemania, Austria, R. Checa, Dinamarca, Eslovaquia, España, Francia, Grecia, Holanda, Italia, Reino Unido y Suiza. En España participan tres CCAA: Andalucía, Castilla-La Mancha y Cataluña dentro del PLAN AVANZA para la modernización de los servicios de las administraciones públicas.

recogida masiva de datos, ni se indicó en ningún momento que pudieran llegar a reutilizarse para otras finalidades, incluso comerciales. Tampoco cabría considerar que la cesión de los datos de los usuarios del sistema público para finalidades no asistenciales fuera el "precio" de la asistencia gratuita, pues es improcedente exigir una contraprestación, ya que esta conllevaría que tal asistencia dejara de ser gratuita.

La HC3, pese a ser recogida y estructurada por los profesionales asistenciales, pertenece al paciente y, por esta razón, las entidades asistenciales reciben peticiones relacionadas con el ejercicio de los derechos que la normativa vigente en protección de datos reconoce a los ciudadanos: derechos de acceso, rectificación, cancelación y oposición (los llamados derechos ARCO). La finalidad de este conjunto de derechos es impedir un tratamiento de los datos personales ilícito y lesivo para la dignidad y el derecho del afectado *(habeas data)*, así como garantizar el ejercicio del más general derecho a la intimidad. Así, los usos de los datos recogidos en la HC3 deben limitarse a la asistencia (junto con los fines científicos —epidemiológicos, investigación y docencia— o encaminados a la mejora de los servicios públicos, que la normativa actual ya autoriza) y es absolutamente preciso establecer garantías que eviten el tráfico de datos y cualquier uso indebido de las empresas del ámbito de la salud (seguros médicos, corporaciones farmacéuticas, entidades financieras y otros). Por ello, el Proyecto VISC +, tal como en estos momentos está previsto que se lleve a cabo, genera dudas importantes, tanto de carácter bioético como estrictamente jurídico, que conviene debatir para prevenir su potencial uso discriminatorio.

Problemas relevantes del proyecto VISC+

1. Denominación equívoca del Proyecto

Consideramos que la propia denominación del proyecto es equívoca y no se adecua al "principio general de lealtad en la recogida y tratamiento de datos", porque induce a pensar que el único resultado del proyecto consiste en mejorar las condiciones de vida y la salud. Cualquier usuario a quien se pida el consentimiento para que sus datos personales de salud sean tratados en el marco de un proyecto así llamado[18] puede pensar, erróneamente, que colabora en un programa que solo le reportará beneficios. Esto puede relacionarse con una práctica contraria al principio jurídico de buena fe, más aun cuando los datos se recaban en un momento especialmente sensible, que puede distorsionar la capacidad de escoger con pleno conocimiento de causa, induciendo a facilitar información que de otro modo no se habría proporcionado. La lealtad es un valor fundamental en el marco de la LOPD, que prohíbe expresamente la recogida y el tratamiento de datos por medios fraudulentos o desleales, y afecta de manera directa al principio de calidad de los datos.

2. Limitación de las finalidades en el tratamiento de los datos

[18] VISC+ significa en catalán "VIVO más".

Como bien se recoge en el informe elaborado por el *Grupo Europeo de Ética de las ciencias y de las nuevas tecnologías* (GEE), de la Comisión Europea, en la recogida y tratamiento de datos de carácter personal las entidades públicas y las privadas deben fundamentar su actividad en el principio de "limitación de la finalidad"[19]; es decir, que este tipo de datos no deben ser recogidos ni tratados para cualquier uso, sino sólo con objetivos específicos y legítimos que se hayan prefijado. Además, los datos no han de estar a disposición de "quien los quiera utilizar" y los ciudadanos deben disponer de mecanismos efectivos para controlar y modificar las informaciones que les conciernen. Insiste también dicho informe en que cualquier posible cesión de datos con fines comerciales sólo debe hacerse con el consentimiento expreso de las personas afectadas, y con conocimiento del tipo de datos que se tratarán, con qué objetivo, durante cuánto tiempo y, si se van a relacionar o cruzar con otros datos procedentes de diferentes fuentes.

Resulta de especial importancia, en el marco del proyecto que aquí se analiza, conseguir una protección gradual de los datos en función de la finalidad del uso, distinguiendo cuidadosamente las finalidades sanitaria, epidemiológica y de investigación y docencia de las finalidades empresariales privadas basadas en la investigación, a las que hay que exigir el nivel de protección más elevado. Ahora bien, el proyecto VISC+ equipara tratamientos que tienen fines totalmente diferentes y esta confusión afecta a la legitimación para tratar los datos sanitarios, personales, que son datos especialmente sensibles y que, en consecuencia, están sometidos a una protección especial.

Como se ha indicado, la ley ampara el uso de datos del paciente para llevar a cabo la asistencia sanitaria y para la investigación y mejora de los servicios públicos. Si se quiere ir más allá y facilitar su utilización con fines no previstos ni autorizados —entre los cuales los intereses comerciales de empresas privadas cuyo producto depende de la investigación, entre otros factores— es necesario un debate social previo sobre la concurrencia de los intereses públicos y privados en la investigación, con el objeto de definir sus contornos éticos y el nivel de protección que se va a reconocer al ciudadano cuando empresas con intereses privados traten datos de salud. El *empoderamiento* del ciudadano se construye con información completa, clara y veraz sobre el uso de sus datos y reconociendo su facultad de controlar el tratamiento, consintiéndolo u oponiéndose al mismo.

3. Habilitación legal *vs.* consentimiento

El contar con la necesaria legitimación es el requisito básico para permitir el acceso de terceros a informaciones o datos que pertenecen al ámbito personal de los afectados. Se debe distinguir entre la legitimación legal, y la voluntaria, basada esta segunda en el consentimiento libremente otorgado. La primera permite tratar los datos con finalidades relacionadas con la atención

[19] EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES: Ethics of Security and Surveillance Technologies. Opinión n. 28, 20 de mayo de 2014. Véase: http://ec.europa.eu/archives/bepa/european-group-ethics/docs/publications/ege_opinion_28_ethics_security_surveillance_technologies.pdf

sanitaria, la calidad y gestión del servicio o fines científicos (epidemiológicos, investigación y docencia). Esta legitimación procedente de una ley se justifica en el interés público, respetando siempre escrupulosamente la confidencialidad de los datos recogidos y obligando a motivar la solicitud de uso de los mismos. Consideramos que el proyecto VISC+, habiendo sido aprobado por un mero "Acuerdo de gobierno", no cuenta con habilitación legal suficiente para la reutilización de datos sanitarios, ya que las leyes de sanidad sólo legitiman para tratar los datos de los pacientes con fines directamente asistenciales, investigadores u organizativos. El segundo tipo de legitimación, la voluntaria, proviene siempre del consentimiento expreso del paciente y es la que se precisa para tratar datos con fines estrictamente privados, es decir, sin interés público evidente; este consentimiento es el que se requiere para utilizar los datos de los usuarios en el desarrollo de las industrias sanitarias, farmacéuticas y de biotecnología, o la promoción y comercialización de sus productos. Consideración especial merecen los datos genéticos por la complejidad que supone su titularidad compartida por un núcleo familiar.

Partiendo de la existencia del "*partenariado* público-privado" en el sistema sanitario e investigador, el problema se centra en cómo se articula la legitimación para usar la información de salud y reutilizarla. A nuestro juicio, debe plasmarse en una ley y es preciso tener en cuenta la gran asimetría —de información e incluso de poder— existente entre el ciudadano, que sufre una enfermedad y necesita curarse, y el profesional que le va a pedir el consentimiento, tanto para procurarle la asistencia médica más adecuada como para el tratamiento de los datos personales de salud. Hay que tener claro que se trata de consentimientos diferentes, y que el acceso a la prestación sanitaria pública no se puede condicionar al consentimiento para tratar datos con otros fines, ni justifica la solicitud de datos adicionales. La conclusión es que la obtención del consentimiento debe someterse a garantías, a fin de compensar la situación de desequilibrio en que se encuentra el usuario de los servicios sanitarios públicos en momentos en los que puede estar especialmente preocupado por su salud —lo que genera una situación de vulnerabilidad— y que piensa *a priori* que todos los datos que se le piden se encaminan a su tratamiento y son necesarios para prestarle la asistencia que necesita y que constituye la razón por la que ha acudido al sistema sanitario.

4. Valor y riesgo

En este contexto, el proyecto VISC+ se presenta con el objetivo de "poner en valor" la enorme cantidad de datos de los que dispone el Departamento de Salud y con el fin de "reutilizar" estos datos para finalidades inicialmente no previstas y que, por ello mismo, el usuario no conoce. La expresión "poner en valor" mencionada tiene interpretaciones diferentes: por un lado, se trataría de poner a disposición de centros de investigación, centros de estudios epidemiológicos y de salud pública, y centros de docencia que así lo soliciten, datos de salud de los ciudadanos con el objetivo de contribuir al progreso del conocimiento en los ámbitos específicos mencionados y también, en último término, a la mejora de la atención sanitaria y la prevención, cosas perfectamente legítimas y deseables y que, reiteramos, la legislación ya permite. Pero, por otro lado, incluye una

interpretación mucho más laxa, que implicaría poner estos *big data* sanitarios a disposición de empresas que difícilmente entrarían en los ámbitos que se acaban de mencionar. Si tenemos en cuenta el principio bien conocido según el cual "*big data is big business",* el escenario más probable sería el de una pura y simple venta de los datos de salud de los ciudadanos en beneficio de la empresa que esté interesada en hacer rentable esta información y que disponga de los medios para hacerlo. Hay que tener siempre en cuenta que la tecnología no es el problema, sino la dirección que le imprime quien la usa y la financia. Así, el marco que plantea el Visc+ permitiría a las empresas implicadas extraer un valor meramente comercial de los datos.

Los riesgos potenciales a que se refiere este Documento no son hipotéticos, ni remotos: es suficiente con analizar la posibilidad de construir perfiles de conducta sobre datos anónimos, que se pueden utilizar en cualquier momento para tomar decisiones automatizadas sobre las personas. Basta un paseo por internet para encontrar una buena cantidad de empresas dedicadas a la compraventa de datos y cómo las que los poseen —originados en la prestación de otros servicios—, crean a su vez otras nuevas empresas y líneas de negocio dedicadas a la reutilización de estos datos consiguiendo perfiles muy precisos mediante sucesivos cruces de información y demás procesos de "enriquecimiento del dato".

5. Control de los datos

Precisamente en razón de evitar la pérdida de control de los datos y los posibles abusos, tiene sentido establecer funciones de *Data Governance* (es decir, el control del tratamiento y la gestión de los datos) que deben corresponder a las entidades públicas, garantes del escrupuloso respeto de los derechos fundamentales de los ciudadanos, con independencia de cómo se suministren los servicios profesionales por parte de las empresas adjudicatarias. Estas funciones de *gobernanza* deben incluir aspectos como: la seguridad y calidad de los datos, la privacidad, los procesos de anonimización, la trazabilidad, las políticas de permanencia de los datos, el enriquecimiento de datos (poniendo limitaciones a las fuentes o bases de datos con los que se pueden relacionar o cruzar). La preferencia por las garantías que otorga lo público deriva de las específicas obligaciones de transparencia y rendición de cuentas que la Administración tiene, si bien la responsabilidad compartida exigiría también transparencia y rendición de cuentas al sector privado. La Declaración Universal de Bioética y Derechos Humanos de la UNESCO, establece —en su artículo 14— el novedoso principio de responsabilidad social en salud que resulta del todo pertinente en este contexto y atañe también a la especial atención que se requiere para evitar los conflictos de intereses en este delicado campo.

El proyecto VISC+ habla de gobernanza, pero meramente de carácter interno mientras que en el presente Documento se propone un control externo y representativo de la sociedad. El mismo Informe de la Autoridad Catalana del Protección de Datos (APDCAT) ya señala que el VISC+ merece un régimen específico de seguridad aún más estricto. Según este informe, el Departament de Salut es "responsable" del tratamiento de los datos, a efectos legales. En cambio

"la entitat" (AQuAS) es “encargada” de su tratamiento. Asimismo, la mecánica prevista en el proyecto VISC+ para la explotación de datos también es motivo de preocupación ya que la empresa adjudicataria recibiría los datos supuestamente anonimizados a cambio de un precio o tasa; hay que advertir que en el mencionado proyecto también se dice que el adjudicatario participaría en el proceso de verificación de la anonimización y en la materialización de un "código anónimo de la persona" antes de transferir los datos a los usuarios finales. Se desconoce si sería esta misma empresa adjudicataria quien se responsabilizaría de “definir, construir y poner en marcha un catálogo de servicios útil, eficiente, competitivo e innovador, de contrastar las necesidades del mercado y los clientes finales, y de definir un plan de difusión y de comercialización, canalizando de manera adecuada la demanda del mercado nacional e internacional”.

Consideramos que, en el modelo VISC+ no queda en absoluto claro si serían las empresas adjudicatarias quienes decidirían a quién traspasan los datos de salud, presuntamente a cambio de compensaciones económicas. Contrapartidas que tampoco se precisan y de las que ni se menciona cómo revertirían en los ciudadanos. Por ejemplo, ¿se cedería la base de datos de enfermos de hepatitis C para desarrollar fármacos que después se pretenderían vender a 60.000 € cada tratamiento? Justamente esta cuestión sería la clave que condicionaría hasta qué punto los clientes o usuarios finales del proyecto estarían dispuestos a contribuir. El uso abusivo y opaco de los datos personales relacionados con la salud genera desconfianza en la población que impide que esos datos se empleen de modo legítimo para fines epidemiológicos, de investigación o docencia.

6. Evaluación del impacto

El proyecto VISC+ no incluye ninguna evaluación del impacto que su puesta en marcha pueda tener sobre el derecho a la intimidad de los usuarios, como sería pertinente teniendo en cuenta que estamos hablando de datos tan sensibles como los de salud, en la línea de lo que propone el proyecto de Reglamento europeo de protección de datos. La valoración del impacto —también ético y social— sería, obviamente, un requisito de cualquier ley que amparara el proyecto VISC+. Tampoco el Proyecto especifica en ninguna parte si los usuarios tienen que dar previamente su consentimiento para el trasvase de datos que implica, ni si entiende que este consentimiento no sería necesario dada la interpretación laxa del concepto de investigación a que se ha hecho referencia más arriba y que no diferencia entre el interés público y el interés privado.

Además, finalmente pero no menos importante, en ningún momento se explica claramente en el VISC+ cómo el beneficio económico que se obtuviera del Proyecto repercutiría favorablemente en los ciudadanos y en el sistema sanitario público. Parece evidente que, siendo los datos tan valiosos, cuando hay beneficios y negocio, los ciudadanos deberían contar con contrapartidas bien establecidas. Consideramos que el lucro es un fin lícito, pero no el bien primordial al que todos los demás valores y derechos deban subordinarse. El Convenio de Derechos Humanos y Biomedicina del Consejo de Europa, en vigor en nuestro país desde el

2000, establece —en su artículo 2— que los intereses de la ciencia o de la sociedad no deben prevalecer nunca sobre los del individuo; sobre esta premisa pivota todo el sistema de ciencia y tecnología y, en especial, el sistema sanitario y de investigación.

7. La experiencia del *National Health Service* británico

Es importante remarcar que las experiencias en la línea del Proyecto VISC + que se han llevado a cabo en países de nuestro entorno han generado conflictos muy significativos que llevan a extremar las medidas de prudencia antes de emprender proyectos de este tipo. En efecto, la supuesta anonimización de datos de salud y atención sanitaria recogidos por el *NHS Information Centre* (NHS-IC), entre 2005 y 2013, no ha impedido que diferentes empresas hayan re-identificado a las personas a quienes hacían referencia estos datos, generando perjuicios diversos; por ejemplo en el precio de las primas de riesgo de los seguros. La consecuencia ha sido una moratoria con el fin de reorganizar el procedimiento de cesión de los datos de manera que sea más transparente y garantice de forma más eficaz los derechos a la intimidad y a la confidencialidad. Actualmente el procedimiento de tratamiento y cesión de los datos está sometido, en todo momento, a auditoría y control público. Para ello se ha creado el *Care Data Advisory Group,* grupo consultor cuya misión es fortalecer la protección de los derechos de los ciudadanos en el ámbito sanitario, y la *National Data Guardian,* a fin de velar por la seguridad de los datos de salud.

RECOMENDACIONES

1. *Generar y potenciar una cultura ciudadana de la privacidad en materia de datos personales.* Acumular información sobre una persona representa adquirir un poder de decisión sobre ésta; por lo tanto, los medios para controlar quién trata nuestros datos, cómo los capta y con qué fin los usa se convierten en instrumentos de libertad personal y colectiva.
2. *Informar y formar sobre el alcance real de que los procesos de anonimización de los datos ya no garantizan la irreversibilidad.* Es factible actualmente la des-anonimización, re-identificación, o revelación de datos personales de conjuntos de usuarios y de usuarios individuales, pues las herramientas informáticas empleadas pueden servir tanto para esa finalidad como para la contraria.
3. *Alertar sobre la necesidad de redefinir el concepto mismo de "datos personales" sobre el que se asienta la legislación actual:* Es importante remarcar que el problema no es únicamente la transformación de los datos considerados personales en un conjunto de datos, pues incluso eliminando estos datos personales es posible llegar a re-identificar a una persona concreta.
4. *Aplicar escrupulosamente el principio que requiere que los datos recabados sean adecuados a la finalidad que motiva su recogida, y exigir que se solicite y se obtenga el consentimiento expreso de los usuarios para la utilización de los datos de salud con fines diferentes a aquellos para los que se obtuvieron, implementando mecanismos eficaces para otorgar, o denegar, dicho consentimiento*. En el caso de datos genéticos, el consentimiento debe ser especialmente riguroso dada la posible afectación a otros miembros del núcleo familiar.
5. *Establecer excepciones a la exigencia de consentimiento informado de la persona en estudios epidemiológicos tales como:* estudios de vigilancia epidemiológica promovidos por las autoridades; estudios de registros sanitarios creados o reconocidos oficialmente por las autoridades sanitarias; estudios en los que los datos estén debidamente anonimizados sin ninguna posibilidad de reidentificación de las personas; y estudios de salud pública de índole observacional, con riesgo mínimo, en donde la obtención del consentimiento informado sea impracticable o extremadamente costosa, siempre que cuenten con el dictamen favorable y de un comité de ética en investigación.
6. *Potenciar el proceso de información y de debate —basado en una honesta ponderación de riesgos y beneficios, ventajas y perjuicios— ante la puesta en marcha de un proyecto como el VISC +,* encaminado a la explotación de datos de salud que pertenecen a los ciudadanos, aunque estén en poder de la Administración, a fin de que puedan pronunciarse sobre la conveniencia de llevar a la práctica tal proyecto.
7. *Establecer mecanismos de control en el tratamiento de datos y concretar las funciones de "Data Governance" que son responsabilidad de los organismos públicos.* Los procesos de anonimización del dato deben llevarse a cabo dentro del perímetro interno de la Administración por las

garantías que otorgan las obligaciones de transparencia y rendición de cuentas que dicha Administración tiene, si bien la responsabilidad compartida exige también transparencia y rendición de cuentas al sector privado, especialmente en las redes mixtas público-privadas.

8. *Controlar de forma específica y reforzada la seguridad en el tratamiento de los datos sanitarios para garantizar en todo momento su correcto uso y evitar la comercialización, que no cuente con consentimiento expreso y no prevea de forma clara la manera en que el beneficio revierta a los ciudadanos.*

   Para ello debe crearse un órgano de gobernanza externo representativo de la sociedad catalana e independiente de los agentes implicados. Sus funciones consistirían en supervisar la gestión de los ficheros en todas las etapas, verificar la trazabilidad, detectar conflictos de intereses y malas prácticas, así como identificar posibles infracciones y responsabilidades e informar públicamente de su actuación y resultados. De esta manera, el ciudadano podría dirigirse a él para saber quién dispone de sus datos personales y con qué finalidad.

   Para llevar a cabo sus funciones dicho organismo debería promover y aplicar un *Código Ético para la Reutilización de Datos de Salud*.

9. *Reforzar la formación en deontología y ética profesional de los implicados en la custodia de los datos*, señalando que toda persona que tenga acceso o trate datos tiene el mismo deber de proteger y promover los derechos y libertades fundamentales de los implicados. El derecho a la intimidad, la confidencialidad y la no discriminación, son pilares del sistema de salud y de investigación.

10. *Recordar que los comités de ética deben contribuir al desarrollo de la cultura de respeto por la intimidad y la confidencialidad de los datos personales ya que, en su ámbito de actuación, son los primeros garantes de estos derechos*. Los miembros de los comités de ética, para poder colaborar en la toma de decisiones ponderadas y proporcionales, deben formarse en los aspectos éticos, técnicos, jurídicos y sociales de las tecnologías y la especial protección que requieren los datos sanitarios, así como prestar atención a no actuar de forma que les convierta en meros mecanismos de cobertura de intereses ajenos.

NORMATIVA DE REFERENCIA

Internacional

- Consejo de Europa: Convenio para la protección de los derechos humanos y la dignidad del ser humano con respecto a las aplicaciones de la Biología y la Medicina (Convenio relativo a los derechos humanos y la biomedicina), hecho en Oviedo el 4 de abril de 1997.
- UNESCO: Declaración universal sobre Bioética y Derechos Humanos, de 19 de octubre de 2005.

Europea

- Carta de los Derechos Fundamentales de la Unión Europea (DOUE núm. 83, de 30 de marzo de 2010).
- Directiva 1995/46/CE, de 24 de octubre, del Parlamento y del Consejo, sobre Protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.
- Directiva 2002/58/CE del Parlamento Europeo y del Consejo, de 12 de julio de 2002, relativa al tratamiento de los datos personales y a la protección de la intimidad en el sector de las comunicaciones electrónicas (Directiva sobre la privacidad y las comunicaciones electrónicas).

España

Sistema sanitario

- Ley 14/1986, de 25 de abril, General de Sanidad.
- Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.
- Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud.
- Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud.
- Ley 33/2011, de 4 de octubre, General de Salud Pública.

- Llei 21/2000 (del Parlament de Catalunya), de 29 de desembre, sobre els drets d' informació concernent la salut i l'autonomia del pacient, i la documentació clínica.
- Llei 16/2010 (del Parlament de Catalunya), de 3 de juny, de modificació de la Llei 21/2000, de 29 de desembre, sobre els drets d'informació concernent la salut i l'autonomia del pacient, i la documentació clínica.

Sistema de investigación

- Real Decreto 223/2004, de 6 de febrero, por el que se regulan los ensayos clínicos con medicamentos.
- Ley 29/2006, de 26 de julio, de garantías y uso racional de los medicamentos y productos sanitarios.
- Ley 14/2007, de 3 de julio, de Investigación biomédica.
- Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.

Protección de datos de carácter personal

- Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.
- Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Reutilización de la información del sector público

- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.
- Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, sobre reutilización de la información del sector público.

DOCUMENT ON BIOETHICS AND BIG DATA: EXPLOITATION AND COMMERCIALISATION OF USER DATA IN PUBLIC HEALTH CARE

PRESENTATION

The *Opinion Group of the University of Barcelona's Bioethics and Law Observatory* was established in 1996 as part of the Bioethics and Law Observatory. Among its objectives is the scientific and interdisciplinary analysis of the ethical, social and legal implications of new technologies and biotechnological and biomedical problems. In addition, it aims to influence the dialogue between the university and society through the transmission of scientific-technical knowledge and the presentation of arguments necessary to participate in a truly informed social debate. With these aims, the *Opinion Group* has already produced twenty-two papers[1] on current issues that continue to divide society or the different scientific communities involved. This work has required identifying problems, comparing arguments and proposing consensus recommendations.

Today, the *Group* presents the position paper*, Bioethics and Big Data: exploitation and commercialisation of user data in public health care.* This paper warns of the need to create a culture of privacy regarding personal data, as this data has become an element or mechanism of control in a digital society. It is now necessary to understand why this data should be protected. We analyse the problem of the exploitation and commercialisation of public health care user data from a bioethical perspective. Based on recognition of the principle of individual autonomy, the paper aims to show that the implementation of Big Data technologies in health, associated with an eventual commercialisation of this data, has a negative impact on the principles of equality and non-discrimination affecting not only the private life of citizens but also the quality and responsibility of the health care and research systems.

The spark for this paper has been the identification of problems in the project Visc+ — *Adding Value to Health Information in Catalonia.* We have identified two types of problems: 1) the possible violation of citizens' rights and 2) a lack of transparency and informed public debate on an issue which may involve the trafficking of personal health data, re-used for purposes other than that of the direct medical treatment of the patient. Thus, we raise questions that are related to this project but that have a wider scope: a) the effectiveness of anonymisation techniques on *datasets*, b) the need to redefine the concept of personal data, given the current capacity to re-identify individuals and c) the impact of these two issues on emerging markets for *big data* and *digital marketing*.

We believe it is necessary to take measures to ensure the rights of all individuals involved and their ability to make free and informed decision. We want to open a debate and formulate proposals to address the paradigm shift implied by these new information technologies, as the

[1] All documents of the *Opinion Group of the University of Barcelona's Bioethics and Law Observatory* are available in PDF format and online at: http://www.bioeticayderecho.ub.edu/documentos (in Catalan, Spanish and English). Consulted 27 January 2015.

actions of government in a democratic society cannot be imposed on citizens without access to prior, accurate and transparent information.

The development of this new position paper was coordinated by Doctors Maria Rosa Llàcer, María Casado and Lídia Buisan, and it has been written by the *Opinion Group of the Bioethics and Law Observatory of the University of Barcelona, the Consolidated Bioethics, Law and Society Research Group of the Generalitat of Catalonia,* with the collaboration of the *Research Group on Private Law, Consumption and New Technologies of the University of Barcelona (GREDINT)*. The names and professional profiles of all those who have participated in its preparation appear at the end of the paper.

GENERAL OBSERVATIONS

The challenges of big data and anonymisation

The term *big data* refers to the treatment of large volumes of data through mathematical algorithms in order to establish correlations among the data, predict trends and make decisions. *Big data* technologies represent a new paradigm and imply organisational changes for firms as well as public administrations. Companies today are more organised around data management than around improving processes. We are witnessing a transition toward *datification* and *monetisation* that involves extracting new value from data and making a profit on it; this is taking place in both the private and public sectors and with both in partnership. This is a trend embedded in a flourishing industry based on knowledge acquired through the re-use of data and its exploitation, and should be taken into account to contextualise the debate and understand this change in model. However, the commitment to innovation cannot ignore ethics and the essential rights of individuals, nor the protection of the public in the context of these new advances. This new situation must be analysed in order to propose a firm level of protection, which will strengthen existing systems and in this way foster a more advanced level of innovation in this area.

It is crucial to point out that until now the premise of the *anonymisation* of data represented a guarantee that made it possible to comply with regulations protecting existing personal data. The understanding has been that a set of anonymised personal data no longer contains data of a personal nature, thus no longer falling under the regulations for personal data protection, which were intended to be quite rigorous in both the EU and Spain. However, after nearly twenty years, and with advances in information technologies, these regulations have become largely obsolete. It is now recognised that anonymisation no longer guarantees the privacy of personal data, as computer engineering techniques now make it possible to re-connect anonymised data to the individual it belongs to[2]. De-anonymisation and subsequent re-

[2] Art. 29 Working Party Opinion 06/2013 on open data and public sector information ('PSI') reuse (1021/00/EN WP207); Art.29 Working Party Opinion 05/2014, on Anonymisation Techniques (0829/14/EN WP216). See: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/ wp216_ en.pdf

NARAYANAN, Arvin; FELTEN, Edward W. "No silver bullet: De-identification still doesn't work", 2014. See: http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf

NARAYANAN, Arvin; SHMATIKOV, Vitaly. "Robust de-anonymization of large sparse datasets". *Security and Privacy*. IEEE Symposium on IEEE, 2008, pp. 111-125.

DE MONTJOYE, Yves-Alexandre, *et al.* "Unique in the Crowd: The privacy bounds of human mobility". *Scientific reports, 2013, vol. 3.*

OHM, Paul: "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization". *UCLA L. Rev.*

identification are possible when technical expertise and means are available; thus, current debates over this problem are largely based on a technical field, from which information and arguments are obtained that directly affect the grounds on which the data-selling industry finds support. It is enough to know that re-identification can be done using the particular values of certain data, up to now not considered personal; for example, it is possible with a zip code, a date of birth and the sex of individuals to re-identify most individuals in a dataset[3]. Just as with fingerprints, certain types of data can unequivocally identify us. The problem this poses is significant: What is personal data and how can we guarantee its protection?[4] How can we avoid a non-personal dataset from being used to identify a person?

We emphasise this point because the business of *enhancing the value of data* depends precisely on the concept of anonymisation, since it is key to avoiding the violation of existing regulations to protect personal data. There has been an ongoing debate about anonymisation, which despite going back some years, we might say has only just begun and is still far from being resolved. This is a debate that, in our opinion, is crucial for society in the 21st century but that does still not have sufficient presence in important forums (legal, ethical, technical, corporate and governmental). It is in these spheres where adequate dialogue must be generated so that this issue can, first of all, be understood and secondly, resolved or at least managed.

At this time, technical evidence clearly reveals that it is possible to identify specific individuals using data from datasets which have undergone anonymisation (or de-identification). An individual or a business can achieve re-identification if they so desire (for economic, corporate or criminal reasons) and if they have the knowledge and technical means to do so (for example, with health care data from a hospital — without personal data — and access to personal data from another dataset — such as a census). It seems clear that in the case of health care data it is easy to find this hypothetical "actor" with the motivation and the resources to do just this, and it is therefore wise to question the validity of initiatives to exchange sensitive data based on anonymisation techniques. In the legal sphere, recourse to "anonymisation", understood as a definitive solution but irremediably in crisis, is facilitated by current regulations for the protection of data, whose origins are found in a European Directive from 1995, well before the emergence of *big data*, and in Spain's State law 37/2007 of 16 November, on the re-use of information from the public sector. However, from the moment that anonymity becomes uncertain, it becomes urgent to establish clear legal grounds for legitimising the large-scale analysis of personal health

2009-2010, p. 731.

[3] SWEENEY L. Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, editor. Data Privacy

Working Paper 3. 2000. http://dataprivacylab.org/projects/identifiability/paper1.pdf. Consulted 27 January 2015.

[4] PAUL M. SCHWARTZ & DANIEL SOLOVE: "The PII Problem: Privacy and a New Concept of Personally Identifiable Information", *86 N.Y.U. L. Rev*. 1814 2011.

data. If this is not done, it will open the door to the unwanted use of this data, as the *owner*, having given his or her consent for certain actions in the areas of health care and research, will in fact lose control of this data and will not be protected, given that with a false conception of the protection of data and secrecy, he or she does not know that this data can be used or passed on for other, undesired or unauthorised, purposes.

This paper does not simply reject this new business model focused on the market (and in which we are already immersed), but rather its aim is to warn us of the risks to individuals and to the public authorities that regulate and control activity in the areas of health care and research. Our society remains unaware of the importance of protecting data and its relationship to the fundamental rights to privacy and non-discrimination. We lack a culture of privacy when it comes to matters of personal data that will enable us to understand how we may be affected by the fact that a corporation accumulates our private information and makes a profit in doing so and, in addition, holds an instrument of power for making decisions that can affect us.[5] For example, large-scale analysis of data can be used to discover the side effects of medicines, but it can also be used to create risk profiles — without the knowledge of those who are affected — that could then be used to "justify" the denial of an insurance policy.

The urgency of this debate over the vulnerability of individuals to the risk of discrimination based on profiles and pattern of behaviour created for purposes that the affected individual has no control over is clear, as is the need for discussion over the adaptation of regulations to the ethical and social challenges raised by the deployment of big data technologies. This is precisely the core of a conflict that must be publicly debated, leading to the creation of a culture of privacy in accordance with these new times and realities.[6]

About the VISC+ project

One example of the re-use of data that we find questionable is the VISC + project, promoted by the Catalan Agency for Health Information, Assessment and Quality (AQuAS) of the Generalitat of Catalonia, which according to its promoters, aims to make health information available to citizens, businesses and for research purposes in order to improve health care services and research and to 'add value' to knowledge.

This project is based on the different existing health care databases in the Catalan health care system, among others the SIDIAP (Information system for the development of research in primary care) and especially the HC3 (Shared clinical history of Catalonia), which collects data

[5] COHEN, J "What Privacy is For" 126 *Harv. L. Rev.* 1904 2012-2013.

[6] RICHARDS NM, KING JH. "Big data ethics". *Wake Forest Law Review* 2014;(49):393-432.

on health care system users' clinical histories and pharmaceutical consumption as well as other relevant information regarding the identification and social and health situation of each individual assisted in the public system. The HC3 also contains information on test results and diagnoses that includes metabolic and biochemical data, as well as genetic diagnostic data to identify carriers of hereditary diseases or to determine risk or susceptibility to complex diseases. These databases contain “user files”, under the responsibility of the Health Department of the Generalitat of Catalonia. Although the macro data files are protected by already existing regulations (in particular, the Organic Law on Data Protection (LOPD) and the regulations it establishes), these regulations are not sufficient, as their application has been superseded by the impact of big data — as has been indicated — and they do not avoid improper and discriminatory uses of the data.[7]

The explicit objectives of the HC3 are: a) to improve the health care assistance provided to the population through a tool that facilitates the work of health care professionals, and b) to provide a new health care model by permitting health care centres in the public health care network access to available relevant information on their patients in a quick, secure and confidential manner. Clearly, HC3 data is extremely sensitive and the collection and treatment of it is justified by its effectiveness as a support in providing good health care, not only in the patient's usual centre but also throughout the Catalan public health care network[8], as the HC3 allows access, under the parameters of confidentiality and security, to the clinical histories held in the health care centres throughout the network. This tool should offer benefits for health care professionals, the population and the health care system itself. However, individuals have the right to know who has access to their personal data and for what ends, and to demand accountability if they believe that their data is being used in an improper way or in a manner different from what they consented to. When the HC3 was established, individuals did not receive sufficient information about this massive data collection, nor were they advised at any time that this data could be re-used for other purposes, including commercial ones. Nor was it considered that the transfer of data from users of the public health care system for other than health care purposes was the "price" for free assistance, as it is inappropriate to require a fee, as this would mean that such assistance ceases to be free.

---

[7] As the *Opinion Group* has already warned in the *Documento sobre pruebas genéticas de filiación* [Paper on genetic parental testing], Barcelona: Signo, 2006. Available in PDF format at: http://www.publicacions.ub.edu/refs/observatoriBioEticaDret/documents/07899.pdf

[8] A project of much greater territorial reach exists, referred to as epSOS [www.epSOS.eu], and developed in and formed by 11 EU countries (Germany, Austria, Czech Republic, Denmark, Spain, France, Greece, Holland, Italy and the United Kingdom) and Switzerland. In Spain, three autonomous communities are participants - Andalucía, Castilla-La Mancha and Catalonia - as part of the PLAN AVANZA for the modernisation of government services.

Although the HC3 data is collected and organised by health care professionals, it belongs to the patients, and therefore, health care agencies can receive petitions related to the exercise of citizens' rights recognised in current regulations on data protection: rights of access, rectification, cancellation and opposition (the so-called ARCO rights). The purpose of these rights is to prevent any use of personal data that is illegal and harmful to the dignity and rights of those affected (*habeas data*), as well as to guarantee the exercise of the more general right to privacy. Thus, the uses of data collected by the HC3 must be limited to providing health care (along with scientific ends — epidemiological, research and teaching — or aimed at improving public services, which current regulations already authorise) and it is absolutely necessary to establish guarantees that avoid the trafficking of data and any other improper uses by businesses in the health care sector (health insurance firms, pharmaceutical companies, financial and other entities). Thus, the VISC + Project, as it is envisioned at the moment, generates significant doubts, both bioethical and legal, that should be discussed in order to prevent the potential discriminatory use of its data.

Problems regarding the VISC+ Project

1.- The misleading name of the project

We believe that the very name of the project is misleading and does not conform to "the general principle of loyalty in the collection and treatment of data", as it leads one to think that the only result of the project is to improve living and health conditions. Any user who is asked to give consent for his or her personal health data to be collected and treated within the framework of a project with such a name[9] might mistakenly believe that he or she is collaborating in a programme that is entirely beneficial. This could be an example of a practice contrary to the legal principle of good faith, even more so when the data are collected at a particularly sensitive time, which may distort the individual's capacity to choose with full knowledge of the facts, thus leading to information being facilitated that would otherwise not be provided. Loyalty is a fundamental value within the framework of the LOPD, which expressly prohibits the collection and treatment of data by fraudulent or improper means, and directly affects the principle of data quality.

2.- Limitation of purposes in the treatment of data

As stated in the report by the *European Group on Ethics in Science and New Technologies* (EGE) of the European Commission, in the collection and treatment of personal data, public and private entities must base their activity on the principle of "purpose limitation"[10]; that is, that this type of

[9] VISC+ in Catalan means "to live more".

[10] EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES: Ethics of Security and

data must not be collected or treated for any purpose, but only for specific and legitimate preset aims. In addition, the data must not be at the disposition of "whoever wants to use them" and citizens must have effective mechanisms to control and modify the information that concerns them. This report also stresses that any possible transfer of data for commercial purposes must only be done with the express consent of the individuals affected and with knowledge of the type of data that will be treated, for what purpose, for what amount of time and if this data is going to be related or crossed with other data from different sources.

It is especially important within the framework of the project analysed here to obtain increased protection of data based on the purpose for which they will be used, carefully distinguishing between health care, epidemiological, research and teaching purposes and the purposes of private research firms, from whom we must demand a higher level of protection. Yet the VISC+ project equates the treatment of data with completely different ends, and this confusion affects the legitimacy of the treatment of personal health care data, which is especially sensitive data and which as a result, falls under special protection.

As has been indicated, the law protects the use of patient data for health care, research and the improvement of public services. If we wish to go further and facilitate its use for other, not authorised purposes — among these, those of private companies, whose products depend on research, among other factors — it is necessary to have a social debate beforehand on the correspondence between private and public interests in research in order to define the ethical contours and the level of protection that will be recognised for citizens when companies with private interests process health care data. Citizens are empowered when they have complete, clear and accurate information about how their data will be used, and when their right to control its treatment, either consenting to it or opposing it, is recognised.

3.- Legal authorisation *vs* consent

Having the necessary legitimacy is the basic requirement for allowing third parties access to the personal information and data of those affected. We have to distinguish between legal legitimacy and voluntary, the latter based on freely given consent. The former allows the treatment of data for purposes related to the provision of health care, the quality and management of services and scientific ends (epidemiological, research and teaching). This legitimacy proceeding from law is justified based on the public interest, always scrupulously respecting the confidentiality of the data collected and requiring a rationale for the use of these data. We believe that the VISC+ project, having been approved by a mere "government ruling", does not have sufficient legal authorisation to re-use health care data, as existing health care laws only legitimise the treatment

Surveillance Technologies. Opinión n. 28, 20 May 2014.
See: http://ec.europa.eu/bepa/european-group-ethics/docs/publications/ege_opinion_28_ethics_security_surveillance_technologies.pdf

of patient data for purposes directly related to health care, research and organisation. The second type of legitimation, voluntary, always comes from the express consent of the patient and is required to treat data for strictly private ends, that is, without clear public interest. This consent is required for the use of user data by the health care, pharmaceutical and biotechnology industries and in the promotion and commercialisation of their products. Genetic data deserves special consideration due to the complexity of its shared ownership by the nuclear family.

Given the existence of a "public-private *partnership*" in the health care and health research system, the problem is how to articulate legitimate use of health information and its re-use. In our judgement, this should expressed in a law, which must also address the great imbalance — in information and even power — that exists between individual citizens, suffering from illness and needing to be cured, and the professionals who are asking for their consent, both to provide the appropriate medical treatment and to treat their personal health data. It must be very clear that there are two different types of consent involved, and that an individual's access to public health care cannot be conditioned on having to give consent to provide data for other ends, or justify the request for additional data. The conclusion is that obtaining consent must be subject to guarantees that offset the imbalance of the situation, as the users of health care services, especially vulnerable given their concern about their health, assume a priori that all of the data they are being asked to provide will be used for their treatment and is necessary to provide the assistance required and for which they have turned to the health care system.

4.- Value and risk

In this context, the VISC+ project is presented with the aim of "adding value" to the enormous amount of data available to the Health Department through "re-using" this data for initially unforeseen purposes and that, for this reason, the user is unaware of. The expression "adding value" can have different meanings: on the one hand, it can refer to providing research centres, epidemiology research, public health centres, and teaching centres access to patient health data in order to contribute to the advancement of knowledge in the specific areas mentioned and, ultimately, to improving health care and prevention. These are goals that are perfectly legitimate and desirable and are permitted by law. On the other hand, however, it can be interpreted more loosely and may involve providing corporations who's purposes do not fit into the areas we have just described access to this health care *big data*. Taking into account that "big data is big business", a likely scenario would be the pure and simple sale of citizens' health data to companies interested in profiting from this information and with the means to do so. It must be kept in mind that technology is not the problem, rather, the issue is the aims of those who use and finance it. Thus, the framework posed by the VISC+ project would allow companies to extract a merely commercial value from the data.

The potential risks we refer to in this paper are not hypothetical or remote: it is enough to consider the possibility of creating behaviour profiles based on anonymous data, which can be used at any time to make automated decisions about individuals. It is enough just to look on the Internet to find the large number of companies dedicated to the buying and selling of data and how those that have data — originated in the provision of other services — are at the same time creating new companies and lines of business dedicated to reusing this data, obtaining very accurate profiles through successive data crossing and other "data enrichment" processes.

5.- Control of data

Precisely in order to avoid abuses and lose of control over data, it makes sense to establish functions for d*ata governance* (that is, controls over the treatment and management of data) that should correspond to public entities, who will act as guarantors of scrupulous respect for the fundamental rights of citizens, and independent of professional services that may be provided by contracted private firms. These functions of governance must include aspects such as: security and quality of data, privacy, anonymisation processes, traceability, data retention policies and data enrichment (placing limits on the sources or databases that it is possible to relate or cross with). The preference for guarantees provided by government derives from the specific obligations it has for transparency and accountability, although shared responsibility would also demand transparency and accountability in the private sector. UNESCO's Universal Declaration of Bioethics and Human Rights establishes in its article 14 the new principle of social responsibility in health, which is pertinent in this context and also applies to the special attention required to avoid conflicts of interest in this sensitive area.

The VISC+ project talks about governance, but only of an internal nature, whereas in the present paper we propose an external control that represents the interests of society. A report by the Catalan Data Protection Authority (APDCAT) already noted that VISC+ deserves a specific even stricter security regime. According to this report, the Health Department is legally "responsible" for the treatment of data. In contrast, "the entity" (AquAS) is "in charge of" its treatment. In addition, the mechanism anticipated in the VISC+ project for the exploitation of data is also cause for concern since a contracted company will receive the supposedly anonymised data in exchange for a price or a fee; it should also be noted that the mentioned project also states that the contracted company will participate in the process of verifying the anonymisation and in the realisation of an "anonymisation code for each individual" before transferring the data to final users. It is not known whether this same contracted company would be responsible for "defining, building and implementing a catalogue of useful, efficient, competitive and innovative services, comparing the needs of the market and end customers, and defining a plan for diffusion and commercialisation, adequately channelling the demands of national and international markets".

We believe that it is not at all clear in the VISC+ model if it will be contracted companies who will decide to whom health data is passed on to, presumably in exchange for financial compensation. Who these counterparts are is also not specified nor what impact they might have on the population. For example, will databases on individuals with hepatitis C be made available for the development of pharmaceuticals that afterwards would be sold for €60,000 per treatment? It is exactly such questions that condition the extent to which end clients or users of the project would be willing to participate. The abusive and opaque use of personal health data generates distrust in the population, which impedes the use of these data for legitimate epidemiological, research or teaching ends.

6.- Impact evaluation

The VISC+ project does not include any assessment of the impact that its implementation may have on the privacy of users along the lines of what is proposed in the draft European regulation on data protection; however, it should, given that we are referring to very sensitive health data. Evaluation of the impact - both ethically and socially - will obviously be a requirement of any law that regulates the Visc+ project. Nor does the project specify anywhere if users have to provide prior consent for the transfer of the data involved, nor is it clear if this consent will not be necessary given the lax interpretation of the concept of research to which previous reference has been made and which does not differentiate between public and private interests.

In addition, lastly, but no less important, no clear explanation is provided in the VISC+ regarding how the economic benefits of the project will favourably impact on the population and the public health care system. It seems clear that, given how valuable the data are, when there are profits the citizens should receive some established compensation. We believe that profit is a licit end, but not the primary end to which all other values and rights should be subordinated. The Convention on Human Rights and Biomedicine of the Council of Europe, in force in Spain since 2000, establishes - in its article 2- that the interests of science or of society should never prevail over those of the individual; the complete system of science and technology and, in particular, the health care and health research system revolves around this premise.

7.- The experience of the UK's National Health Service

It is important to emphasize that experiences along the lines of the VISC+ project that have been carried out in other countries in Europe have generated significant conflicts that call for extreme prudence before undertaking projects of this type. In effect, the supposed anonymisation of health care data gathered by the UK's *NHS Information Centre* (NHS-IC), between 2005 and 2013, did not keep different companies from re-identifying the persons to whom the data referred, generating diverse harmful effects, for example, in the price of insurance premiums. As a result, a moratorium on transferring data was instituted with the aim of re-organising the

procedures for doing so, so that the process would be more transparent and the rights to privacy and confidentiality more effectively guaranteed. Currently, the procedure for treating and transferring data is subject, at all times, to auditing and public control. A *Care Data Advisory Group*, a consulting group whose mission is strengthening the protection of citizens' rights in the health care sphere, was created to do this, along with a *National Data Guardian*, to ensure the security of health data.

RECOMMENDATIONS

1. *Generate and strengthen a public culture of privacy on matters of personal data*. Accumulating information about a person is acquiring decision-making power over him or her; therefore, the means to control who treats our data, how it is captured and with what ends it is used becomes an instrument of personal and collective freedom.
2. *Inform and educate about the real reach of data anonymisation processes as they no longer guarantee irreversibility.* It is now feasible to de-anonymise, re-identify or reveal the personal data of group and individual users, as the IT tools employed can serve for both this end and its opposite.
3. *Warn about the need to redefine the very concept of "personal data" upon which current legislation rests*: It is important to stress that the problem is not only the transformation of what are considered personal data in a dataset, as even eliminating these personal data it is possible to re-identify a concrete person.
4. *Scrupulously apply the principle that requires that data collected be appropriate for the ends that motivated their collection, and demand that the express consent of users for the use of their health data with ends different from those for which they were solicited and obtained, implementing effective mechanisms for granting or denying said consent*. In the case of genetic data, the consent must be particularly rigorous, given the possible effect on other members of the nuclear family.
5. *Establish exceptions to the demand for informed consent from persons in epidemiological studies such as*: epidemiological surveillance studies promoted by authorities; studies of medical record created or officially recognised by health authorities; studies in which the data are properly anonymised without the existence of any possibility of re-identification of persons; and studies on public health of an observational nature, with minimum risk, in which obtaining informed consent would be impracticable or extremely costly, always when they have received a favourable opinion from a committee on ethics in research.
6. *Strengthen the information process and debate - based on an honest weighing of risks and benefits, advantages and harm - before implementing a project like the VISC+*, aimed at the exploitation of health data belonging to the citizenry, although they are in the possession of the government, so that they can decide on whether to implement such a project.
7. *Establish control mechanisms for the treatment of data and concretise the functions of "data governance" that are the responsibility of public bodies*. The processes of anonymising data should be carried out within the inner perimeter of the administration for the guarantees that the obligations of transparency and accountability of said administration provide, although

shared responsibility also demands transparency and accountability from the private sector, especially in public-private networks.

8. *Specifically control and strengthen security in the treatment of health data to guarantee its correct use at all times and to avoid commercialisation without express consent and that does not clearly benefit the public.*

   To do this an external governing body, representative of Catalan society and independent of the actors involved, should be created. Its functions should consist in supervising the management of data files at all stages, verifying their traceability, detecting conflicts of interests and bad practices, as well as identifying possible infractions and responsibilities and publicly reporting on its own actions and results. In this way, individuals can go directly to this body to know who has their personal data and for what ends.

   To carry out its functions this body should promote and apply a *Code of ethics for the re-use of health data*.

9. *Strengthen education in deontology and professional ethics of those involved in the custodianship of data,* indicating that all persons that have access to or treat data have the same duty to protect and promote the rights and fundamental freedoms of the public. The right to privacy, to confidentiality and to not be discriminated against are pillars of the system of health and research.

10. *Remember that ethics committees must contribute to the development of a culture of respect for privacy and the confidentiality of personal data, as, in their sphere of action, they are the first guarantors of those rights*. The members of ethics committees, so that they can collaborate in making thoughtful and appropriate decisions, must be educated in the ethical, technical, legal and social aspects of technologies and the special protection that health care data requires. They must also give special attention to not acting in ways that lead them to become mere mechanisms for outside interests.

MEMBRES DEL GRUP D'OPINIÓ DE L'OBSERVATORI DE BIOÈTICA I DRET QUE HAN ELABORAT AQUEST DOCUMENT

Maria Rosa Llàcer
Catedràtica de Dret Civil, Universitat de Barcelona. Directora del Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya.

María Casado
Catedràtica d'Universitat, especialitat Filosofia del Dret, Moral i Política, Universitat de Barcelona. Directora de l'Observatori de Bioètica i Dret, del Màster en Bioètica i Dret de la Universitat de Barcelona i del Grup de Recerca Consolidat "Bioètica, Dret i Societat" de la Generalitat de Catalunya. Titular de la Càtedra UNESCO de Bioètica de la Universitat de Barcelona. Membre de la Comissió de Bioètica de la Universitat de Barcelona.

Lídia Buisan
Metgessa, especialista en anestesiologia, reanimació i tractamet del dolor i Doctora en Dret, Universitat de Barcelona. Antiga Professora d'Ètica Mèdica, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Pilar Antón
Antiga Professora Titular de Legislació i Ètica Professional de l'Escola Universitària d'Infermeria, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret i de la Comissió de Bioètica de la Universitat de Barcelona.

Àngels Avecilla
Ginecòloga. Cap de secció de Salut Sexual i Reproductiva, Badalona Serveis Assistencials. Màster en Bioètica i Dret, Universitat de Barcelona. Membre del Comitè d'Ètica d'Investigació de l'Hospital Germans Trias i Pujol.

Anna Badia
Catedràtica de Dret Internacional Públic i Directora del Màster en Estudis Internacionals, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Immaculada Barral
Professora Titular de Dret Civil, Universitat de Barcelona. Membre del Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya.

Blanca Bórquez
Advocada. Màster en Bioètica i Dret i Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Maria Jesús Buxó
Catedràtica d'Antropologia Cultural, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret i de la Comissió de Bioètica de la Universitat de Barcelona.

Lluís Cabré
Metge i Cap de la Unitat de Cures Intensives i d'Urgències de l'Hospital de Barcelona. President de l'Associació de Bioètica i Dret de la Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona, de la Comissió Deontològica del Col·legi Oficial de Metges de Barcelona i del Comitè de Bioètica de Catalunya.

Mirentxu Corcoy
Catedràtica de Dret Penal i Coordinadora del Màster en Dret Penal, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret, Universitat de Barcelona.

Fernando García López
Metge, especialista en Nefrologia. Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Madrid. Màster en Bioètica i Dret, Universitat de Barcelona.

Ricardo García Manrique
Professor Titular de Filosofia del Dret, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Carmelo Gómez
Professor Titular de Filosofia del Dret, Moral i Política, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Maria Dolors Gramunt
Professora Titular de Dret Civil, Universitat de Barcelona. Membre del Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya.

Carlos Humet
Metge i Director de l'Hospital de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Itziar de Lecuona
Professora Lectora del Departament de Salut Pública, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret i de la Comissió de Bioètica de la Universitat de Barcelona. Membre del Comitè d'Ètica en Investigació Clínica de l'Hospital Clínic de Barcelona i del Grup d'Experts LERU en Integritat en la Recerca.

Gemma Marfany
Professora Titular de Genètica, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Joaquim Martínez Montauti
Metge i Coordinador del Servei de Medicina Interna de l'Hospital de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Isabel Miralles
Professora Titular de Dret Civil i Secretària General de la Universitat de Barcelona.

Esther Mitjans
Professora Titular de Dret Constitucional, Universitat de Barcelona. Antiga directora de l'Autoritat Catalana de Protecció de Dades.

Mónica Navarro-Michel
Professora de Dret Civil i Vicedegana de la Facultat de Dret, Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Ismael Peña-Lopéz
Professor de Dret i Ciència Política, Universitat Oberta de Catalunya. Membre del Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya.

Glòria Pérez
Metgessa especialista en Medicina Preventiva i Salut Pública.

Francesca Puigpelat
Catedràtica de Filosofia del Dret, Universitat Autònoma de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Rosa Ros
Metgessa. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Albert Royes
Secretari de la Comissió de Bioètica de la Universitat de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Gemma Rubio
Professora de Dret Civil, Universitat de Barcelona. Membre del Grup de Recerca Consolidat "Dret Privat, Consum i Noves Tecnologies" (GREDINT) de la Generalitat de Catalunya.

Ana Sánchez Urrutia
Consultora de Bioètica del Secretari Nacional de Ciència i Tecnologia de Panamà. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.

Josep Santaló
Catedràtic de Biologia Cel·lular i President de la Comissió d'Ètica en Experimentació Animal i Humana de la Universitat Autònoma de Barcelona. Membre de l'Observatori de Bioètica i Dret de la Universitat de Barcelona.