

# APROXIMACIÓN PRÁCTICA AL USO DE LAS TECNOLOGÍAS BIG DATA EN EL SECTOR SALUD

*Méndez García, M.*  
*Hospital Clínic de Barcelona*

*Roda Faura, M.*  
*Hospital Clínic de Barcelona*

**SUMARIO\*:** 1. Conceptos previos. Que entendemos por tecnologías Big Data y especificidades del sector salud. 2. De la perspectiva normativa a la práctica en el uso de las tecnologías Big Data en el sector salud. 3. Propuestas prácticas para el desarrollo de proyectos Big Data por parte de las instituciones sanitarias. 4. Bibliografía.

## RESUMEN

Este trabajo pretende abordar a través de una breve aproximación práctica la problemática que existe a nivel legal para dar respuesta al uso actual que se realiza desde el sector salud por parte de las tecnologías Big Data, ya sea por motivos asistenciales o para investigación. Dadas las especiales características del sector se analizarán los principales problemas desde una perspectiva de la protección de datos pero también desde una perspectiva de la protección del conocimiento que hay detrás de estas tecnologías, así como los retos de futuro. Se concluye el trabajo con una serie de indicaciones prácticas a tener en cuenta por los centros sanitarios para poder poner en marcha proyectos Big Data.

## PALABRAS CLAVE

Big Data, investigación, salud, intimidad.

---

\* Para la realización de este trabajo se ha contado con el apoyo del Proyecto de Investigación BIGDATIUS (Uso de datos clínicos ante nuevos escenarios tecnológicos y científicos -Big Data-. Oportunidades e implicaciones jurídicas. UPV/EHU. Ref.:DER2015-68212-R), financiado por el Ministerio de Economía y competitividad y Fondo Europeo de Desarrollo Regional (FEDER).

## 1. CONCEPTOS PREVIOS. QUE ENTENDEMOS POR TECNOLOGÍAS BIG DATA Y ESPECIFICIDADES DEL SECTOR SALUD.

### Definición Big Data

Un elemento previo a analizar es la definición de que entendemos por Big Data y las características del sector salud que condicionan su uso.

Quando hablamos de Big Data hablamos del “conjunto de tecnologías, algoritmos y sistemas empleados para recolectar datos a una escala y variedad no alcanzada hasta ahora y a la extracción de información de valor mediante sistemas analíticos avanzados<sup>1</sup>”.

Desde un punto de vista técnico, una de las definiciones más comúnmente aceptada de las tecnologías Big Data es la acuñada por Laney Douglas en 2001 que las define en base a Velocidad, Variedad y

---

1 Definición dada por el Código de Buenas prácticas en protección de datos para proyectos Big Data editado por la Agencia Española de Protección de Datos

Volumen, es la denominada de las 3 V:

- Velocidad: el procesamiento de la información se realiza a gran velocidad.
- Variedad: uso de gran variedad y tipología de datos, así como procedentes de varias fuentes.
- Volumen: se analizan grandes volúmenes de información.

Esta definición, realizada desde un punto de vista técnico, se complementa con el término Conocimiento, ya que podemos tener grandes volúmenes de información, sistemas muy potentes para analizarlos, pero debemos tener Conocimiento que nos permita interpretar esos datos, para que puedan ser de alguna utilidad. Este último dato, es especialmente relevante en el sector salud, ya que se trata de aplicar el método científico, sino analizamos los datos, con Conocimiento, podemos tener grandes volúmenes de información sin utilidad alguna.

Así mismo debemos tener en cuenta que un mismo dato en el ámbito sanitario puede tener diferentes significados o trascendencia en función del origen, forma de recogida o entorno, por lo que debemos analizar estos datos con el conocimiento necesario. Por ejemplo un dato tan sencillo como la temperatura puede tener diferentes relevancias en función de cómo se haya tomado (axilar, anal o bucal) a la franja de edad de a quien pertenezca el dato o la precisión de los instrumentos con que se tomó.

### **Características del tratamiento de información en el sector salud mediante tecnologías Big Data.**

a) Un elemento que condiciona el uso de las tecnologías Big Data en el sector salud, es el formato en que encontramos la información, ya que para poder utilizar tecnologías Big Data necesitamos disponer de información en un mismo lenguaje estructurado.

Si bien es verdad que existe cada vez más información en lenguaje estructurado como por ejemplo los CMBD (Conjunto Mínimo Básico de Datos), también existe mucha información que se encuentra en lenguaje natural, como por ejemplo los cursos clínicos. Pero incluso a veces ese tipo de información estructurada se encuentra en formatos diferentes, por ejemplo la codificación de las tarjetas sanitarias es distinta en las diferentes comunidades autónomas.

b) En cuanto al origen y tipología de datos, toda la información que encontramos en el sector salud es especialmente sensible, y desde un punto de vista de la normativa de protección de datos se exige las máximas cautelas en cuanto a su tratamiento.

Conjuntamente con la información de salud, encontramos la información de tipo genética, el uso de la que está en auge. El *Reglamento (UE) 2016/679 del Parlamento y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas relativo al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE*, define «datos genéticos», como todos los datos personales relativos a las características genéticas de una persona que hayan sido heredadas o adquiridas, (...) que proporcionen una *información única sobre la fisiología o la salud de esa persona, obtenidos en particular del análisis de una muestra biológica de la persona de que se trate*;

Este tipo de información, se caracteriza por:

- Son datos obtenidos a través del análisis de muestras biológicas humanas, sometidas por ellas mismas a un régimen específico de obtención y conservación. En este sentido, la *Ley 14/2007, de 13 de julio de Investigación Biomédica* establece una serie de requisitos.
- Son datos singulares que proporcionan información única acerca del sujeto fuente de la muestra, y no sólo afectan al sujeto del que proviene la muestra sino también a su familia.
- El análisis de datos genéticos proporciona una gran cantidad de información sensible, y el uso del mismo sin las cautelas adecuadas puede dar lugar a comportamientos discriminatorios basados en ese tipo de información.

c) La anonimización de la información de salud en general y la genética específicamente en entornos Big Data genera una amplia discusión, ya que hay posiciones que apuntan que siempre se podrá saber a quién pertenece la información. En este sentido debemos acudir al *Dictamen 04/2007 del Grupo de Trabajo del Artículo 29* en que analiza el concepto de dato personal y cuando se puede considerar anonimizado.

Por este motivo, tal y como expondremos más adelante la anonimización por sí sola no basta en estos entornos, se debería acompañar de medidas

adicionales enfocada a minimizar el riesgo de re identificación, no deseado. Hablaríamos de seudonimización, más que de anonimización. En este sentido es a destacar la definición que realiza el Reglamento Europeo del concepto de dato seudonimización como “tratamiento de datos personales de manera tal que ya no pueda atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y que esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable”.

Aquí también podemos introducir otra variable en cuanto a la anonimización, la necesidad o no de que esta sea reversible, para poder notificar resultados que puedan ser relevantes para la salud de los sujetos.

d) Finalmente, debe destacarse que cuando hablamos de uso de tecnologías Big Data en el sector salud, encontramos una gran casuística de proyectos:

- Uso de datos para proyectos asistenciales impulsados por ley y de tipo gubernamental (p.ej es el caso del proyecto de Historia Clínica Compartida de Cataluña, programa de receta electrónica<sup>2</sup>).
- Uso de datos para proyecto con finalidades de investigación de tipo gubernamental (p.ej es el caso del programa PADRIS<sup>3</sup>).
- Uso de datos para proyectos internos de investigación de los centros sanitarios públicos.
- Uso y cesión de datos a terceros por parte de las instituciones sanitarias para el desarrollo de proyectos de investigación.
- Uso de datos para grandes proyectos internacionales de datos genéticos (por ejemplo es el caso de los grandes proyectos de secuenciaciones genéticas como el “Chronic Lymphocytic Leukemia (CLL) Genome Project”<sup>4</sup>).

Cada caso debe analizarse por separado, ya que las condiciones de uso de los datos de los personales son distintas. No es lo mismo utilizar los datos

2 [http://salutweb.gencat.cat/ca/ambits\\_tematicos/linies\\_dactuacio/tecnologies\\_informacio\\_i\\_comunicacio/](http://salutweb.gencat.cat/ca/ambits_tematicos/linies_dactuacio/tecnologies_informacio_i_comunicacio/)

3 [http://aquas.gencat.cat/es/proyectos/analitica\\_dades/](http://aquas.gencat.cat/es/proyectos/analitica_dades/)

4 <http://www.cllgenome.es/welcome-cll-genome-web-site>

de los pacientes para ser incorporados en la Historia Clínica Compartida de Cataluña, donde existe una habilitación legal y los datos son identificados, que la petición de uso de datos de repositorios oficiales como son los CMBD por parte de empresas privadas.

Por tanto como conclusión a este primer punto podemos afirmar que el uso de tecnologías Big Data en el sector salud se caracteriza por:

- a) El análisis de grandes volúmenes de datos no tiene sentido si no va acompañado de los conocimientos que permitan su correcta interpretación.
- b) Los datos se pueden obtener de multitud de fuentes, pero tiene que garantizarse que su formato y características sean equivalentes.
- c) Los datos deben estar anonimizados o se debe legitimar su uso.
- d) Los datos de salud y genéticos, son datos especialmente sensibles y como tal la normativa de protección de datos les concede una especial protección.
- e) Existe una gran diversidad de proyectos de Big Data que su tipología y características condicionan el uso de los datos.

## 2. DE LA PERSPECTIVA NORMATIVA A LA PRÁCTICA EN EL USO DE LAS TECNOLOGÍAS BIG DATA EN EL SECTOR SALUD.

Atendiendo al análisis realizado en el punto anterior debemos distinguir dos normativas que inciden en el análisis de los Proyectos Big Data en el sector salud, por un lado la normativa que afecta a la protección de la intimidad del paciente y por otro lado la protección del conocimiento aportado por los profesionales sanitarios.

### 2.1 De la protección de la intimidad de los pacientes

Desde un punto de vista de la protección de datos debemos tener en cuenta la siguiente normativa:

- Ley Orgánica 15/1999, de Protección de Datos Personales (en adelante “LOPD”).

- Real decreto 1720/2007, Reglamento de desarrollo de la LOPD (en adelante “RLOPD”)<sup>5</sup>.
- Reglamento (UE) 2016/679 del Parlamento y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas relativo al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (en adelante el “Reglamento Europeo”).
- El Reglamento Europeo será plenamente aplicable a partir del 25 de mayo del 2018, por lo que vamos a referenciar conjuntamente el Reglamento Europeo y la LOPD.

Y en lo que se refiere a los datos de tipo genético no podemos olvidar que el origen de los mismos son las muestras humanas, por lo que también deberemos tener en cuenta lo establecido por:

- Ley 14/2007, de 13 de julio de Investigación Biomédica (en adelante “LIB”).
- Real Decreto 1716/2011, de 18 de noviembre, por el que se establecen los requisitos básicos de autorización y funcionamiento de los biobancos con fines de investigación biomédica y del tratamiento de las muestras biológicas de origen humano, y se regula el funcionamiento y organización del Registro Nacional de Biobancos para investigación biomédica (en adelante “RLIB”).

Así mismo también debemos tener en cuenta el *Documento de trabajo sobre datos genéticos adoptado el 17 de marzo de 2004 por el Grupo del artículo 29 sobre protección de datos* y el anteriormente mencionado *Dictamen 05/2014 sobre técnicas de anonimización adoptado el 10 de abril de 2014 por el Grupo de Trabajo sobre Protección de datos del Artículo 29*.

De acuerdo con lo establecido en la LOPD y el Reglamento Europeo, los datos de salud y los datos genéticos se podrán tratar en las siguientes condiciones:

- Si los datos están anonimizados no se consideran dato personal y por tanto no se les aplicará la normativa de protección de datos. En este sentido debemos hacer especial referencia al concepto de seudonimización.

- Existencia de consentimiento expreso del titular de los datos.
- El tratamiento es necesario para proteger un interés vital o interés público. El Reglamento Europeo, prevé que para determinadas finalidades como pueden ser las relacionadas con un interés público esencial, las de investigación científica, las relacionadas con la atención sanitaria o social, o las relativas a salud pública, los tratamientos serán posibles en las condiciones que determine la legislación europea o nacional. Esta excepción se concreta en la habilitación legal para llevar a cabo los tratamientos asistenciales por los profesionales de la salud, es decir cuando los pacientes ingresan en un centro asistencial para recibir tratamiento no deben prestar su consentimiento para que sus datos se utilicen con finalidades asistenciales. En este caso únicamente se les debe informar del uso que se hace de sus datos, y pedir el consentimiento para otros usos o cesiones no autorizadas por ley.

Estos criterios que desde el punto de vista teórico son muy claros, se tornan en extrema complejidad cuando lo queremos aplicar en entornos Big Data. A continuación exponemos algunos ejemplos prácticos, que anteriormente hemos enunciado para ilustrar estas dificultades:

#### **a) Uso de datos en el proyecto de Historia Clínica Compartida de Catalunya (HC3) y en el Programa de Receta Electrónica.**

En el primer caso el Proyecto HC3 busca la mejora de la atención de los pacientes mediante una herramienta que facilite el trabajo a los profesionales de la salud que los deben atender, y en el segundo caso se trata de un sistema que integra los procesos de prescripción y dispensación de la prestación farmacéutica mediante las TIC y el trabajo en red y que establece mecanismos enfocados a racionalización en el uso de los medicamentos. Ambos proyectos desde un punto de vista legal no presentan problemas, ya que su base jurídica son dos normas con rango legal, en el primer caso la *Ley 29/2006, de 26 de julio, de garantías y uso racional de los medicamentos y productos sanitarios*, y por otro lado, y en el segundo caso por la *Ley 21/2000, de 29 de diciembre, sobre los derechos de información concerniente a la salud y la autonomía del paciente, y a la documentación clínica*.

<sup>5</sup> A fecha de este trabajo estas dos normas están vigentes y pendientes de ser modificadas para que no se contradigan con el Reglamento Europeo de Protección de Datos.



**b) Uso de datos en el proyecto PADRIS.**

El programa PADRIS es la nueva denominación que se da al fallido programa VISC+. El objetivo del PADRIS es impulsar y facilitar la investigación, la innovación y la evaluación de la salud mediante el acceso y el análisis de la información del Sistema Sanitario Integral de Utilización Pública de Cataluña (SISCAT), bajo la responsabilidad del AQUAS (Agencia de Calidad y Evaluación Sanitaria de Catalunya).

Este programa se basa en los principios, de respeto a las personas, solidaridad entre pacientes, justicia, eficiencia, transparencia, investigación responsable y protección de datos personales. La información del Programa PADRIS pues se obtendrá del sistema sanitario catalán, y los usuarios del mismos podrán ejercer su derecho y oponerse a que sus datos se utilicen para el programa PADRIS (Opt-out).

Los diferentes centros que quieran acceder a los datos deberán firmar un convenio con el AQUAS, comprometiéndose al cumplimiento de las condiciones de acceso al programa. Cada solicitud concreta de acceso al PADRIS deberá estar debidamente justificada y pasará por el filtro de un Comité de Ética de la Investigación (CEI).

En cuanto a la anonimización, se aprecian detalles interesantes como por ejemplo que se reconoce que habrá datos que por su naturaleza no se puedan anonimizar, por ejemplo datos genómicos, en este caso será imprescindible el consentimiento de su titular. Así mismo se prevé que cada proyecto deberá contener un estudio estadístico del riesgo de reidentificación, por lo tanto se reconoce el hecho que la anonimización es un concepto dinámico, es decir en función del uso que se dé a los datos existirá riesgo de re identificación.

Pero lo que si podemos comprobar es que los tres proyectos tienen un denominador común y es que no existe un consentimiento específico del titular de los datos para su uso, si bien si existe una información y/o una anonimización del dato. El hecho que no exista una consentimiento expreso por parte del titular de los datos (menos en el caso que hemos indicados de proyectos de investigación que necesitan usar datos identificados en el marco del programa PADRIS) es precisamente porque la necesidad del mismo dificultaría las virtudes de las tecnologías de análisis masivo de los datos.

**c) Uso de datos por parte de un Hospital o un conjunto de Hospitales para realizar un proceso de mejora de la asistencia de los pacientes.**

Cada vez es más frecuente que se utilicen las tecnologías Big Data para generar proyectos que reviertan en la mejora de la asistencia a los pacientes. En este tipo de proyectos no se suelen solicitar el consentimiento de los pacientes, sino que se anonimiza.

En este caso no podemos afirmar que se trate de finalidades asistenciales, pero que inciden directamente en la mejora de la asistencia a los pacientes. Su finalidad se podría decir que está relacionada con la mejora de la asistencia.

En este tipo de proyectos el riesgo más elevado a ser valorado, es la intervención de terceros con intereses económicos y el uso de la información que puedan realizar.

En este punto resulta especialmente importante, además de valorar la importancia y relevancia del proyecto y el beneficio indirecto que puede suponer en la asistencia de los pacientes, garantizar la anonimización y que no se emplearan técnicas de re identificación o cruce de datos para intentar averiguar a quien pertenece la información.

**d) Uso de datos de por parte de un centro sanitario o conjunto de centros sanitarios en el marco de un proyecto de investigación interna.**

El uso de tecnologías Big Data para investigación quizá sea el más complejo de analizar, ya que en el caso del uso de datos para investigación, sin tecnologías Big Data, más allá de aquellos proyectos que por sí ya disponen de un consentimiento específico (p.ej ensayos clínicos y estudios prospectivos) es una práctica extendida en el sector que los estudios retrospectivos internos, es decir sin participación de terceros, donde no se tiene contacto con el paciente se lleven a cabo sin el consentimiento de este, y de forma codificada que no anonimizada.

El estudio se realiza de forma codificada, sin que se pueda saber quién es el paciente, pero el investigador suele guardar la clave por si necesita volver a identificar al paciente, por dos motivos:

- Por una información especialmente relevante para la salud de los pacientes cabría plantearse la necesidad desde un punto de vista ético de comunicar estos resultados.
- Por motivos de metodología a veces para llevar a cabo el estudio se necesita volver al origen de los datos para revisar nuevas informaciones que inicialmente no se detectaron que fueran relevantes.

Esta situación práctica, no tiene una base legal clara, aunque a la práctica suele aceptarse si el paciente está debidamente informado de esta práctica, se garantiza que en el transcurso de la investigación sólo se utilizará información codificada y que únicamente el investigador podrá acudir a la información origen en casos muy limitados.

Cuando trasladamos este escenario a los proyectos Big Data se tornan aún en más complejos estos problemas, y más si añadimos el elemento de dato genético que dificulta la consideración del dato como anónimo.

En estos casos cobran especial importancia los elementos de anonimización y participación de Comités éticos en su evaluación, tal y como expondremos en el punto tercero de este trabajo.

Así pues, y a modo de resumen, casos como la HC3 o el proyecto de Receta Electrónica son proyectos dirigidos a la mejor asistencia del paciente y con base legal, que desde un punto de legitimación no generan grandes dificultades. Donde realmente encontramos los principales problemas es en el uso de las tecnologías Big Data para proyectos de investigación realizados por Hospitales con o sin la intervención de terceros. Estas dificultades aumentarán cuando los datos que se utilicen sean de tipo genético, ya que justificar la anonimización aún resulta más complejo

## **2.2 Protección del conocimiento**

Tal y como hemos expuesto en el primer punto de este trabajo, además de disponer de grandes tecnologías y datos estructurados, para hacer un correcto análisis de los mismos es necesario que exista un conocimiento que los sepa interpretar correctamente.

En el caso de los proyectos Big Data del sector salud este conocimiento los aportan los profesionales

sanitarios, por tanto y de conformidad a lo establecido en la normativa de propiedad intelectual e industrial partimos de que los resultados de la investigación en el marco de su relación laboral son de la institución donde presta sus servicios (artículo 15 Ley 11/1986, de 20 de marzo de Patentes, y artículo 51 Real Decreto Legislativo 1/1996, de 12 de abril, texto refundido de la Ley de Propiedad Intelectual).

Esta previsión se matiza o modifica a través del artículo 14 de la *Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación*, que prevé la participación en los beneficios que pueda obtener la institución en la que presta servicios en investigador, por el trabajo llevado a cabo por el mismo.

En este sentido se deberá proteger este conocimiento de los profesionales de la institución, y en consecuencia de la institución.

## **3. PROPUESTAS PRÁCTICAS PARA EL DESARROLLO DE PROYECTOS BIG DATA POR PARTE DE LAS INSTITUCIONES SANITARIAS**

A la espera de un futuro cambio normativo que concrete en las especificidades de los proyectos Big Data, las instituciones sanitaria pueden seguir una serie de guías o cautelas que les permitan desarrollar este tipo de proyectos con un mínimo de seguridad jurídica:

### **a) Información y derechos de los pacientes.**

Este elemento es básico para garantizar el uso de adecuado de los datos, es informar a los pacientes de forma clara y adecuada del uso que se va a dar a sus datos. En este sentido los pacientes que entran en un centro sanitario se les informa del tratamiento que van a recibir sus datos, este sería el momento en que se debería dar la información de cualquier uso presente y futuro de los mismos, por ejemplo se les podría informar de:

- Uso de los datos en proyectos gubernamentales con base legal, como por ejemplo HC3 y firma electrónica.
- Uso de sus datos de forma anonimizada o seudonimizada en proyectos gubernamentales como por ejemplo PADRIS.

- Uso de sus datos para proyectos internos de investigación relacionados con las patologías que sufren de forma anonimizada o codificada.
- También se les debería informar de la posibilidad de ejercer sus derechos.

Si el paciente dispone de toda la información del uso que se va a dar a sus datos, y estas finalidades son legítimas, ya sea por qué se anonimiza o porque responden a un interés público y que no se contradice con el mismo, y además conocen la posibilidad de ejercer sus derechos, esto supone un plus de legitimación en el tratamiento de la información.

#### **b) Técnicas de anonimización o seudonimización y medidas enfocadas a evitar la reidentificación.**

No sólo se deben implantar unos sistemas de anonimización robustos en el sentido *Dictamen 05/2014 sobre técnicas de anonimización adoptado el 10 de abril de 2014 por el Grupo de Trabajo sobre Protección de datos del Artículo 29*, sino que dado el entorno en que nos encontramos y en especial en los datos de tipo genético, contractualmente se deberían establecer unas garantías y medidas que aseguraran que en especial en los proyectos colaborativos no se van a emplear técnicas para re identificar la información, en este sentido hablaríamos de seudonimización.

#### **c) Creación de comités de evaluación de proyectos Big Data o ampliación de funciones de los Comités de ética de la Investigación.**

Es imprescindible que los proyectos de Big Data, antes de implantarse sean examinados por un Comité multidisciplinar que analice la conveniencia del proyecto desde un punto de vista técnico, jurídico y ético. En el caso de los proyectos de investigación la parte ético-jurídica puede ser analizada por el Comité de Ética de la Investigación del centro, y es conveniente que incorpore un experto en materia de Big Data.

Resulta especialmente relevante la intervención de los Comité de Ética o de otro creado *ad hoc* para esta finalidad, en aquellos casos que no exista ni consentimiento ni base legal para

explotar los datos, en este caso el Comité deberá analizar, en base a unos criterios predeterminados, si existe una legitimación para poder tratar la información y el si el grado de anonimización o seudonimización es suficientemente fiable. Este caso lo podríamos asimilar a la excepción, que según la LIB, que pueden conceder los Comités de Ética de la Investigación para utilizar muestras sin consentimiento del titular de los datos. Los CEI deberían elaborar un protocolo<sup>6</sup> de actuación dónde fijen los criterios en base a los que se evaluarán este tipo de proyectos.

Otro elemento que deberán valorar estos Comités, con independencia de la existencia del consentimiento, es la necesidad de llevar a cabo este tipo de proyectos, ya que bajo la etiqueta de Big Data, se llevan a cabo múltiples proyectos, de los que sólo se debería permitir llevar a cabo los que sean verdaderamente útiles. Se trataría de incorporar una valoración desde el punto de vista de oportunidad del proyecto para garantizar que no se destinen recursos, en un sector donde son tan limitados como el de la investigación, a finalidades superfluas o de escasa utilidad.

#### **d) Creación de políticas de propiedad intelectual.**

Finalmente, y como elemento conectado a la protección del conocimiento las instituciones a través de la previsión realizada por el artículo 14 de la *Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación*, pueden establecer mecanismos prácticos para garantizar que la institución tendrá protegido el conocimiento que aporta cuando participe en este tipo de proyectos con terceras empresas. Uno de estos elementos son las Políticas de Propiedad intelectual y industrial que pueden implantar los centros y en las que se establezcan las bases y los circuitos a seguir en los proyectos colaborativos con otras instituciones y la participación que los investigadores pueden llegar a tener si el proyecto genera beneficios.

6 Evaluación por un CEI de proyectos de investigación con tecnología BIGDATA. Guillermo Alcalde y Iciar Alfonso. Proyecto BÍ DATIUS. <http://www.bigdatius.com/documentos/evaluacion-por-un-cei-de-proyectos-de-investigacion-con-tecnologia-bigdata/>

#### 4. BIBLIOGRAFÍA:

- 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 d diciembre, de protección de datos de carácter personal.
- Ley 14/2007, de 3 de julio, de Investigación Biomédica.
- Real Decreto 1716/2011, de 18 de noviembre, por el que se establecen los requisitos básicos de autorización y funcionamiento de los biobancos con fines de investigación biomédica y del tratamiento de las muestras biológicas de origen humano, y se regula el funcionamiento y organización del Registro Nacional de Biobancos para investigación biomédica.
- Reglamento (UE) 2016/679 del Parlamento y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas relativo al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE
- Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.
- Ley 29/2006, de 26 de julio, de garantías y uso racional de los medicamentos y productos sanitarios.
- Ley 21/2000, de 29 de diciembre, sobre los derechos de información concerniente a la salud y la autonomía del paciente, y a la documentación clínica.
- Declaración Internacional sobre los Datos Genéticos Humanos de Naciones Unidas. Paris, 16 de octubre de 2003.
- Instrumento de Ratificación del 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 de Oviedo*), hecho en Oviedo el 4 de abril de 1997.
- Dictamen 05/2014 sobre técnicas de anonimización adoptado el 10 de abril de 2014 por el Grupo de Trabajo sobre Protección de datos del Artículo 29.
- Statement on Statement of the WP29 on the impact of the development of Big Data on the protection of individuals with regard to the processing of their personal data in the EU, adopte on 16 Septiembre 2014. Article 29 Data Protection Working Party.
- Documento de trabajo sobre datos genéticos adoptado el 17 de marzo de 2004 por el Grupo del artículo 29 sobre protección de datos.
- Marco de actuación para el uso compartido responsable de datos genómicos y relativos a la salud de la Alianza Global para la Genómica y la Salud. Versión de 10 de septiembre de 2014.
- Código de Buenas prácticas en protección de datos para proyectos Big Data editado por la Agencia Española de Protección de Datos.
- Web proyecto PADRIS  
- [http://aquas.gencat.cat/ca/projectes/analitica\\_dades/](http://aquas.gencat.cat/ca/projectes/analitica_dades/)
- Web proyecto BIG DATIUS  
- <http://www.bigdatius.com/>