

Open data



Carlos Pintos

Open data

Reutiliza la información pública
para crear una nueva empresa



coordinadores de la colección:
Juan Vicente García Manjón y
José Luis Marín de la Iglesia

OPEN DATA

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

No está permitida la reproducción total o parcial de este libro, ni su tratamiento informático, ni la transmisión de ninguna forma o por cualquier medio, ya sea electrónico, mecánico, por fotocopia, por registro u otros métodos, sin el permiso previo y por escrito de los titulares del Copyright.

QR code es una marca registrada por Denso Wave, inc.

DERECHOS RESERVADOS 2012, respecto a la primera edición en español, por

© Netbiblo, S. L.

netbiblo

www.netbiblo.com

NETBIBLO, S. L.

c/. Rafael Alberti, 6 bajo izq.

Sta. Cristina 15172 Oleiros (La Coruña) – Spain

tlf: +34 981 91 55 00 • fax: +34 981 91 55 11

www.netbiblo.com

editorial@netbiblo.com

Miembro del Foro Europeo de Editores

ISBN: 978-84-9745-487-2

Depósito Legal: C-1785-2012

Directora Editorial: Cristina Seco López

Editora: María Martínez

Imagen interior: © Photosani

Imagen cubierta: © Sergey Ilin

Producción Editorial: Gesbiblo, S. L.

Impreso en España – Printed in Spain

El autor

Carlos Pintos es actualmente coordinador de proyectos en la entidad pública empresarial Red.es adscrita a la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información. Ingeniero informático con perfil de especialización en Ingeniería del software ha colaborado en multitud de proyectos innovadores relacionados con nuevas tecnologías en distintas áreas de negocio.

Con más de 15 años de experiencia profesional, ha desempeñado cargos relevantes tanto en el sector público como en la empresa privada. A lo largo de su trayectoria ha prestado sus servicios profesionales en distintas empresas, entre las que se encuentra Accenture, Eresmas Interactiva, la Biblioteca Nacional Española, Entel y Red.es, donde actualmente desarrolla proyectos de impulso al desarrollo de la Sociedad de la Información en España.

Pocket Innova

La serie POCKET INNOVA nace con el objetivo de acercar la disciplina de la innovación a directivos, académicos, estudiantes y técnicos de empresa interesados en conocer el qué y porqué de la innovación en nuestros días.

En un mundo en constante cambio y evolución, la innovación se ha convertido en la respuesta natural de aquellos que desean aportar soluciones creativas, basadas en el conocimiento y que aporten valor a las organizaciones y a la sociedad.

POCKET INNOVA pretende abordar, con cercanía y rigor, las distintas temáticas que componen la innovación. Para ello, la colección se estructura en cinco grandes bloques temáticos:

- **Tecnología.** Esta sección tiene como objetivo acercar a los lectores la información sobre las tecnologías emergentes, los últimos avances y las aplicaciones a nuevos productos y servicios.
- **Financiación.** Desde las subvenciones hasta el capital riesgo, los *business angels*, el VII Programa Marco de I+D y un largo etcétera, conforman los mecanismos de financiación de la I+D+i, los cuales serán abordados en la serie de una forma clara y práctica.
- **Actores.** Existen multitud de organismos y expertos que trabajan en materia de innovación en España, desde los organismos públicos hasta los privados, fundaciones, centros tecnológicos, expertos, redes... El quién es quién de la innovación en España y en Europa es el objetivo de esta sección.
- **Gestión de la innovación.** Sección central de la serie, la cual incluye metodologías prácticas que se pueden aplicar a la gestión de la innovación en las organizaciones. Desde la implantación de sistemas de I+D+i hasta la gestión de proyectos tecnológicos; desde la vigilancia tecnológica hasta la propiedad industrial y la transferencia tecnológica, todos ellos son temas cruciales para la puesta en marcha de organizaciones innovadoras.
- **Capital humano y creatividad.** Las personas son el elemento central de la innovación, su conocimiento, la creatividad, el trabajo en equipo, el liderazgo para la innovación y muchos otros temas componen una atractiva temática que todos aquellos que quieran adentrarse en el mundo de la innovación deben conocer.

1_ Información y globalización	
1_1 Internet, acceso a datos, globalización	12
1_2 Información del sector público	13
2_ ¿Qué es la reutilización? Open data	
2_1 ¿Por qué reutilizar?	18
2_2 ¿Para qué reutilizar?	21
3_ Cómo reutilizar	
3_1 Antecedentes históricos del Open data y la reutilización	26
3_2 Qué información se puede reutilizar	28
4_ La reutilización en España	
4_1 Reutilización en la empresa pública	38
4_2 Proyecto Aporta	42
5_ El sector infomediario	
5_1 Quién reutiliza	50
5_2 Características de las empresas infomediarias	53
5_3 Qué ofrecen las empresas infomediarias	58
5_4 Repositorio de datos	59

6_ Tecnología asociada a la reutilización

6_1	Principios básicos de partida	68
6_2	Formatos de reutilización	73
6_3	Raw data	81

7_ Marco jurídico y licencias

7_1	Marco jurídico y licencias	88
7_2	Modalidades de reutilización y licencias aplicables	97

8_ Reutilización e innovación

8_1	Definir la actividad de una empresa infomediaria	103
8_2	Realizar la caracterización de los clientes	104
8_3	Definir las fuentes de los datos.....	106
8_4	Definir las estrategias de comercialización	108
8_5	Realizar un plan de negocio	110

Bibliografía	113
---------------------------	-----





1_Información y globalización



1_1 Internet, acceso a datos, globalización

Desde siempre se ha escuchado la frase “La información es poder”, esta frase implica un modelo de gestión empresarial en el que los procedimientos, productos, catálogos, diseños, etc. no se hacían públicos. Formaban parte de los activos de la empresa, sujetos la mayoría de las veces a estrictas normas del copyright y por ello eran celosamente guardados.

Pese a que históricamente esta ha sido la forma de actuar, en ocasiones ha traído consigo grandes fracasos. En 1971 la compañía Sony inventó el primer reproductor de vídeo en color, lo cual llevó al desarrollo y producción de los vídeos Betamax en 1975. No obstante, Sony con el objetivo de controlar el mercado a través de esta tecnología, estableció unas condiciones de uso de la misma que hizo que a nivel mundial solo fueran aceptadas por 12 compañías, pese a ofrecer mayor calidad que la tecnología que un año después desarrolló la competencia. Este nuevo estándar, se abrió al mercado y fue apoyado por 40 fabricantes. En consecuencia, los vídeos VHS ocuparon mayor cuota de mercado y poco a poco relegaron al Betamax a una segunda posición hasta su desaparición en el mercado ciudadano. En otras palabras, un estándar tecnológico de calidad superior al de los competidores (de hecho hasta la entrada del almacenamiento digital y la desaparición de las cintas de vídeo se seguía utilizando en entornos profesionales) se vio abocado a su desaparición por una política empresarial proteccionista.

Quizás este fue un caso muy conocido, pero desde luego no el único. Hoy en día la filosofía ha cambiado, en un mundo empresarial donde globalización es una de las palabras más escuchadas, la empresa se enfrenta a retos en los cuales aparecen nuevos mercados, nuevas oportunidades y, al mismo tiempo, nuevos agentes que compiten directamente entre ellos. Por ello desde el punto de vista de la información, ha perdido valor su posesión frente al disponer de acceso a la misma. Y el motivo es simple, el volumen de esta información es tan ingente y cambia a tal velocidad que solo su productor puede mantenerla, actualizarla y garantizar su validez.

Y es en este escenario donde la capacidad de acceso al conocimiento y a la información es lo realmente valioso; donde cobra peso Internet.

Internet es el gran facilitador que nos garantiza la consulta de esa información, nos ofrece una plataforma a través de la cual tanto una multinacional como una pequeña pyme, esté dónde esté, pueden acceder a los mismos datos y competir ambas en igualdad de condiciones. Así, sería posible que una empresa con sede, pongamos, en Valladolid, sea la que tenga acceso a la información sobre las noticias de la NBA (la liga profesional de baloncesto americana) en Estados Unidos y se convierta en el primer proveedor independiente de esta información a nivel mundial. Por cierto, esta empresa se llama Hoopshype¹, fue creada en 2002 y adquirida en 2006 por Fantasy Sports Ventures por una cifra no desvelada².

1_2 Información del sector público

Independientemente de lo anterior, surge una cuestión que hay que responder, ¿quién es el encargado de generar y mantener la información? Es obvio que es mucho peor tener acceso a una información errónea o desactualizada que no tenerlo a ninguna información, especialmente si no se es consciente de este hecho.

Por ello la información del sector público ha tomado un peso específico tan importante. Partimos de un conjunto de premisas, en primer lugar en el ejercicio de sus funciones los distintos ministerios y agencias públicas generan gran cantidad de información: mapas de diversa índole, datos demográficos, información meteorológica, estudios económicos, etc. No hay duda de que uno de los sectores con mayor proyección económica actualmente es el asociado a los contenidos digitales, y este sector encuentra en la Administración Pública un productor de información fiable, completa y de calidad.

Pero adicionalmente hay otro punto a tener en cuenta que, aun en el caso de que esa información no fuera la mejor, no hay que olvidar el hecho de que esa información proviene de una fuente oficial gubernamental lo que

¹ www.hoopshype.com

² García Manjón, J.V. (2010).

le proporciona una presunción de validez y calidad que no tiene ninguna otra fuente. Es más, en muchos casos, y esto es importante tenerlo en cuenta, aunque no sea información totalmente actualizada, sí es la información oficial y, por lo tanto, es la que debe utilizarse.

Desde hace años, nuestros responsables políticos han sido conscientes de la importancia de la información del sector público. Ya en el año 2006 la Comunidad Económica Europea elaboró el estudio MEPSIR, *Measuring European Public Sector Information Resources*³, en el que se estima el valor del mercado de la reutilización entre los 26 y los 47 mil millones de euros solo en la Unión Europea. Dicho de otra forma, las entidades públicas, en el ejercicio de sus funciones, generan información, pagada con el dinero de los ciudadanos, que es susceptible de ser reutilizada. Esa información, por lo tanto, no tiene ningún sentido que no sea reutilizada para generar riqueza para mejorar la economía del país y, en consecuencia, producir un retorno hacia esos ciudadanos.

Actualmente los mercados locales están copados, mucha gente intenta repetir modelos de negocio que antes de empezar ya están agotados. Es necesario buscar nuevas oportunidades, e Internet ahí es el elemento diferenciador: el canal de distribución, la fuente de aprovisionamiento, el medio en el que publicitarse...

Sin duda hay otros, pero hoy en día ningún canal como Internet ofrece la posibilidad de innovar, crear y desarrollar un negocio. Si a esto añadimos los cambios legislativos que hacen que más información pública se ponga a nuestra disposición, estamos en un momento ideal para poder innovar y crear un negocio de éxito.

³ http://ec.europa.eu/information_society/policy/psi/mepsir/index_en.htm

2_¿Qué es la reutilización? Open data



De acuerdo a la Wikipedia, Open data es “una filosofía y práctica que persigue que determinados datos estén disponibles de forma libre a todo el mundo, sin restricciones de copyright y patentes u otros mecanismos de control”.

Básicamente consiste en que las empresas, las administraciones y los particulares ofrezcan de forma libre los datos que producen en el ejercicio de sus actividades. No importa si se consideran datos de especial calidad. Sin duda, podemos afirmar que la información que una entidad produce es útil; aún en el caso de que esa entidad no sepa sacarle partido, ofrezcámosla a la comunidad y alguien habrá con una idea capaz de generar riqueza a partir de ella. Con esto no queremos decir que el productor de la información que se ofrece esté renunciando totalmente a los derechos que posee como generador de la misma (más adelante trataremos este tema en un capítulo específico) pero sí que puede establecer licencias y/o permisos para definir el uso que se debe dar a la información, estableciendo cláusulas por las que, por ejemplo, quien desee utilizar la información no pueda tergiversar su contenido y cuando la utilice deba citar la fuente origen de la información, generando de esta forma un retorno directo hacia el productor de la misma.

Haciendo un poco de historia, podemos comentar que el Open data como tal surgió fuera del entorno empresarial, se centraba más en el mundo académico donde distintos investigadores fueron conscientes de que, al liberar el resultado de sus investigaciones, esto traía consigo que otros investigadores pudieran aportar valor o incluso que surgieran nuevas vías de investigación a partir de los datos originales. Se basa en principios tan conocidos como “los datos pertenecen a la raza humana”.

Es obvio que la composición del genoma humano, por ejemplo, no es propiedad de una compañía. El disponer de datos obtenidos a partir de una investigación sobre su composición, puede ofrecer una ventaja competitiva respecto a otras empresas de cara a la producción de medicamentos pero, por otro lado, no se puede patentar un hecho como es en este caso la composición del genoma humano. Pero esta filosofía va mucho más allá del mundo académico...

Por todos es conocido el desastre que en 2010 asoló la isla de Haití. Un terremoto de magnitud 7 en la escala Richter, con un epicentro situado solo a 15 kilómetros de la isla asoló la ciudad de Puerto Príncipe. La situación en la que se encontraba la zona era francamente dramática; Haití es el país más pobre de América Latina: más de 300.000 fallecidos, 1.500.000 personas sin hogar. Las distintas iniciativas internacionales, loables sin duda, se encontraron con la dramática situación en la que se encontraba el país. Coordinar cualquier tipo de actuación en la zona era una tarea titánica por falta de medios de todo tipo. Revisando los mapas digitales disponibles de la zona del desastre, se disponía de una información bastante pobre.

En esos momentos la compañía GeoEye, cuyo negocio se construye a partir de fotos cartográficas obtenidas vía satélite, siendo consciente del valor y del poder de su información y conociendo la necesidad que había de la misma, liberó de forma gratuita las imágenes de satélite que disponía de esta ciudad, antes y después del terremoto. Esta información fue puesta a disposición de la comunidad internacional, de forma que geógrafos de todo el mundo comenzaron a elaborar de forma conjunta un mapa de Puerto Príncipe, recogiendo información sobre la ciudad antes del terremoto y el estado de la misma tras él (edificios, servicios públicos, infraestructuras, etc.), todo ello publicado en tiempo real generando de esta forma el mapa de referencia para cualquier agencia, periodista u organización no gubernamental que quisiera desempeñar su trabajo o llevar a cabo acciones humanitarias en la zona.

Sin duda los motivos que movieron a GeoEye, al igual que a otras muchas compañías y particulares, eran totalmente altruistas. Nadie puede dudar del valor social de esta información y las consecuencias de su liberalización, pero si nos centramos en el punto de vista empresarial, ¿cuánto costaría una campaña de marketing a nivel mundial donde en prensa, radio y televisión se hable de GeoEye, del trabajo que desempeñan y el compromiso con la sociedad en general y con la recuperación de Haití en particular?

2_1 ¿Por qué reutilizar?

En el escenario económico mundial actual, nos encontramos que Occidente está en algunos segmentos empresariales en una situación de desventaja frente a otras economías emergentes en las cuales el coste de mano de obra y de la producción en general son muy inferiores. En parte también por la propia madurez del proceso productivo y los costes asociados al mismo, que están muy ajustados, y por la reducción de los costes de transporte.

Actualmente, en la industria automovilística los fabricantes europeos están trasladando parte de su producción a países orientales o de Europa del Este, dejando en las fábricas tradicionales la tarea de montaje final del vehículo. Y realmente si no se produce todo el vehículo en las nuevas fábricas orientales es, en parte, por las ayudas y condiciones establecidas con los gobiernos locales y el impacto social que tendría el cierre de fábricas en dichos países pese a que, inevitablemente, surgen noticias de reestructuración y reducción de personal en las mismas.

Este hecho hace que surja la necesidad de nuevos mercados donde la innovación permita la creación de nuevos productos que, de forma natural, puedan generar nuevos servicios y valores de cara a mejorar la economía local.

Y si hay un segmento empresarial virgen que destaque en esta área, es el sector de las empresas infomediarias, un nicho por explotar. Este sector empresarial está compuesto por un conjunto de empresas que, utilizando información generalmente del sector público, generan valor a través de servicios, aplicaciones y productos.

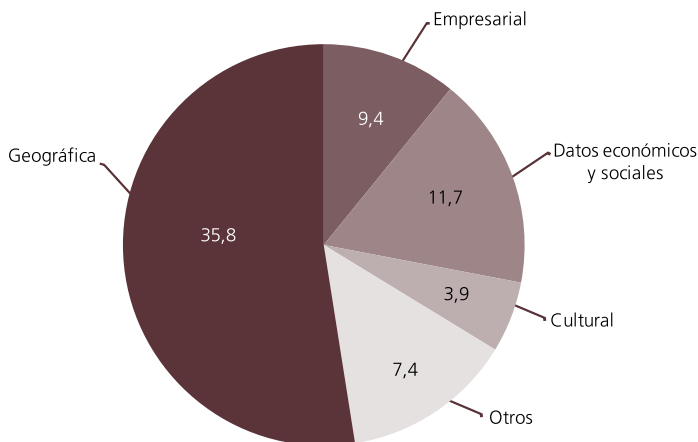
Si somos una gran empresa que durante años ha desarrollado su actividad en un determinado sector, es posible que dispongamos de una base de conocimiento e información que nos sirva para generar contenidos y comerciar con ellos. Pero lo cierto es que en la gran mayoría de los casos no es así, una empresa de nueva creación es difícil que disponga de esa base de conocimiento con

lo que la solución es reutilizar el conocimiento disponible ajeno a nuestra empresa y utilizar la innovación para ofrecer nuevos productos y servicios a partir de él. De esta forma se puede entrar de forma competitiva en mercados maduros, o incluso crear nuevos mercados y competir en igualdad de condiciones con otras empresas con posibilidades de éxito.

El principio puede parecer complicado, ten una idea y busca la forma de llevarla a cabo, pero es sencillo: se basa en encontrar algo que podamos hacer antes que nadie o mejor de lo que se hace con procedimientos tradicionales. Por ejemplo, en Inglaterra se creó www.upmystreet.com que, a partir de GoogleMaps como base cartográfica, utiliza información pública proporcionada por distintas entidades para ofrecer un servicio demandado. Reutiliza información sobre escuelas, accesibles de forma gratuita a través del Ministerio de Educación, datos de criminalidad del Ministerio del Interior, índices de polución del Ministerio de Medio Ambiente, etc. Así, combinando todos estos datos, esta página web ofrece la información georreferenciada de forma que a partir de una dirección postal se obtiene acceso a los datos de criminalidad, polución, la lista de escuelas en la zona incluyendo las notas medias por asignaturas, los servicios cercanos (hospitales, parques, etc.)... A partir de una idea, han ofrecido esta información de forma gratuita y rápidamente gran cantidad de inmobiliarias han comenzado a utilizarla como herramienta de soporte al proceso de venta de inmuebles. Desde la web solo con los anunciantes y ofreciendo un valor añadido a esos clientes que demandaban un servicio un poco más especializado han conseguido crear un negocio de éxito. De hecho, la página ha llegado a tener más de 1.000.000 de visitas mensuales.

Las personas a las que este concepto, en principio tan abstracto, les resulte nuevo pueden no ser conscientes del volumen de negocio del que se está hablando. Tal y como comentábamos anteriormente, el volumen potencial estimado en 2006 del mercado de la reutilización de información del sector público se encuentra entre los 26 y los 47 mil millones de euros solo en la Unión Europea. Y, para dar un dato más reciente, en el

Figura 2.1. Valor de los distintos tipos de información del sector público en la Unión Europea (en miles de millones de euros).



Fuente: Informe Pira (2000).

Estudio de caracterización del Sector Infomediario del 2012, se realiza una proyección sobre el total de empresas del sector estimando que el volumen de ingresos de las empresas infomediarias solo en España oscila entre los 330 y los 550 millones de euros.

Se ha nombrado algún caso de éxito que nos puede motivar a investigar en las posibilidades del Open data y la reutilización de información del sector público. Pero obviamente hay otros motivos. Cualquier persona que haya constituido una empresa o incluso que haya estudiado el hacerlo es consciente de los costes asociados de dicha tarea así como los necesarios para progresar. Los costes de la información del sector público son muy bajos cuando no inexistentes, y esa es nuestra materia prima. Ciertamente es que dicha información requiere un tratamiento, pero se trata de establecer los procedimientos, mecanismos y herramientas (lo cual implica una inversión inicial) ya que posteriormente los costes de gestión y tratamiento proporcionalmente son mucho menores.

2_2 ¿Para qué reutilizar?

En este momento surgen dos puntos de vista opuestos pero complementarios. Por un lado están los productores de la información y, por otro lado, los consumidores de la misma.

Desde el Ministerio de Industria, Energía y Turismo a través de la Secretaría de Estado de Telecomunicaciones y Sociedad de la Información, se han hecho eco de las directrices europeas y su trasposición a la legislación española y en los últimos años han lanzado varias iniciativas de cara a impulsar la implantación de la cultura de la reutilización entre la Administración Pública. Como iniciativa estrella cabe destacar la realización de un proyecto de cara a fomentar el uso de la reutilización de la información del sector público, el Proyecto Aporta⁴. Uno de los resultados de este proyecto es la publicación de una guía de reutilización (que actualmente ha sido actualizada con otro documento) donde se establecen, desde la perspectiva de la Administración Pública, todos los aspectos que una entidad pública debe conocer para reutilizar su información.

Si nos referimos a las razones que llevan a una entidad pública a querer reutilizar, anteriormente ya hemos esbozado los motivos. En primer lugar la información de la que una entidad pública es garante no es de su propiedad. Es el depositario de la misma, es la encargada de mantenerla y de velar por su integridad, pero esa información ha sido pagada con dinero público o, lo que es lo mismo, ha sido financiada a partir de los impuestos de los ciudadanos y, por lo tanto, éticamente hablando, estos son los propietarios últimos de dicha información. Por ello la entidad pública pierde el derecho a retener esa información que, en base a nuestras consideraciones, no le pertenece (siempre y cuando no lo haga para asegurar una serie de supuestos como podrían ser garantizar la privacidad de los ciudadanos o por motivos de seguridad nacional, entre otros).

El objetivo de cualquier Gobierno es mejorar la calidad de vida de sus ciudadanos y hoy en día sabemos que en gran parte esto depende de tener una economía saneada y, en consecuencia, del bienestar económico de

⁴Antiguamente, www.aporta.com, que ha sido sustituido por datos.gob.es para adaptarse a la nueva normativa.

las familias. Todo ello pasa por la creación de empresas y negocios que impulsen esos objetivos, trayendo consigo la creación de riqueza que inevitablemente revierte en el bienestar de la sociedad en general. Es un proceso cíclico. Por ello, es obligación de la Administración Pública el poner los medios para que este proceso se realice. La información del sector público es un elemento a tener en cuenta (por supuesto no el único) de este proceso y por ello no puede ser despreciado o ignorado.



Es responsabilidad de las empresas públicas el poner a disposición de los ciudadanos y empresas su información.

Adicionalmente hay otras consecuencias más directas que justifican la reutilización. La primera se refiere a la generación de transparencia de cara al ciudadano y a las empresas. En España, al igual que en otros muchos países, el ciudadano percibe demasiado a menudo a las distintas entidades públicas como entes cerrados, opacos donde, sin duda, los empleados públicos realizan su labor, pero esta no es visible. Sí se perciben algunos de los resultados pero no se es consciente del trabajo y el esfuerzo que hay detrás de todo ello. Si a esto sumamos los casos de corrupción que se han detectado en determinados ámbitos de la Administración, nos encontramos con un escenario en el cual la percepción por parte del ciudadano dista mucho de ser la deseada desde esas mismas administraciones.

Sin duda, existen distintas formas de mejorar la relación con el ciudadano pero todas inevitablemente pasan por un acercamiento hacia él ofreciendo servicios y generando transparencia. En este aspecto la administración electrónica está haciendo una inmensa labor y, como complemento a esta administración electrónica, debe estar la apertura de la información hacia ciudadanos o empresas. La información que se libera no tiene por qué ser la información final sino que pueden ser distintos estudios, mapas, artículos, informes, etc. que han servido

como base para la prestación de los servicios por parte de la entidad pública, cualquier tipo de documentación generada en el ejercicio de sus funciones. El ofrecer toda esa información traslada a los destinatarios de la misma una visión de cuál es el trabajo desempeñado para la prestación de los servicios. De esta forma, la imagen de la entidad pública hacia los destinatarios de la información mejora tanto en lo que se refiere a la calidad del trabajo desempeñado como a la transparencia en la gestión de la Administración Pública.

Por otro lado, el proceso de acercamiento, de liberación de datos, tiene distintos pasos. Existen distintas formas de ofrecer esta información a sus destinatarios como se verá más adelante, pero se elija una forma u otra, cualquiera de ellas implica una consolidación de los activos de la entidad. En cierto modo esto implica un proceso metodológico compuesto por una serie de etapas que, sin ser obligatorias, de forma individual, son un beneficio en sí mismo. En primer lugar trae consigo una homogeneización a través de la identificación de los activos, su ubicación y estado. Una agregación a través de la integración de todos estos activos para su uso. Esta agregación está directamente relacionada con un proceso de depuración detectando duplicidades. Todo ello implica una mayor facilidad para el mantenimiento de la información.

Adicionalmente, el Open data trae unos beneficios intrínsecos para la empresa que se han ido esbozando a lo largo del documento. El primero de ellos es que supone una base para la creación de servicios y productos innovadores —de hecho ha traído consigo la creación de un nuevo tipo de empresas, las llamadas empresas infomediarias—, puede servir como base de conocimiento para la toma de decisiones en la empresa y, en definitiva, es la base para la generación de riqueza en la empresa, en particular, y en el conjunto de toda la economía en general.

Analizándolo a más bajo nivel, como comentaba anteriormente, otro de los beneficios, sobre todo para estas empresas infomediarias, es el de la reducción de costes; si en nuestros negocios tenemos acceso a la materia prima de forma gratuita o a un coste residual⁵ nos

⁵De acuerdo a la trasposición de la directiva europea a la Ley 37/2007, las entidades públicas pueden aplicar una tasa o precio público por la información que ofrecen, tasa que debiera cubrir al menos los costes del servicio o la actividad pero desde la *Guía Aporta* se recomienda “facilitar la información en medios electrónicos de forma gratuita para fomentar la reutilización”.

situamos en una posición ventajosa de cara a garantizar la viabilidad de la empresa ya que se reducen enormemente nuestros costes de operaciones.

Por último hay que citar que en otro caso la falta de información accesible del sector público puede **provocar una ventaja competitiva** de las empresas extranjeras ante las locales, que pueden recurrir a su propia experiencia de la situación y posicionarse de forma ventajosa perjudicando sobre todo a las pymes.

En definitiva, a través de la reutilización de la información las empresas van a conseguir incentivar la economía, mediante el uso del activo económico que nos aporta la Administración. Todo ello desarrollado en la Sociedad de la Información.

3_Cómo reutilizar



3_1 Antecedentes históricos del Open data y la reutilización

Realmente se desconoce cuándo surge el Open data y la reutilización, a lo largo de la historia hemos podido observar ejemplos que siguen la misma filosofía.

El físico americano Joseph Henry nació en Albany en 1797, criado en una familia de pocos recursos desde muy joven se vio en la necesidad de trabajar con el fin de ayudar a su familia. Por diversas circunstancias volvió a estudiar y terminó trabajando en el campo de la electricidad especializándose en el magnetismo. A partir de un descubrimiento realizado por un zapatero inglés Williams Sturgeon en 1825, el cual enrollando alambre de cobre en una barra de hierro con forma de herradura y haciendo pasar una corriente eléctrica por el mismo conseguía que el hierro actuara como un imán, Henry consiguió, a través de diversas mejoras, convertir el “juguete” de Sturgeon en un potente electroimán. A partir de la corriente proporcionada por una batería eléctrica estándar, consiguió levantar más de 300 kg de hierro en un primer intento y, posteriormente, hasta una tonelada.

No obstante, Joseph Henry era un científico a la vieja usanza y creía en la filosofía del Open data, aunque no la conociera con ese nombre. Las leyes de la ciencia y cualquier beneficio obtenido a partir de ellas es patrimonio de la humanidad y deben ser compartidas. De esta forma, el electroimán y otros inventos posteriores basados en el mismo principio, como el relé e incluso una primera versión de un motor eléctrico, los puso a disposición de quien quisiera utilizarlos sin realizar ninguna patente. El telégrafo de Morse (con el que John Henry colaboró), el teléfono de Alexander Graham Bell y, en realidad, se podría decir que cualquier aparato que utilice un motor eléctrico, están basados en los conocimientos, inventos y experiencias compartidas por John Henry. ¿Alguien es capaz de imaginar cómo podrían haber cambiado la evolución científica en general y las telecomunicaciones en particular en el mundo si John Henry hubiera sido proteccionista con sus descubrimientos?

En agosto de 2010, la *BBC History Magazine* publica un artículo⁶ en el que se relata cómo Florence Nightingale, enfermera británica durante la Guerra de Crimea (1854-1856), utilizó datos estadísticos proporcionados por el Ministerio de Defensa Británico para generar un gráfico que representaba los datos de las bajas británicas y sus causas. En este gráfico se pudo comprobar que el porcentaje de muertes en batalla era siete veces menor que las muertes en los hospitales. Con esta base, Florence Nightingale consiguió los apoyos para mejorar las condiciones higiénicas, alimentarias y de equipamiento de los soldados reduciendo de esta forma la tasa de fallecidos en un 75 %. Este primer ejemplo de reutilización de datos pura y dura es obvio que no tiene fines empresariales, pero nos sirve como modelo para poder comprobar que la reutilización de la información no es un concepto nuevo.

No obstante, Open data tal y como lo conocemos es un concepto reciente, aunque científicos como John Henry hubieran seguido su filosofía tiempo atrás. Tal y como se comentaba en la introducción, surge en entornos académicos. Durante la Guerra Fría cualquier tipo de intercambio de información académica entre Este y Oeste estaba totalmente interrumpida. Pero en 1957 se celebró el Año Internacional de la Geofísica, evento en el que participaron prácticamente todos los países del mundo. Se celebraron numerosos actos de gran importancia científica, pero para nuestro interés cabe destacar la creación por parte del *International Council for Science* de un conjunto de centros de datos a nivel mundial que permitía clasificar, catalogar, almacenar y distribuir información entre los distintos centros y los investigadores.

Respecto a la Administración Pública, todavía tuvo que pasar bastante tiempo para que los políticos se dieran cuenta del potencial de la información del sector público, su posible reutilización y el valor que podía generar. No fue hasta 1998 cuando desde la Comunidad Europea se redactó el *Libro Verde sobre la información del Sector Público en la Sociedad de la Información* (1998). Este documento establece la situación en ese momento a nivel europeo, recoge ejemplos de distintos Estados Miembros y enumera una lista de cuestiones a tener en

⁶ www.historyextra.com/issue/august-2010

cuenta de cara a la redacción de una normativa en materia de reutilización. Pero, por encima de todo, describe el valor potencial de la información del sector público y la necesidad de su explotación y reutilización de cara a desarrollar la industria de los contenidos digitales basada en información pública.

Con el fin de impulsar su aplicación, se publicó en el año 2000 el estudio *Commercial exploitation of Europe's public sector information* de la Comisión Europea (Informe Pira), que refleja la importancia económica de este recurso. En él se hace una primera aproximación del valor de la información del sector público, señalando que representaba en el año 1999 casi el 1 % del producto nacional bruto en la Unión Europea.

Este estudio desencadenó la redacción e implantación de la Directiva 2003/98/CE con la finalidad de explotar el potencial de la información pública en todos los Estados Miembros a través de la regulación del acceso de los ciudadanos a la información administrativa y la forma en la que las entidades públicas podían y debían adoptar medidas para difundir esa información así como la forma en la que las entidades privadas pueden comercializar la información. El objetivo final era la creación de un marco homogéneo que garantizara unas condiciones de equidad para todos los agentes reutilizadores en todos los países de los Estados Miembros. Esta directiva fue finalmente traspuesta al ordenamiento jurídico español en la Ley 37/2007, del 16 de noviembre, sobre reutilización de la información del sector público que no es de obligado cumplimiento.

Tras la ejecución del Proyecto Aporta, que trataremos más adelante, se han visto las carencias de la trasposición de la directiva y como consecuencia se ha aprobado con fecha 8 de noviembre de 2011 el Real Decreto 1495/2011 que desarrolla la Ley 37/2007, incorporando importantes mejoras que serán descritas más adelante.

3_2 Qué información se puede reutilizar

De acuerdo a la Ley 37/2007, toda información generada por la Administración Pública es reutilizable, siempre y cuando no se encuentre excluida de forma expresa tal

y como se indica en su artículo 3. Adicionalmente, el Real Decreto 1495/2011 establece que las entidades del sector público estatal deberán, a través de su sede electrónica, informar sobre los documentos reutilizables elaborados o custodiados por ellas.

De acuerdo a la citada ley, son las Administraciones Públicas las que tienen la decisión final sobre qué información es susceptible de ser reutilizada. A este respecto, la publicación del Real Decreto ha supuesto un paso adelante, ya que en el Capítulo II del artículo 2 establece que cualquier entidad pública con información reutilizable deberá autorizar la reutilización de los documentos elaborados o custodiados por ellos e incluidos en el ámbito de aplicación del Real Decreto.

Existe una limitación establecida por ese ámbito de aplicación que está motivada por la necesidad de proteger los intereses públicos o de terceros liberando información que los pueda dañar. Aún así, cuando no se realice dicha liberación, esta decisión debe estar debidamente motivada y documentada, aunque la experiencia nos dicta que muchas veces esa “motivación” depende más de criterios arbitrarios. Por ello la aprobación del nuevo Real Decreto ha dado la vuelta a la situación estableciendo que **por defecto toda la información del sector público es reutilizable sin necesidad de solicitar ningún tipo de autorización**. Claro está, siempre y cuando esta administración no tenga la obligación de no liberar la información y proteger los intereses públicos y privados.

Los casos en los cuales la información **no** puede ser liberada son los siguientes:

1. Información excluida por intereses públicos

El objetivo que se persigue con esta exclusión es preservar el funcionamiento de la Administración Pública y que no se vean afectadas las decisiones que desde la misma se puedan realizar. Algunos de los casos a los que nos referimos son:

1. Actuaciones políticas: se refiere a los expedientes que contengan información sobre actuaciones del Gobierno del Estado o de las comunidades autónomas en el ejercicio de sus competencias constitucionales

no sujetas a derecho administrativo. Podríamos citar por ejemplo:

- Defensa nacional y seguridad del Estado: resulta obvio que la Administración Pública no puede publicar información que afecte a la seguridad o defensa nacional, a la seguridad interior o a la lucha antiterrorista.
 - Política monetaria: toda aquella documentación que refleje las actuaciones o decisiones tomadas por la Administración en el ejercicio de sus funciones que tengan como objetivo estudiar, dirigir o controlar el coste del dinero con el fin de preservar la estabilidad económica es información que en ningún caso debe liberarse, ya que podría directamente afectar a los mercados.
 - Investigación de delitos: siempre y cuando con ello estemos preservando la protección de los derechos y libertades de terceros o incluso las investigaciones en curso.
2. Información que incluye materias clasificadas por ejemplo por el Consejo de Ministros, que pudieran incluirse en alguno de los puntos descritos anteriormente (por ejemplo, la defensa del Estado).
3. Información registral: en este caso (por ejemplo, el Registro Civil) ya existe una ley que regula qué información recogida en estos organismos tiene carácter público y cuál tiene carácter privado.

Estas dos últimas son otras categorías de información de acceso limitado, regulado el mismo a través de normas sectoriales, y que, en cierto modo, están relacionadas con las exclusiones anteriormente mencionadas.

2. Información excluida por intereses de terceros

Las Administraciones Públicas poseen numerosas fuentes de información que incluyen bien información relativa a terceros, bien información elaborada por terceros. Estas terceras partes tienen unos derechos que deben ser respetados, por lo cual la directiva establece normas que, a través de la limitación del acceso público a esta información, los garantizan.

Con carácter general, la Ley 37/2007 no se aplicará respecto a documentos para cuyo acceso se requiera ser titular de un derecho o interés legítimo. Estos se podrían clasificar de la siguiente forma:

1. Información que contenga datos de carácter personal.

El acceso a la información que contenga datos personales está limitado a sus titulares que, de acuerdo a la Ley de Protección de Datos, tienen derecho a rectificar, cancelar u oponerse a la misma dentro de los términos establecidos en la propia ley. Está claro que se puede acceder a información catastral a través de la Oficina Virtual de Catastro, permitiéndose incluso descargas masivas, pero lo que nunca se nos va a facilitar son los datos personales recogidos en sus bases de datos. No obstante, una persona física mediante el uso de un certificado digital podrá tener acceso a sus datos, actualizarlos, etc. Otros ejemplos de información que entrarían dentro de esta categoría son los datos sanitarios, datos personales referentes a la intimidad, datos electorales, etc.

2. Información cuyo acceso o reutilización pueda afectar intereses de terceros.

Este punto está particularmente orientado a la protección del secreto comercial o industrial (patentes, marcas, modelos industriales, etc.) en dos vías. Por un lado se refiere a la información que las empresas tienen que entregar a la Administración Pública, debiendo esta preservarla protegiendo a esas empresas y garantizando la libre competencia. Puede ocurrir el caso en que parte de esa información deba ser hecha pública por diversos motivos, pero en ningún caso puede ser reutilizable.

La segunda vía se refiere ya no al acceso sino a la reutilización de la información que pueda afectar a intereses de terceros. Es importante citar que aquí se incluyen las obras sujetas a los derechos de propiedad intelectual, derechos de autor, derechos de uso y reproducción y todos los derechos conexos.

La normativa aplicable tal y como se indica en la Ley 37/2007 recoge que esta no afecta a los derechos de propiedad intelectual o industrial de terceros y

excluye de su aplicación aquella información sobre la que existan dichos derechos de propiedad intelectual o industrial.

En el caso de los derechos de propiedad intelectual de las propias administraciones y organismos del sector público, aplica la misma normativa. Los garantes de esos derechos son esas administraciones y organismos y, por lo tanto, no deben ser reutilizados. No obstante la misma ley respecto a las Administraciones Públicas recomienda que, aún ejerciendo sus derechos de protección intelectual, faciliten la reutilización de la información mediante la adopción de licencias libres.

3. Información cuya reutilización pueda afectar a intereses colectivos.

Este punto en concreto está especificado dentro de la Ley 37/2007 y se aplica a los contenidos creados o custodiados por las siguientes entidades:

- 1. Entidades de radiodifusión:** la información conservada por las entidades que gestionen los servicios esenciales de radiodifusión sonora y televisiva y sus filiales.
- 2. Entidades de educación e investigación:** la información conservada por instituciones educativas y de investigación, tales como centros escolares, universidades, archivos, bibliotecas, centros de investigación, etc.
- 3. Entidades culturales:** la información conservada por instituciones culturales tales como museos, bibliotecas, archivos históricos, orquestas, óperas, ballets y teatros.

No obstante hay que indicar que desde la Comunidad Europea se está tramitando una nueva directiva (en el momento de la publicación de este libro ha sido sometida a consulta pública y está a la espera de aprobación) que revisa la actual Directiva 2003/98 y modifica estos últimos supuestos. De tal manera que amplía los supuestos de información susceptible de ser reutilizada a información almacenada en museos, archivos y bibliotecas y que no esté sujeta a derechos o condiciones que lo impidan.

Pero lo que de verdad nos interesa no es lo que no se puede reutilizar sino lo que **sí** se puede reutilizar. Ya en el Informe PIRA (*Commercial exploitation of Europe's public sector information*) mencionando anteriormente, se hace un análisis del tipo de información pública que puede estar disponible para su reutilización. De esta forma se establecen nueve categorías que definen la tipología y nos pueden dar una idea del volumen y potencial de la información del sector público:

1. Información económica y comercial: dentro de este epígrafe encontraremos todo tipo de información financiera, información de las empresas y estadísticas económicas referidas a la inflación o desempleo. Podemos citar por ejemplo la encuesta de población activa realizada por el INE.
2. Información geográfica: datos hidrográficos, información sobre la utilización de tierras, datos de calidad ambiental, tales como mapas y datos meteorológicos. En esta categoría disponemos de boletines hidrológicos, sistemas de información de aguas subterráneas, diversos datos de las distintas confederaciones hidrográficas de toda la geografía nacional, etc. Es destacable el esfuerzo realizado para la publicación de una amplísima base de datos de mapas del Instituto Geográfico Nacional.
3. Información agrícola y pesquera: información sobre cosechas y datos sobre utilización de fincas, ingresos de las explotaciones y utilización de los recursos, producción pesquera e información sobre las piscifactorías así como diversa información proporcionada por el Ministerio de Agricultura como, por ejemplo, información puntual sobre la evolución del mercado español, europeo y mundial de cultivo de cereales, permitiendo seleccionar entre diversos productos: conjunto de cereales, cebada, maíz, trigo blando y trigo duro.
4. Información social: información demográfica, investigación de actitudes, datos sobre la salud y enfermedades y datos censales. Es destacable de nuevo el esfuerzo del INE ofreciendo toda la información que produce

en formato digital, información tanto numérica como metodológica o descriptiva para mejorar la comprensión de los datos.

5. Información jurídica y legal: información sobre normativa y jurisprudencia, legislación, indicadores de criminalidad y condenas, etc. Cabe destacar en esta categoría el BOE, ofreciendo sus publicaciones a través de su página web, de hecho han limitado su producción en papel para centrarse en la publicación en formato digital. Adicionalmente, el CENDOJ ofrece un amplio fondo de jurisprudencia debidamente tratado y adaptado para ajustarse a la normativa de la Ley 37/2007 referente al tratamiento de datos privados.
6. Información científica: investigación producida en las universidades, investigación financiada por fondos públicos, patentes e investigación dentro de los departamentos gubernamentales. Como ejemplo destacable tenemos a TESEO, la base de datos de tesis doctorales españolas, leídas y publicadas desde 1976.
7. Información cultural (parte de ella será reutilizable tras la aprobación de la revisión de la Directiva 2003/83): materiales en los museos, galerías de arte, información turística, recursos en las bibliotecas, etc. En esta tipología, existen gran cantidad de colecciones como podría ser la Galería de la Colección del Museo Nacional del Prado. No obstante hay que citar que esta información debe ser tratada con cuidado en lo que se refiere a su uso respecto a los derechos de autor y explotación.
8. Información política: comunicados de prensa, actas de reuniones de gobierno, libros blancos y verdes, estrategias, etc. Cabe destacar el esfuerzo realizado desde las Cortes Generales en aras de la transparencia publicando una gran fototeca, sesiones del pleno en vídeo, discursos del presidente del gobierno, notas de prensa, intervenciones, etc.
9. Información relativa a la gestión empresarial: dentro de esta categoría se incluye la información y publicaciones relativas a la creación de empresas, a la captación y a la formación de los trabajadores, a la gestión medioambiental y la calidad y la excelencia empresarial.

No obstante, desde el punto de vista práctico algunos de estos segmentos no han sido explotados mientras que otros, cuya categoría de información todavía no ha sido detallada en la anterior clasificación, sí han sido compartidos por Administraciones Públicas, como es el caso de la información relativa a transportes (información del tráfico, estado de carreteras, precio de carburantes, etc.).



4_La reutilización en España



4_1 Reutilización en la empresa pública

Hasta finales del año 2011, como comentábamos, la transposición de la directiva europea a la Ley 37/2007 no era de obligado cumplimiento, de ahí que, en general, no se habían dedicado todos los esfuerzos necesarios desde las administraciones en la implantación de las recomendaciones de la ley que, de hecho, era una gran desconocida. En España no fue hasta el desarrollo del Proyecto Aporta por parte de la entidad pública Red.es, adscrita a la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información dependiente del Ministerio de Industria, Turismo y Energía, cuando se elaboró una guía donde se describen claramente los procedimientos para poder generar una cultura de reutilización en una Administración Pública.

Dicha guía, basada en las directivas de la citada ley, proporciona a las Administraciones Públicas un documento de referencia, cubriendo todos los puntos clave para facilitar la reutilización de la información del sector público y para poder establecer los cambios procedimentales e incluso organizativos con el objetivo de crear una cultura de reutilización dentro de esa administración. Sin duda, supuso un impulso enorme en esta línea que ha sido alabado a nivel europeo, de hecho ha sido referenciado en varias ocasiones desde la *European Public Sector Information Platform*, plataforma auspiciada por la Unión Europea para impulsar la reutilización de la información del sector público a nivel europeo.

Pero es importante ir por partes, este proyecto es el fruto del esfuerzo de un equipo de trabajo dirigido desde la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información. Para desarrollar el programa se contó con la colaboración de diversas Administraciones Públicas que participaron activamente en los trabajos que permitieron alcanzar los resultados finales.

Desde sus orígenes, se fue consciente de la necesidad de involucrar a las entidades productoras de la información así como a destinatarias de los resultados finales del proyecto de cara a conseguir cumplir los objetivos fijados en el mismo. Entidades como el Boletín Oficial del Estado,

la Dirección General del Catastro (Ministerio de Economía y Hacienda), el Instituto Nacional de Estadística (INE), el Ministerio de Vivienda, la Agencia Estatal de Meteorología (AEMET), el Centro Nacional de Información Geográfica (CNIG), etc. trasladaron sus experiencias permitiendo definir el estado del arte, el manual de reutilización y, posteriormente, la guía de reutilización que era, después de todo, el objetivo final de la iniciativa.

No obstante, antes de entrar en más detalle en el Proyecto Aporta es necesario conocer cuál era la situación en España en lo que se refiere a la reutilización de la información dentro de la Administración Pública. Ya en el año 2006, desde la Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información, se impulsó la elaboración de un estudio que recogiera cuál era el estado de la reutilización de datos del sector público en España en ese momento⁷.

El documento analiza cuál es la situación desde el punto de vista de la reutilización recorriendo todos los posibles escenarios: el usuario final, la Administración Pública y las empresas infomediarias sobre las que hablaremos más adelante.

El análisis de la situación en un total de 176 entidades públicas permite observar que ya en ese momento existe gran cantidad de información que es puesta a disposición de los agentes que quieran reutilizarla. No obstante esta información accesible, pese a estar en formato digital no está en un formato fácilmente reutilizable, siendo el formato pdf el más común. Si analizamos el tipo de información que en este momento está disponible se trata de información final, es decir, información que ha sido elaborada expresamente para su liberación: estadísticas, mapas, informes, normativas, etc. Obviándose cualquier tipo de documentación que haya sido utilizada para la generación de cada uno de esos productos.

La mayor parte de las publicaciones disponibles para su reutilización eran de libre distribución (sin necesidad de abonar tasa o precio público alguno), no obstante existían ciertas revistas, catálogos e información que tenían un coste asociado a través de una tasa cuyo importe era definido por el propio organismo. Dependiendo del organismo y del tipo de información esta tasa podía oscilar entre 1 y 60 €.

⁷ Everis (2006), *Estudio sobre la situación en España de la Reutilización de la Información del Sector Público*. Madrid.

Pese a la existencia de dicha información publicada, es importante resaltar que el grueso de la información reutilizable no se encontraba en formato digital, siendo el acceso a la misma bajo demanda, a través de catálogos de centros de documentación y bibliotecas. Para la información publicada generalmente no existía limitación en su acceso en Internet excepto para ciertos casos en los que se requiere una suscripción para el envío de la misma.

En las fechas en las que el informe fue redactado se establecía que el 75 % de la información pública elaborada por la Administración en el ejercicio de sus funciones era susceptible de ser reutilizada. Aunque la falta de orden y criterio en la oferta de información reutilizable era un hecho palpable, ya había un conjunto de entidades que entendían la importancia de la reutilización y la necesidad de fomentar esta cultura dirigiendo sus esfuerzos en esta línea.

Entre estas entidades se encontraban el Boletín Oficial del Estado, el Centro Nacional de Información Geográfica, la Dirección General del Catastro, el Instituto Nacional de Meteorología o el Ministerio de Industria. Todas estas entidades presentaban la característica de tener gran parte de su información informatizada. Esto facilitaba enormemente la gestión y puesta a disposición de esta información hacia empresas infomediarias y ciudadanos en general.

No obstante, hay que tener en cuenta que hablamos de los orígenes del Open data en España, así, por citar un ejemplo, la Dirección General del Catastro disponía en sus bases de datos en 2006 de 124 millones de datos de información literal: datos de titulares, datos urbanos, de construcción y de parcelas rústicas. Adicionalmente, se incluían en la base de datos 9.200 mapas cartográficos digitalizados, 6.200 sobre cartografía rústica y 3.000 de cartografía urbana. La oferta de estos datos, dado que hasta el 2007 no se traspuso la directiva europea a la legislación nacional, era limitada. Los casos en los cuales podías disponer de esta información estaban circunscritos a aquellos supuestos permitidos por las leyes vigentes. Este hecho se mantiene hoy en día pero la evolución en la jurisprudencia y la tecnología ha hecho que más datos y con mayor facilidad de acceso estén actualmente disponibles.

Otro caso digno de mención es el Instituto Nacional de Meteorología que ofrecía diversos tipos de servicios a través de su página web. Por la naturaleza del organismo es obvio indicar que la naturaleza de la información a publicar es de carácter medioambiental, tanto meteorológica como climatológica. En este caso el acceso a las publicaciones se realizaba a través de un catálogo que te permitía la adquisición de las guías, informes, estudios, etc., todo ello accesible mediante el pago de una tasa. Respecto a la información accesible a través de Internet, sobre todo en lo que se refiere a las prestaciones sobre información meteorológica, predicciones, aviso de fenómenos climatológicos potencialmente peligrosos, etc., estaba disponible para cualquier ciudadano y/o empresa de forma libre con la única condición de indicar el origen de los datos y, por supuesto, no desvirtualizar su contenido.

Pero en estos ejemplos todavía hay muchos puntos que no se ajustaban a la directiva europea vigente (que no había sido traspuesta al marco normativo español). Sin embargo el informe establece un conjunto de buenas prácticas que, adelantándose a la normativa española, ya se ajustaba a la directiva europea.

Cabe destacar el Centro de Información Judicial (CENDOJ), organismo a través del cual se realiza la puesta a disposición de las sentencias del Tribunal Supremo, Tribunales Superiores de Justicia y otras resoluciones judiciales. Todo ello con las debidas garantías de igualdad, protección de datos, etc. dado que las sentencias son anonimizadas para proteger la intimidad de las personas citadas en las mismas. En el 2006 ya se estableció un sistema de tasas a partir del cual las empresas y particulares tenían acceso previo pago a la totalidad o a parte de las sentencias. Se establecen varios modos de acceso a toda la información facilitando su descarga y tratamiento, aunque el formato entregado no era demasiado usable (pdf) frente a los estándares que se defienden en este libro. Hay que indicar que aunque en el 2006 estaba permitido, el CENDOJ mantenía un conjunto de acuerdos de publicación exclusivos con unas empresas (a través de concursos públicos, pero exclusivos al fin y al cabo) para el conjunto de sentencias que publicaba. Hoy en día de

acuerdo a la nueva directiva de la Unión Europea esto no estaría permitido y de hecho se ha producido una revisión por parte de las autoridades comunitarias para asegurar su desaparición.

No obstante, volviendo a la situación actual, vamos a hablar de iniciativas más recientes. Más concretamente del Proyecto Aporta.

4_2 Proyecto Aporta

El Proyecto Aporta marcó un antes y un después en lo que es la reutilización de datos del sector público en España. Más allá de la trasposición de la Ley 37/2007, que al fin y al cabo era una obligación por parte de las instituciones españolas, el Proyecto Aporta fue la primera iniciativa que de manera activa promovía la reutilización a nivel institucional. Esta iniciativa supuso un impulso para la creación de una cultura en la Administración Pública, el primer impulso real.

El objetivo de este proyecto era la creación de una guía de reutilización que ayudara a la Administración Pública a realizar todos los cambios necesarios para poder ofertar su información para su reutilización.

El primer paso que se acometió en el marco del proyecto fue el análisis de la situación actual. Durante la ejecución de esta tarea se identificaron las iniciativas más relevantes que fueron llevadas a cabo tanto a nivel nacional como internacional. El objetivo era claro, por un lado, se quería identificar a los principales organismos generadores de información para, a partir de las experiencias de estas entidades, identificar las necesidades de las empresas susceptibles de reutilizar esta información, las empresas infomediarias de las que hablaremos posteriormente. A partir de todas estas experiencias se concretaron las principales características de todas y cada una de ellas tratando de determinar las acciones a llevar a cabo para favorecer el desarrollo del sector empresarial relacionado con la reutilización de información del sector público.

Pero... ¿cómo se consolidaron las conclusiones obtenidas en el análisis del estado del arte? En primer lugar a través de la extracción de un conjunto de conclusiones

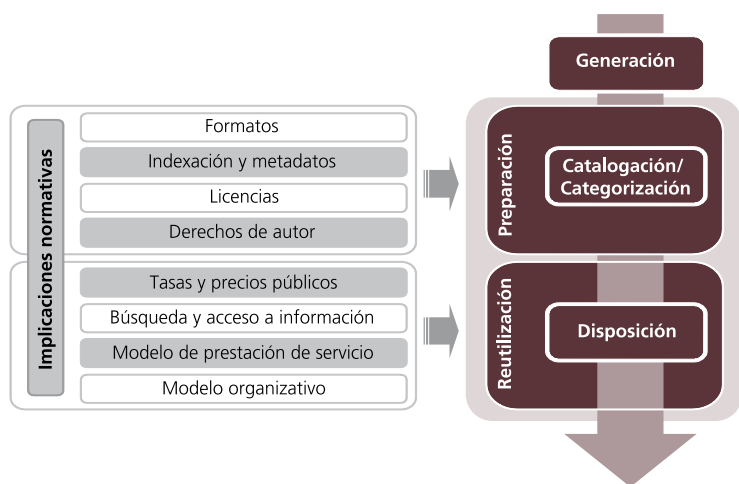
que podrían aplicarse en trabajos posteriores. Estas estaban organizadas en un conjunto de áreas orientadas a:

- Crear una cultura de reutilización mediante acciones de concienciación e implicación de las administraciones y sus gestores a todos los niveles, definiendo incluso relaciones entre el sector público y privado de cara a fomentar la reutilización.
- Llevar a cabo un impulso dentro de las organizaciones públicas en aras de crear y/o modificar los procesos de negocio para establecer un modelo que permitiera la generación y reutilización de información pública.
- Establecer y definir los tipos de licencia a los cuales las distintas organizaciones podían acogerse para poner a disposición de empresas y ciudadanos la información a reutilizar. Aquí hay que decir que el Proyecto Aporta fue pionero ya que, pese a ofrecer diversas modalidades, recomendaba siempre licencias abiertas que a la larga han sido incluidas dentro de normativas europeas posteriores.
- Definir un modelo de costes que permitiera establecer las tasas o precios públicos bajo los que se podía ofrecer la información.
- Y por último y no menos importante, concretar los formatos y tecnología necesaria para poner a disposición de las empresas y ciudadanos la información de forma que esta pueda ser reutilizada adecuadamente.

Todas ellas tomaron forma en un *Manual de reutilización*, documento de trabajo del Proyecto Aporta que sirvió como base para la definición de las acciones y para la obtención de información de distintos agentes, aunque no fue finalmente consolidado y publicado como un documento resultado del proyecto, y en el que se define la cadena de valor de la información a reutilizar. Se puede observar en la Figura 4.1 cómo se tienen identificadas tres tareas:

1. La generación de la información.
2. La preparación a través de una catalogación o categorización que facilite su gestión con herramientas informáticas.
3. La reutilización a través de la puesta a disposición de la información para todo aquél que desee reutilizarla.

Figura 4.1. Cadena de valor de la información.



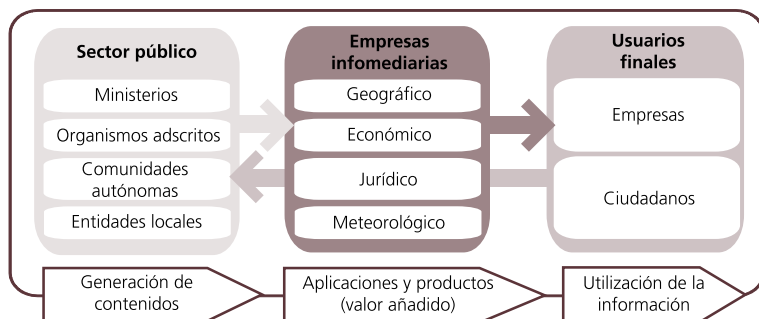
Fuente: Secretaría de Estado de Administraciones Públicas del Ministerio de Hacienda y Administraciones Públicas y Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información del Ministerio de Industria, Energía y Turismo.

Para cada una de las fases se identifican aquellos puntos a tener en cuenta en cada una de ellas, de esta forma de cara a la *preparación* es necesario tener en cuenta los derechos a los cuales la información está sujeta, las licencias aplicables para cada caso o incluso los metadatos que acompañarán a la información de cara a facilitar su gestión a través de medios informáticos.

Del mismo modo, la *puesta a disposición* de la información reutilizable deberá tener en cuenta las tasas o precios públicos aplicables, la forma de buscar y acceder a la información o incluso el modelo de prestación de servicio y el modelo organizativo en la Administración Pública correspondiente.

A través de cada una de las fases de la cadena de valor, el *Manual de reutilización* identifica la tipología de información reutilizable, los procedimientos de catalogación y categorización, los tipos de licencias aplicables, etc. De cara a la puesta a disposición establece un conjunto de

Figura 4.2. Agentes que intervienen en el proceso de reutilización de la información.



Fuente: Secretaría de Estado de Administraciones Públicas del Ministerio de Hacienda y Administraciones Públicas y Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información del Ministerio de Industria, Energía y Turismo.

modelos de tasas y precios públicos que actualmente se han quedado desfasados dado que desde la Unión Europea cada vez se impulsa más la necesidad de ofrecer la información reutilizable de forma gratuita.

Desde el punto de vista técnico se identifica un modelo de metadatos que recoge las tipologías de información, un modelo de licencia, etc.

Por último, identifica una propuesta de modelo organizativo que establece un modelo orgánico y funcional junto con los procedimientos operativos necesarios para su gestión y la formación y capacitación necesaria para poder acometer todas estas tareas. No voy a entrar en más detalle dado que lo que nos interesa es lo que se puede hacer con la información y no la forma en la cual las entidades públicas la pueden reutilizar.

El paso final del proyecto fue, a partir del *Manual de reutilización*, la construcción de la *Guía Aporta sobre reutilización de la información del sector público* que de forma más amigable sirviera de guía para las Administraciones Públicas. En ella se identifican los agentes que intervienen en el proceso de reutilización.

Como se puede observar (Figura 4.2) va un paso más allá de la cadena de valor de la información reutilizable dentro de la empresa pública, incluyendo a las empresas infomediarias y a los ciudadanos y empresas como consumidores finales de esta información.

La guía define claramente qué información es reutilizable, establece la información que no puede ser directamente reutilizable (por ejemplo aquella que contenga datos de carácter personal) y define aquellos procedimientos que, de acuerdo a la Ley Orgánica de Protección de Datos, sirven para adaptar la información (sin tergiversar su contenido) y de esta forma hacerla publicable y accesible por cualquier empresa o ciudadano.

Aunque en la Ley 37/2007 se establece que no es necesario realizar un esfuerzo desproporcionado para poner a disposición la información reutilizable, desde la guía se recomienda que las entidades públicas ordenen, cataloguen, clasifiquen e incluso indexen la información que tiene que ser reutilizada, ya que únicamente de esta forma puede ser usada directamente mediante medios telemáticos, esta misma tendencia se ha mantenido en toda la normativa nacional y europea tramitada con posterioridad a la citada ley.

A lo largo de toda la guía se establecen recomendaciones, por ejemplo respecto a la puesta a disposición de la información en los términos que comentamos en el párrafo anterior. Como mínimo, se debe poner a disposición la información, no obstante se recomienda la adecuada ordenación de la misma, así como la generación de listados e índices que ayuden a su gestión. Pero de cara a fomentar la reutilización de forma clara lo más adecuado sería definir metadatos que permitieran catalogar la información e incluso digitalizar aquella que no estuviera informatizada.

Respecto a las modalidades en las cuales la información es ofrecida se presentan tres opciones:

- Sin que esté sujeta a ningún tipo de condición.
- Condicionando la reutilización a un conjunto de licencias tipo (gratuitas o no).
- Previa solicitud.

En la primera de las modalidades, simplemente la información que se va a descargar tiene asociado un

aviso legal que indica las condiciones a las que está sujeta la reutilización.

Respecto a las licencias tipo, las trataremos con más detalle en un capítulo específico, pero aun así, independiente del tipo de licencia, todas deben incluir, de acuerdo a la Ley 37/2007 la siguiente información:

- Finalidad concreta para la que se concede la reutilización: comercial o no comercial.
- Duración de la licencia.
- Obligaciones del beneficiario.
- Obligaciones del organismo concedente.
- Responsabilidades de uso de la licencia.
- Modalidades financieras: gratuita o contraprestación económica aplicable.

Por último, respecto a la tercera opción, modalidad previa solicitud, lo que incorpora la directiva y que se refleja en la guía es que existe un procedimiento claramente establecido en el cual se estipulan, por ejemplo, los plazos de respuesta en los que un organismo debe gestionar esta solicitud, las modalidades de resolución e incluso los costes asociados.

Aquí de nuevo la *Guía Aporta* recomienda la puesta a disposición de la información reutilizable sin ningún tipo de prestación económica, de forma gratuita.

Hemos recorrido el contenido de la guía de reutilización y las recomendaciones que incluye. No hay que olvidar que es un documento orientado a las Administraciones Públicas de cara a fomentar la reutilización, pero no por ello hay que quitarle importancia ya que nos ofrece una información vital sobre los distintos escenarios que podemos encontrarnos en las Administraciones Públicas de cara a disponer de información reutilizable para nuestras empresas. Aun teniendo en cuenta que está basada en los contenidos de la Ley 37/2007 (fue redactada antes de la aprobación de la jurisprudencia más reciente), lo cierto es que muchas Administraciones Públicas han recibido formación respecto a esta guía y la Ley 37/2007 y han seguido sus directrices.

Más adelante, con la aprobación del Real Decreto 1495/2011, desde el portal datos.gob.es se publicó la *Guía de aplicación del Real Decreto 1495/2011 por el*

que se desarrolla la Ley 37/2007 sobre reutilización de la información del sector público⁸, para el ámbito del sector público estatal. Este documento, con el mismo objetivo que la *Guía Aporta*, actualiza la información ofrecida en la misma adaptándola al nuevo decreto y ofreciendo información adicional que complementa la primera. Dentro de este documento, es destacable la definición que se da de una estrategia de reutilización de información para el sector público identificando en su caso los responsables de dicha estrategia dentro de la entidad así como los perfiles con sus tareas y responsabilidades para acometerla.

Sin embargo, pese a que este documento no ha tenido la difusión de la Guía Aporta inicial no hay que ignorar su indudable importancia y el esfuerzo realizado desde las entidades públicas participantes, Ministerio de Industria, Energía y Comercio, y Ministerio de Hacienda y Administraciones Públicas, ya que actualmente es la guía de referencia para la aplicación del Real Decreto.

⁸ <http://datos.gob.es/datos/sites/default/files/PLANCISP-GRD-07.3.4.doc> (junio 2012).

5_El sector infomediario



Nuevos modelos empresariales han surgido desde la aparición de Internet. La red ha sido un impulsor de ideas que ha traído consigo una revolución en la forma en la que se llevan a cabo los negocios. Cuando desde un gobierno se impulsa una iniciativa como es el Open data, es inevitable que la oportunidad que supone despierte la imaginación de los emprendedores deseosos de crear nuevas empresas.

En 1999 John Hagel III y Marc Singer publicaron el libro *Net Worth*, en el cual ya analizan los cambios tecnológicos que trae consigo Internet respecto a su influencia en el mundo empresarial. En el libro se introducen nuevos modelos derivados de la aparición del comercio electrónico y es donde por primera vez se utiliza el término “Infomediario”, que viene a significar “intermediario de información”.

Podemos decir que el sector que basa sus modelos de negocio en la reutilización de la información (sea del sector público o no) es el sector infomediario.

Su operativa es teóricamente sencilla, se trata de obtener información del sector público, información que se puede obtener de forma gratuita o pagando un precio por ella, y **aportarle valor**, crear algún tipo de producto o servicio que es ofrecido a terceros obteniendo un beneficio.

5_1 Quién reutiliza

En el año 2011 en el marco del Proyecto Aporta se realizó un *Estudio de caracterización del Sector Infomediario* que ha sido actualizado en 2012⁹, en él se estima que el volumen de negocio de las empresas del sector infomediario en España estriba entre los 330 y 550 millones de euros (similar a una industria en auge como es la del videojuego), un volumen nada despreciable que sitúa a este sector en el punto de mira de cualquier emprendedor deseando desarrollar una nueva idea de negocio.

⁹ www.ontsi.red.es/ontsi/es/estudios-informes/estudio-de-caracterización-del-sector-infomediario-en-españ-edición-2012

De acuerdo al estudio MEPSIR, que ya ha sido citado anteriormente en este libro, se destacan los siguientes ámbitos de información de cara a la reutilización:

- Negocio/económico.
- Geográfico.
- Jurídico.
- Meteorológico.
- Socialdemográfico o estadístico.
- Transportes.

Actualmente cuando se crea una empresa en España, entre otros trámites, es necesario asignarle un código CNAE. El CNAE es la **Clasificación Nacional de Actividades Económicas**¹⁰, a través de esta clasificación se asigna un código a cada actividad económica que puede realizarse. Generalmente este código (que suele ser de 5 dígitos) se utiliza en muchos formularios e impresos, tanto oficiales como a nivel de empresa. Básicamente nos sirve para identificar la actividad empresarial que se está realizando. No obstante no existe un código asociado a la actividad infomediaria *per se*. Las empresas que, por ejemplo, extraen información del BOE para agregarle valor y venderla, usarán el CNAE correspondiente por ejemplo a “Servicios jurídicos”, de ahí que sea difícil poder identificar cuántas empresas están actualmente reutilizando información.

De hecho, en la revisión del *Estudio de caracterización del Sector Infomediario* publicada en 2012 se ha depurado el método de identificación de empresas infomediarias (ampliamente descrito en dicho documento), identificando a nivel nacional un censo de 150 empresas infomediarias que han participado en dicho estudio.

Por ello estas empresas, debido a que comparten experiencias, fuentes de información, necesidades y expectativas, han decidido constituirse en asociaciones de cara a poder aprovechar sinergias. Cabe destacar en España ASEDIE (Asociación Multisectorial de la Información)¹¹ que es una asociación compuesta por empresas cuya actividad resulta de la reutilización de la información del sector público, así como empresas públicas reutilizadoras.

Como organización, desde sus orígenes ha dirigido sus esfuerzos a fomentar la reutilización presentando propuestas de modificación a los distintos proyectos de

¹⁰ Esta clasificación se puede consultar en el Instituto Nacional de Estadística: www.ine.es/jaxi/menu.do?type=pcaxis&path=/t40/casrev&file=inebase

¹¹ www.asedie.es

ley de las distintas iniciativas que se han ido realizando desde la Directiva 2003/98/CE hasta la actual revisión de la directiva a nivel europeo en fase de aprobación en el 2012.

El objetivo de esta asociación pasa por la identificación de problemas y barreras existentes en la reutilización, por ello como miembros de esta organización se encuentran tanto empresas públicas proveedoras de información como empresas privadas usuarias de la misma. Como objetivo último, tal y como describen en varios documentos, la asociación persigue la mayoría de los principios que se han descrito en este libro:

1. Obtener una mejor información, más completa y exhaustiva que permita la aparición de nuevos productos y servicios de cara a los ciudadanos, la creación de empleo y la mejora de la economía.
2. Incrementar la seguridad jurídica y comercial a través de una transparencia de la información disponible del sector público y del establecimiento de un conjunto de principios que todos los socios deben seguir de cara a reutilizar la información del sector público.
3. Fomentar la investigación, tanto a nivel empresarial como del ciudadano: la gran aportación de valor en el conocimiento general que supondrá abrir al público y a las entidades infomediarias el acceso a la información disponible, tanto en el ámbito cultural como social.

No obstante, esta asociación, centrada en la reutilización de datos del sector público no es la única. Es importante citar también a ADIGITAL (Asociación Española de la Economía Digital)¹². Tal y como se indica en su página web, a esta asociación pertenecen las principales empresas relacionadas con la economía digital, el comercio electrónico y el marketing directo, relacional e interactivo. Con más de 500 miembros de distintas empresas pertenecientes a sectores de los segmentos anteriormente citados, persiguen el desarrollo de la economía digital formando, orientando y acompañando a empresas a través de distintos servicios de formación y consultoría. La reutilización de la información pública es una de las áreas que cubren realizando tareas de intermediación con las entidades públicas, por un lado, y, por otro, persiguiendo la aprobación de leyes

¹²www.adigital.org

que favorezcan la creación de empresas y la explotación de datos por parte de la mismas.

Existen otras asociaciones o fundaciones como la Fundación CTIC¹³ o la asociación AMETIC¹⁴ que, al igual que ADIGITAL, incluyen dentro de sus áreas de actuación el fomento de la reutilización de la información del sector público.

Pero para encontrar la asociación con más peso en la reutilización de la información en la administración tenemos que irnos a Europa. La PSI Alliance¹⁵ es una asociación centrada exclusivamente en el impulso al desarrollo de un mercado de negocio justo y equitativo basado en la información pública reutilizable. Esta alianza está compuesta por compañías privadas y asociaciones (ASEDIE es una de ellas) manteniendo contacto directo con la Administración Pública a nivel local, regional y europeo en aras de obtener los mayores beneficios para todos los agentes participantes: las empresas públicas, las empresas privadas y los consumidores finales.

Esta asociación participa de forma consultiva en la elaboración y/o revisión de la normativa persiguiendo cambios en la misma cuando es necesario. Busca, analiza y recompensa las buenas prácticas, de hecho en 2012 le ha concedido al Instituto Geográfico Nacional el “5 stars PSI award”¹⁶ a la excelencia en la provisión de datos públicos. Sus eventos y conferencias anuales son de obligado seguimiento para cualquier emprendedor que quiera introducirse en el sector infomediario.

5_2 Características de las empresas infomediarias

Estas características han sido brevemente descritas en el estudio sobre el sector infomediario descrito anteriormente. Desarrollaremos las mismas a través de este capítulo. Hay unos puntos que las caracterizan.

Tamaño de las empresas

No se puede establecer un patrón en cuanto al tamaño de las empresas tanto respecto al número de empleados como al volumen de negocio o de facturación. No

¹³ www.fundacionctic.org

¹⁴ www.ametic.es

¹⁵ www.psialliance.eu

¹⁶ <http://epsiplatform.eu/content/psi-alliance-5-star-award-instituto-geografico-nacional>

obstante cuando se trata de empresas pequeñas, el servicio que se apoya en la reutilización suele ser un servicio de la empresa adicional a otras actividades desarrolladas por la misma. Tenemos por ejemplo el caso de una pequeña empresa en la Comunidad Autónoma de Cantabria, Iteisa, que ya desde hace años está ofreciendo a través de la web www.elpreciodelagasolina.com un servicio a través del cual un ciudadano puede averiguar en su localidad cuáles son las estaciones de servicio que ofrecen el precio más barato para adquirir combustible. No obstante se trata de una empresa dedicada al desarrollo de software, aplicaciones web y comercio electrónico.

Origen de la información

Las empresas infomediarias españolas utilizan información proveniente mayoritariamente de organismos nacionales. Por citar algún ejemplo, existen empresas que en España obtienen del BOE (Boletín Oficial del Estado) o del CENDOJ (Centro de Documentación Judicial) la jurisprudencia española. Esta información es tratada para aportarle valor y ofertarla a través de servicios a particulares y empresas.

Estas empresas componen un importante segmento que ofrece información y servicios jurídicos a partir de la información reutilizada. Pero esto no quiere decir que no haya otro tipo de oportunidades.

Tenemos que citar el ejemplo de www.euroalert.net un portal de la empresa vallisoletana Gateway S.C.S. que ofrece información de convocatorias, programas, iniciativas, licitaciones y oportunidades comerciales publicadas en el DOUE (Diario Oficial de la Comunidad Europea). Los servicios que Euroalert presta provienen de la Comunidad Europea y se ofrecen a países de toda la Comunidad Europea.

Tipología de clientes

Las empresas infomediarias ofrecen servicios fundamentalmente a empresas y autónomos, aunque esto no es óbice para que esa misma información se muestre a ciudadanos. Muchas veces a estos últimos se les ofrece de forma gratuita, pero al mismo tiempo limitada (veremos estos casos en el siguiente punto). Incluso siguiendo con

el ejemplo anterior, www.euroalert.net, se encuentran casos de empresas públicas que utilizan los servicios de Euroalert para mostrar información pública proporcionada por una empresa privada, que la reutiliza, le aporta valor y la oferta.

Estrategias de comercialización

Lo cierto es que la información reutilizable se obtiene por lo general a través de medios telemáticos, de esta forma se facilita su gestión y tratamiento. La explotación de los datos y los modelos de comercialización es uno de los puntos donde más diversidad encontramos entre las distintas empresas. Los modelos que habitualmente solemos encontrar, y que hay que estudiar para averiguar cuál es el que mejor encaja en nuestro modelo de negocio, son los siguientes:

- *Gratuito sin restricciones*: este modelo se caracteriza por que la información se ofrece de forma gratuita sin generar ingresos. Este tipo de estrategia se suele producir en empresas que quieren utilizar el servicio ofertado a partir de la reutilización de información del sector público como un valor añadido, es decir, los ingresos vienen a través de otros servicios. El sector infomediario es uno de los pocos sectores donde esto es viable ya que una vez realizada la inversión consistente en desarrollar los programas informáticos que recuperan la información, la obtención de la materia prima (la información) tiene un coste residual o inexistente.
- *Gratuito mediante registro*: este modelo es similar al anterior, pero de cara a controlar el acceso y uso que se hace de la información es necesario que el usuario se identifique, registrándose en las aplicaciones como usuario y debiendo aceptar las condiciones de uso que se impongan. Estas condiciones nunca deberán ser abusivas, e incluso deben ser beneficiosas para ambos ya que el objetivo es establecer con nuestro cliente una relación *win-win*.
- *Gratuito mediante publicidad*: en este caso se trata de un modelo clásico que es, al mismo tiempo, el más utilizado. El objetivo es obtener una masa de clientes que, dado que desde nuestra aplicación se oferta una

tipología de producto concreto (información meteorológica, jurídica, etc.), estará claramente segmentada. A partir de ahí en nuestra web se ofrecerá publicidad que tendrá más posibilidades de éxito que la de una web más generalista al estar claramente segmentada. No obstante también es cierto que este mismo factor que nos beneficia al tener un tráfico segmentado hace a su vez que el volumen de visitas descienda con lo cual los ingresos por publicidad son menores.

- *Acceso/uso/trabajo realizado*: este es el modelo que más se acerca al modelo de ingresos tradicional. El usuario abona una cantidad por el uso que hace de los servicios que se ofrecen en el portal.
- *Suscripción lineal*: este es sin duda el modelo más utilizado en empresas infomediarias. Los clientes abonan una cantidad por un servicio que se presta a lo largo del tiempo. Un ejemplo es la empresa a la que anteriormente hacíamos referencia, Euroalert. La empresa recoge diariamente la información sobre las licitaciones del DOUE. Estas licitaciones se categorizan, se clasifican y se almacenan. Euroalert ha desarrollado una tecnología llamada Widget que permite presentar de forma sencilla en la página web del cliente las licitaciones a nivel europeo relacionadas con la temática concreta. Así, por ejemplo, cuando un usuario accede a la página web de Puertos del Estado y entra en la sección correspondiente a licitaciones europeas, está accediendo a los sistemas de Euroalert que, sobre la página de Puertos del Estado, muestran la información solicitada sobre todas las licitaciones relacionadas con actividades portuarias en la Unión Europea. Esta información aporta mucho valor a los usuarios de Puertos del Estado, la entidad pública necesitaría dedicar diariamente varios recursos para revisar y recuperar esta información del DOUE, por un importe periódico pueden ofrecer un excelente servicio a un bajo coste.
- *Suscripción con modalidades*: este modelo es una combinación de modelos anteriores: ofrece servicios puntuales por un precio y un servicio continuado por un importe inferior a la suma de todos los trabajos puntuales.

- *Servicios freemium*: se trata de un modelo en el que una parte del servicio o producto se ofrece de forma gratuita a todos los usuarios, mientras que los usuarios que desean utilizar el servicio avanzado deben pagar por el mismo (el mismo término lo define ya que es una combinación de las palabras inglesas *free*, gratis, y *Premium*). Lo complicado de este modelo es encontrar el punto de inflexión en el cual se ofrece un servicio gratuito a los clientes que, pese a ser bueno, no aporte todo el valor que incluye el servicio Premium de pago, lo que hará que para un porcentaje de ellos compense el abonar una suscripción mensual para obtener este servicio extra. Una proporción mínima adecuada sería conseguir un 10 % de usuarios Premium frente a un 90 % de usuarios gratuitos, aunque cualquier mejora a dicha proporción sería bienvenida.

Existen variantes a este modelo como establecer limitaciones por tiempo de uso, este por ejemplo es el modelo al que ha tenido que migrar el servicio de música on-line Spotify. Esta empresa se caracterizaba por ofertar un producto gratuito que era tan bueno que la mayoría de los usuarios no contemplaban la necesidad de migrar al servicio de pago. Spotify ofrecía un acceso ilimitado a una inmensa base de datos de temas musicales en un servicio gratuito mediante publicidad. De esta forma no obtenía la tasa de migración a clientes Premium que esperaba por lo que optó por limitar el número de horas diarias que el servicio estaba disponible en modo gratuito con lo que muchos usuarios que ya disfrutaban del servicio gratuito optaron por pagar una pequeña cuota mensual y disponer del servicio sin limitación, aprovechando además los servicios extra que, de otra forma, no eran tan interesantes.

Es importante destacar que no hay un modelo que sea mejor que otro, simplemente son diferentes, en base al tipo de cliente, tipo de producto, el mercado al que nos dirigimos, etc. será más conveniente optar por un modelo u otro. De una forma o de otra es el tipo de producto y cliente el que nos tiene que definir la estrategia de comercialización y no al revés.

5_3 Qué ofrecen las empresas infomediarias

En el año 2009 se creó el portal www.aporta.es (desde julio de 2012, <http://datos.gob.es>) desde donde se comunicaban y dinamizaban todas las iniciativas relacionadas con reutilización de información del sector público en España y, más concretamente, el Proyecto Aporta. El objetivo del portal es el fomento de la apertura y reutilización de la información del sector público en España y el desarrollo de productos y servicios de valor socioeconómico añadido a partir de los datos que custodian los departamentos, organismos, entidades y agencias oficiales de nuestro país, tal y como se describe en el propio portal.

Esta iniciativa se orienta fundamentalmente a dos vías. Una primera vía dedicada a la creación de un catálogo de datos de reutilización que abordaremos más adelante y una segunda, que es la que nos interesa en este punto, conformada por un registro de ejemplo de utilidades desarrolladas a partir de información reutilizable del sector público. En el momento de escribir este libro existen registradas como ejemplos 41 aplicaciones de reutilización de información del sector público desarrolladas por distintas empresas españolas, algunas han sido descritas en este libro. Las categorías entre las que se encuentran estas aplicaciones son las siguientes:

- ☛ Transporte (13).
- ☛ Sector público (8).
- ☛ Ciencia y tecnología (5).
- ☛ Medioambiente (3).
- ☛ Industria (2).
- ☛ Comercio (2).
- ☛ Energía (2).
- ☛ Cultura y ocio (1).
- ☛ Hacienda (1).
- ☛ Demografía (1.)
- ☛ Legislación y justicia (1).
- ☛ Economía (1).
- ☛ Urbanismo e infraestructuras (1).

Un ejemplo es la web de Tuitrafico¹⁷, se trata de una iniciativa de uso gratuito en la que se utilizan los catálogos de datos de la Dirección General de Tráfico de incidencias en las carreteras, del Ministerio de Industria para obtener el precio de los carburantes por estación de servicio y del Instituto Nacional de Meteorología para obtener los datos meteorológicos y las alertas causadas por ese motivo. A partir de esta información, se utiliza el servicio de Twitter para publicarlos ofreciendo una información del tráfico tanto vía web como a través de esta aplicación. Adicionalmente, el usuario tiene la opción de, a través de georreferenciación, establecer su ubicación actual y de esta forma ser consciente de cualquier incidencia que se produzca en su viaje. Además se ha establecido un canal para que los usuarios, en el momento que encuentren una incidencia en su viaje, la puedan reportar al portal que automáticamente la incorpora al registro de incidencias que serán reportadas a su vez a otros usuarios.

Otro ejemplo claro es el de la web www.goolzoom.com, esta web ofrece un servicio a usuarios a través del cual ordena al menos 3.388 servidores WMS (*Web Map Services*). Se trata de un estándar auspiciado bajo la directiva *Inspire* (impulsada desde la Unión Europea) para el acceso a información geográfica a través de Internet. Desde la web se tiene acceso a un servicio freemium en el cual se puede consultar toda la información ordenada y organizada de los servidores WMS a los que la web tiene acceso. Adicionalmente, usuarios Premium pueden disponer de servicios adicionales como la impresión de planos, información catastral, elaboración de informes de tasaciones, descarga de mapas en formato GIS, etc.

No obstante existen muchos más servicios y empresas que no están incluidos en el presente repositorio de aplicaciones. Lo más importante no son ya las empresas o aplicaciones que ya existen sino de dónde obtienen los datos que necesitan para ofrecer sus servicios.

5_4 Repositorio de datos

Un repositorio reutilizable es un almacén de datos propiedad de una entidad pública que está disponible para su consulta por aquellos usuarios (ciudadanos o

¹⁷<http://trafico.lainformacion.com>

empresas) que lo requieran. Los datos son ofrecidos por entidades públicas de forma individual. No obstante, si somos conscientes de la existencia de algún dato que no ha sido hecho público por parte de sus productores, tenemos la posibilidad de solicitárselo. De acuerdo a la última revisión de la directiva de la Unión Europea, a menos que este dato entre dentro de alguno de los epígrafes descritos en el Capítulo 3 de este libro, es un dato reutilizable por defecto.

De cara a facilitar esta tarea, los principales catálogos de información han sido recogidos en el portal datos.gob.es para que sean fácilmente accesibles y reutilizables. Existen más de 275 catálogos accesibles desde este portal. Aunque esta cantidad, todavía pequeña, se tiene que ver incrementada en breve de acuerdo con el Real Decreto 1495/2011, donde se establecen plazos de publicación obligada de nuevos catálogos de datos en el portal, tal y como explicamos en el Capítulo 6.

No vamos a listar estos 275 catálogos aunque creemos que sí es importante detallar las distintas categorías de información que están accesibles en base al tipo de catálogo que está publicado. De esta forma cabe hablar de las siguientes temáticas:

Ciencia y tecnología

Esta categoría cobra especial importancia con la entrada en vigor de la nueva directiva pendiente de aprobación por parte de la Unión Europea y que es uno de los valores diferenciadores de esta última actualización, ya que en el momento en que se apruebe pasará a contener catálogos de datos que anteriormente no eran reutilizables. Incluye datos sobre telecomunicación, tecnología, I+D+i, patentes y, tal y como comentábamos, repositorios relacionados con datos de proyectos de investigación que han sido financiados con fondos públicos (hasta ahora inaccesibles).

Comercio

Como su nombre indica, en esta categoría se engloban todos aquellos repositorios que contengan información relacionada con actividades comerciales, como la información de importación y exportación proporcionada

por el ICEX, información de Cámaras de Comercio, normativa relacionada con la ordenación del comercio, registros de actividades comerciales o participantes en las mismas, etc.

Cultura y ocio

Al igual que en ciencia y tecnología, este apartado tiene un gran impulso con la aprobación de la nueva revisión de la directiva europea ya que a toda la información relacionada con literatura, cine, música, danza y teatro, hay que incluir todos aquellos libros digitalizados en bibliotecas que no estén sujetos a derechos de autor. Habría que citar por ejemplo la gran cantidad de documentos que están almacenados en el nodo español de la Biblioteca Digital Europea, que no es otro que la Biblioteca Digital Hispánica gestionada desde la Biblioteca Nacional¹⁸.

Demografía

Información demográfica real o estadística sobre la población. Parte de esta puede ser obtenida, entre otras fuentes, del Instituto Nacional de Estadística que elabora detallados informes configurables accesibles desde su página web.

Deporte

Repositorios que contengan todo tipo de información relacionada con esta temática, desde ayudas y becas a deportistas hasta listados de federaciones y asociaciones o ubicación, características e información sobre instalaciones deportivas, centros de enseñanza deportiva, etc.

Economía

Esta categoría contiene almacenes de datos con información económica y financiera de todo tipo: financiación local y autonómica, estudios de evolución de precios de determinados mercados como el tecnológico, acceso a la biblioteca virtual del Ministerio de Economía y Hacienda donde se pueden consultar documentos a texto completo, información del mercado de valores, información contable como el balance de cuentas de compañías aseguradoras, etc.

¹⁸<http://bdh.bne.es/bnerearch>

Educación

Apartado que contiene tanto información para estudiantes de todos los niveles y profesorado, como material que es susceptible de ser usado como material educativo. Por ejemplo se ofrece un acceso al catálogo avanzado de la Biblioteca Digital de la Real Academia de la Historia desde el que se puede acceder a documentos históricos (incluso incunables y manuscritos) de altísimo valor y no accesibles a través de otras vías (excepto a expertos autorizados) por motivos de conservación.

Empleo

En esta categoría se ofrece acceso a distintas páginas de entidades públicas donde se puede consultar la información sobre las contrataciones que van a ejecutar. Así, por ejemplo, tenemos acceso al perfil del Contratante de la entidad pública empresarial Red.es o a las ofertas de empleo del sistema portuario. Además, se puede obtener acceso a información estadística sobre la población activa, datos de empleo y desempleo directamente del Instituto Nacional de Estadística o información para el fomento del autoempleo, protección del desempleo, etc.

Energía

Aquí tenemos información tan diversa como los listados de los electrodomésticos en España con su calificación de eficiencia energética, acceso a un directorio de empresas de energías renovables o acceso a un listado de los documentos aprobados por la CNE (Comisión Nacional de la Energía) referente a cualquiera de las áreas posibles: gas, electricidad, energía nuclear o energía “verde” (obtenida a través de un desarrollo sostenible).

Hacienda

Engloba esta categoría toda la información relacionada con la Agencia Tributaria, desde impuestos y recaudación hasta la consulta sobre la doctrina tributaria, consulta de bienes enajenados por la Agencia Tributaria incluyendo las fechas de las próximas subastas, información de recaudación o información catastral, etc.

Industria

En esta categoría podemos encontrar todos aquellos almacenes de datos relacionados con la actividad empresarial e industrial, tanto a nivel de pymes como de grandes empresas. Tendremos disponible un buscador de prospecciones tecnológicas industriales cubriendo sectores tan diversos como la automoción, la electrónica, el sector del metal, etc. Se tiene acceso a la biblioteca de Industria y Energía e incluso a publicaciones, fotografías y vídeos de la Empresa Nacional de Residuos Radioactivos, a la legislación histórica sobre propiedad industrial, etc.

Legislación y justicia

Una de las categorías con más información y mejor organizada de todas gracias al esfuerzo del BOE (Boletín Oficial del Estado) y del CENDOJ (Consejo General del Poder Judicial). Aquí tendremos acceso a todo tipo de jurisprudencia histórica y actualizada a todos los niveles: sentencias del Tribunal Constitucional, Audiencia Nacional, Tribunal Superior de Justicia, Audiencia provincial; información completa de leyes, reales decretos, etc.; bases de datos de legislación extranjera, propuestas legislativas, informes y comunicaciones de la Comunidad Económica Europea; publicaciones del Ministerio de Justicia o legislaciones nacionales y europeas sobre telecomunicaciones...

Medioambiente

En un mundo donde cada vez más la ecología y el respeto por el medioambiente marca la directriz de muchas oportunidades para empresas y negocios, entre los repositorios que se pueden encontrar tenemos información diversa de las distintas confederaciones hidrográficas (Guadalquivir, Ebro, etc.); umbrales de información climatológica por provincias; boletines hidrológicos o bases de datos de ayudas a los sectores agroalimentario, pesquero y medioambiental; así como una amplísima información proveniente del Instituto Español de Oceanografía entre la que se encuentra su base cartográfica, su repositorio gráfico y audiovisual o las bases de datos de trabajos de investigadores y publicaciones periódicas.

Medio rural y pesca

Incluye, entre otros temas, almacenes relacionados con sanidad vegetal y animal como el Boletín de Sanidad Vegetal sobre Plagas; información sobre producciones agrícolas y ganaderas; información sobre la regulación de los mercados como la información sobre los precios percibidos, pagados y los salarios agrarios; información sobre el etiquetado de la producción; denominaciones de origen e indicaciones geográficas protegidas o registro de etiquetado de un sector cárnico concreto como puede ser el vacuno; información sobre espacios protegidos y las especies que habitan en ellas o información sobre acuicultura.

Salud

Aquí encontramos repositorios con información diversa. Un repositorio básico en este área es el banco de datos del Ministerio de Sanidad, Servicios Sociales e Igualdad con información sobre el barómetro sanitario, datos de la Encuesta Nacional de Salud, diversas estadísticas y el catálogo de centros de atención primaria del Servicio Nacional de Salud así como el catálogo de hospitales. Del mismo modo encontramos información sobre salud ambiental, salud mental, enfermedades raras, etc.

Sector público

Esta es la categoría donde más diversidad de información se puede encontrar. Básicamente tenemos información completa relacionada con los organismos: organigramas, presupuestos, contrataciones y licitaciones. Del mismo modo se encuentra información sobre boletines oficiales y normativa relacionada con el ejercicio de la función pública. Pero tenemos también otra información tan interesante como es la recopilación de la presencia en las redes sociales de la Administración General del Estado. Se dispone de un localizador de oficinas de correos o incluso un localizador de envíos de correos.

Seguridad

Aunque obviamente esta información está limitada por los principios descritos en el Capítulo 3, no dejamos de disponer de información relacionada con los cuerpos

y fuerzas de seguridad del Estado. De esta forma no tendremos acceso a investigaciones abiertas, pero sí tendremos acceso a la localización de comisarías de policía, la búsqueda de plazas en las fuerzas armadas o información tan dispar como un buscador de cartas náuticas del Instituto Hidrográfico de la Marina. También entrarían en este epígrafe los almacenes relacionados con las instituciones penitenciarias.

Sociedad y bienestar

Esta categoría recoge todos los repositorios relacionados con aquellos segmentos de la sociedad relacionados con la cultura del bienestar: mayores, niños, sectores desfavorecidos, discapacitados, etc.

Transporte

Segmento importante y con grandes posibilidades de negocio, engloba los almacenes de información relacionados con control de tráfico en distintos medios, estado de las carreteras y vías circulatorias, siniestralidad y control de tráfico. Pero no solo aplicado al segmento del automóvil, sino también aviación civil, transporte de mercancías y marina mercante. Localizadores de ITV o Jefaturas Provinciales de Tráfico, consulta de estado de circulación de carreteras, consulta sobre el saldo de los puntos del carné de conducir, el listado de buques en determinados puertos o información de vuelos de Aena son algunos de los almacenes de datos.

Turismo

Respecto a esta categoría se recoge en el portal información referente a ocupación hotelera a través de buscadores de alojamientos turísticos. También se dispone de información sobre oficinas exteriores de información turística, comportamientos y hábitos turísticos, información de hostelería, etc.

Urbanismo e infraestructuras

Se trata de un área altamente desarrollada, programas como Urbanismo en Red de la entidad pública Red.es, la oficina virtual del catastro o el Instituto Geográfico

Nacional, galardonado con el premio “5 star PSI award” en 2012, ofrecen catálogos de información detallados y completos. El impulso de la directiva *Inspire* desde la Comunidad Europea ha sentado las bases tecnológicas para la puesta a disposición de la información relacionada. Información cartográfica, mapas georreferenciados, información catastral, etc. son algunos de los datos accesibles a través de estos almacenes.

Vivienda

Esta última categoría se refiere a todos los almacenes de datos relacionados con el sector de la construcción y el sector inmobiliario. Alguno de los ejemplos de servicios de datos son el buscador de viviendas de la Sociedad Pública de Alquiler o la mediateca de la entidad pública SEPES.

Estas categorías y repositorios son ejemplos de algunos de los servicios que están disponibles. A partir de estos, son los emprendedores los que tienen que diseñar servicios, construir productos que puedan aportar valor y generar riqueza y empleo. Una combinación de mapas georreferenciados del Instituto Geográfico Nacional, información sobre el precio de los carburantes del Ministerio de Industria, información del estado de las carreteras por parte de la DGT e información meteorológica del AEMET, pueden servir como base para construir servicios para transportistas. Esto es solo un simple ejemplo que ilustra perfectamente la potencialidad de estos almacenes de datos.

Lo más importante actualmente es que, de acuerdo a lo establecido en el Real Decreto 1945/2011, las entidades públicas comiencen a incorporar nuevos repositorios. En esta línea, desde el portal <http://datos.gob.es> se ha publicado un procedimiento y plantillas que ayudan en esta tarea. Esto traerá, sin duda, nuevas oportunidades de negocio.

Por último, destacar que este portal ha sido galardonado en enero de 2012 como el portal europeo de apertura de datos más amigable en el concurso de la red LAPSI (red profesional dedicada a los aspectos legales de la información del sector público) por su diseño atractivo, usabilidad y difusión de la reutilización de los datos públicos.

6_Tecnología asociada a la reutilización



6_1 Principios básicos de partida

El 8 de mayo de 2012 se celebró en San Sebastián, el día Open data, enmarcado en el Congreso Internacional de Ciudadanía Digital. En el marco de este evento se elaboró un decálogo (en formato borrador en el momento de la redacción de este libro) que recoge los puntos básicos que tienen que ser seguidos de cara a fomentar la reutilización:

1. Publicar datos en formatos abiertos y estándares.
2. Usar esquemas y vocabularios consensuados.
3. Realizar un inventario en un catálogo de datos estructurado.
4. Mostrar los datos de manera accesible desde direcciones web persistentes y amigables.
5. Exponer un mínimo conjunto de datos relativos al nivel de competencias del organismo y su estrategia de exposición de datos.
6. Tener compromiso de servicio, actualización y calidad del dato, manteniendo un canal eficiente de comunicación.
7. Monitorizar y evaluar el uso y servicio mediante métricas.
8. Exponer los datos bajo condiciones de uso no restrictivas y comunes.
9. Evangelizar y educar en el uso de datos.
10. Recopilar aplicaciones, herramientas y manuales para motivar y facilitar la reutilización.

De este decálogo queremos poner especial énfasis en los seis primeros puntos por su directa relación con este capítulo:

1. Publicar datos en formatos abiertos y estándar

Cualquier iniciativa enmarcada en el Open data tiene que publicar sus conjuntos de datos en formatos abiertos, no comerciales. De esta forma se garantiza que cualquier ciudadano o empresa pueda tener acceso a los mismos sin necesidad de adquirir software o licencias de ningún tercero. Además, estos formatos tienen que ser adecuados para permitir su reutilización por parte de los destinatarios.

Existen formatos abiertos que no son reutilizables. En 2007, se aprueba la Ley del Suelo 8/2007 en la que en la

disposición novena, “Modificación de la Ley Reguladora de las Bases del Régimen Local”, se establece que:

“Las Administraciones Públicas con competencias en la materia publicarán por medios telemáticos el contenido actualizado de los instrumentos de ordenación territorial y urbanística en vigor, del anuncio de su sometimiento a información pública y de cualesquiera actos de tramitación que sean relevantes para su aprobación o alteración.

En los municipios menores de 5.000 habitantes, esta publicación podrá realizarse a través de los entes supramunicipales que tengan atribuida la función de asistencia y cooperación técnica con ellos, que deberán prestarles dicha cooperación”.

Esta disposición en la Ley del Suelo tenía por objeto la difusión de la información urbanística entre los ciudadanos generando transparencia y, al mismo tiempo, favoreciendo la reutilización de la información del sector público. La disposición partía con unas premisas que le beneficiaban, como era la aprobación en abril de 2007 de la directiva *Inspire* que tenía como objetivo la creación de una infraestructura de datos espaciales a nivel europeo y que regulaba la forma en la que la información geográfica era mostrada y los datos eran transmitidos a través de Internet mediante el uso de los servicios inteoperables que los componían.

No obstante, en España la directiva de la Ley del Suelo no establecía condiciones excepto que la información debía estar publicada. De ahí que muchos ayuntamientos optaran por la medida fácil y rápida para cumplir la ley: publicar el planeamiento urbanístico en formato pdf que, aunque es un estándar abierto, ni estaba georreferenciado, ni cumplía ninguna de las premisas de la directiva *Inspire*, ni cubría los objetivos que pretendía la Ley del Suelo, ya que el mapa del planeamiento en formato pdf no puede ser tratado de forma automática.

2. Usar esquemas y vocabularios consensuados

De cara a fomentar la reutilización, además de los mencionados formatos abiertos y estándar, se debe definir una estructura de datos que siga un convenio o esquema predefinido, en el caso por supuesto de que existan.

De esta forma, cuando se creó en España el proyecto de la Biblioteca Digital Hispánica en la Biblioteca Nacional en la cual ya hay publicados más de 75.000 fondos accesibles por Internet, era necesario crear un modelo que fuera capaz de dar soporte a la información y a las necesidades de búsqueda. Por ello se optó por utilizar estándares bibliotecarios para la catalogación y publicación en Internet de esos catálogos. La plataforma soporta estándares como SCORM (del inglés, *Sharable Content Object Reference Model*) que permite crear objetos pedagógicos estructurados. De esta forma a través de este estándar se satisfacen principios como:

- Accesibilidad.
- Adaptabilidad.
- Durabilidad.
- Interoperabilidad.
- Reusabilidad.

Cuando ocurra el caso de que los esquemas y vocabularios no existan y sea necesario crearlos, se deberán seguir convenios y esquemas establecidos construyendo unos modelos que serán publicados y difundidos de cara a que cualquier otra entidad pueda seguir el mismo modelo o pueda reutilizarlo si lo cree adecuado.

En la entidad pública empresarial Red.es, a solicitud del entonces Ministerio de Industria, Turismo y Comercio (hoy Ministerio de Industria, Energía y Turismo), se inició un proyecto destinado a facilitar a las entidades locales (ayuntamientos, diputaciones, cabildos, mancomunidades, etc.) la posibilidad de gestionar el planeamiento urbanístico en formato digital publicándolo en Internet para que pudiera ser accesible para todos los ciudadanos y empresas. En el marco del convenio creado para la ejecución de los trabajos¹⁹, Red.es consultó a un grupo de expertos que le ayudaron a definir un modelo de datos que fuera capaz de dar cabida a la realidad del planeamiento urbanístico en todo el territorio nacional. No hay que olvidar que las competencias de urbanismo recaen en las entidades locales, aunque el planeamiento tiene que cumplir la normativa de la comunidad autónoma y, al mismo tiempo, esta normativa está supeditada a la Ley del Suelo de ámbito nacional.

¹⁹ Puede consultarse en www.urbanismoenred.es

Esto puede servir para hacernos una idea de la dificultad del problema a acometer, no es lo mismo los “conceptos” urbanísticos en una comunidad autónoma con su propio idioma que en otra con otro distinto. No es lo mismo el planeamiento en un municipio de costa (donde hay que considerar playas, el mar, zonas de protección para la pesca, etc.) que uno de montaña con sus términos propios. Desde el grupo de expertos se llegó a una solución consensuada sobre la que se basaron los trabajos. Esta solución está publicada en la página anteriormente citada accesible para las entidades locales que quieran participar y planteen sistematizar su propio planeamiento, para las que hayan participado y se planteen redactar un nuevo planeamiento urbanístico sistematizado o para empresas que quieran adoptar la solución y desarrollar nuevas herramientas que le den soporte o integrarlo con otras herramientas existentes.

3. Realizar un inventario en un catálogo de datos estructurado

Cualquier iniciativa Open data debe tener un punto de consulta donde se incluya un inventario con información descriptiva y técnica sobre los conjuntos de datos que se exponen. Este es el principio del catálogo que hemos descrito anteriormente, catálogo que aunque no está todo lo actualizado que debiera (esta tarea recae en los autores de los datos), sí presenta una estructura común y estándar y establece unos procedimientos para que los distintos autores (entidades públicas) puedan publicar nuevos catálogos de datos. Los metadatos que informan sobre cada conjunto de datos tienen que seguir una estructura común y estándar. En el caso del catálogo de almacenes de datos de datos.gob.es, para cada uno de los registros se guarda la siguiente información:

- Nombre del catálogo.
- Descripción detallada del mismo.
- Condiciones de reutilización del catálogo.
- Categoría (entre las listadas en el capítulo anterior).
- Fecha de la última actualización (del catálogo al que hacemos referencia).
- Frecuencia en que el catálogo es actualizado.

- Cobertura geográfica.
- Detalle geográfico.
- Tarifa aplicable para la reutilización de la información.
- Idiomas en lo que se encuentra la información del catálogo.
- Lista de distribuciones en la cual está accesible el catálogo (formatos y/o servicios a través de los cuales podemos acceder a la información del catálogo).

Del mismo modo, en el caso de que existan necesidades especiales de clasificación (como puede ocurrir en Urbanismo en Red, donde es necesario ajustarse a una normativa), se tienen que compartir las especificaciones para asegurarnos que los datos están correctamente catalogados.

4. Mostrar los datos de manera accesible desde direcciones web persistentes y amigables

Este punto se trata de un tema meramente técnico. Existen portales estáticos y dinámicos, esto significa que las direcciones a través de las cuales se accede a las distintas páginas web y/o catálogos pueden ser dinámicas (es decir varían en base a distintos criterios y son diferentes para distintos usuarios) o estáticas (se trata de direcciones en Internet fijas). Las Administraciones Públicas tienen que optar por este último modelo, esto es así ya que de otra forma se perderían las referencias a las fichas o incluso al propio catálogo ya que la dirección del mismo, en caso de direcciones dinámicas, irá variando a lo largo del tiempo. Es necesario que el acceso a todos los recursos sea persistente a lo largo del tiempo para evitar que estas referencias se pierdan más adelante.

Del mismo modo, es deseable que las direcciones de acceso sean claras y bien definidas para que los reutilizadores, solo con ver la dirección, tengan claro cuál es el contenido de la página a la que se refiere. Por ejemplo, en www.bne.es/es/Catalogos/BibliotecaDigital, “_bne_” hace referencia a la Biblioteca Nacional Española, “_es_” indica el idioma de la página, “_ Catalogos_” la sección catálogos y, por último, “_ BibliotecaDigital_” la biblioteca digital. Este es un perfecto ejemplo autoexplicativo del modelo de dirección a seguir.

5. Exponer un mínimo conjunto de datos relativos al nivel de competencias del organismo y su estrategia de exposición de datos

Con la aplicación del Real Decreto 1495/2011, se establece la obligación de que en el plazo de un año las entidades públicas deben definir un plan de publicación. Este plan de publicación debe incluir la estrategia de publicación de la entidad pública y sería el sitio adecuado para exponer los datos solicitados en el catálogo. Tras la aprobación del Real Decreto, la fecha límite para la publicación de esta estrategia es el 8 de noviembre del 2012.

Lo que se pretende con este punto es que cada administración tenga clara, tanto para su propia organización como para los reutilizadores, la estrategia de exposición de los conjuntos de datos. En esta estrategia se debe incluir la prioridad de publicación y tener en cuenta los deseos e intereses de los usuarios finales de la información.

6. Tener compromiso de servicio, actualización y calidad del dato, manteniendo un canal eficiente de comunicación

Tomando como base la estrategia de publicación definida por la Administración, lo que es fundamental es el mantenimiento de los almacenes de datos publicados. Es necesario cumplir con unos requisitos con el fin de garantizar la validez de la información a lo largo del tiempo, así la Administración debe mantener un mínimo de calidad y servicio en su iniciativa Open data, manteniendo lo expuesto en dicha estrategia de publicación y comprometiéndose con el colectivo reutilizador. Debe establecer un canal eficiente de comunicación que permita la interacción bidireccional organismo público-reutilizadores, garantizando el mantenimiento del servicio a través de un catálogo de datos actualizado.

6_2 Formatos de reutilización

Existen diferentes opiniones sobre cuál es el formato más adecuado para fomentar la reutilización. Por lo que a este autor respecta, el formato adecuado es el que interesa al usuario final. Dependiendo del uso para

el que esté destinado el dato que se quiere reutilizar, el formato adecuado es uno u otro.

No hay duda de que lo más recomendable son los formatos abiertos. Cuando hablamos de un formato abierto, estamos haciendo referencia a una forma de gestionar, tratar y almacenar datos informáticos que suele estar definida por una organización que establece las especificaciones necesarias para realizarlo sin establecer ningún tipo de restricciones legales y económicas. Normalmente está impulsado por organizaciones no lucrativas, aunque es cierto que en algunos momentos, cuando estas organizaciones promovían un estándar o un software libre que estaba haciendo excesiva competencia a otro fabricante, este ha terminado por adquirir los derechos sobre estos desarrollos para controlar su evolución (y la competencia sobre sus productos).

Las ventajas de los estándares abiertos es que pueden ser utilizados, implementados y accedidos desde cualquier aplicación que lo desee sin necesidad de adquirir una licencia de software o pagar ningún tipo de derecho al propietario del formato. Del mismo modo, esta aplicación tiene la garantía de que, en caso de que el formato del estándar abierto cambie, esta variación será debidamente documentada para que pueda ser incorporada a la misma.

El objetivo es claro, y es este objetivo lo que dicta que sea la mejor opción para la reutilización, se trata de que el acceso a los datos esté garantizado a largo plazo sin que se tengan que tener en consideración licencias, derechos o cualquier otro aspecto legal. Hay que tener en cuenta que aquí estamos hablando del **formato** de los datos y no del **contenido** de los mismos, este contenido sí puede estar sujeto a derechos y licencias. Por ejemplo, el último libro publicado por Arturo Pérez Reverte en formato ePub (el estándar abierto para libros digitales o ebooks), es un documento cuyos derechos de autor recaen en el escritor aunque el formato sobre el que se presenta y está soportado no esté sujeto a licencias.

Del mismo modo, a través del uso de formatos abiertos (que son un subconjunto de los estándares abiertos)

evitamos el depender de una tecnología concreta para acceder al mismo.

Todos estos principios apuntan directamente a los objetivos del Open data y de la reutilización de datos, ya que su uso fomenta la libre competencia, la distribución entre los distintos agentes y la distribución de la información.

Existe gran número de estándares abiertos. En el catálogo de datos que comentábamos en capítulos anteriores se pueden observar los distintos formatos en los cuales la información está disponible, realizando incluso una búsqueda de todos los catálogos que ofrecen información en un formato concreto. Ofrecer esta información junto con el catálogo es un aspecto básico para garantizar la reutilización. Los formatos que se ofrecen en este catálogo son los siguientes:

- HTML: acrónimo inglés de *HyperText Markup Language*, se trata de un lenguaje que nos sirve para estructurar textos y presentarlos de forma que puedan ser visualizados a través de navegadores web. Se trata de uno de los formatos más extendidos y populares de información accesible a través de un navegador. Ofrece para el reutilizador la ventaja de establecer una estructura de documento que permite su tratamiento de forma automática.
- XHTML: formato de la familia del XML que pretende desbancar al HTML en la especificación de páginas web, presenta con su predecesor la ventaja de que exige una especificación más estricta para representar el contenido del documento de ahí que sea más sencillo el desarrollar un programa que facilite la reutilización de este tipo de documento.
- XML: acrónimo en inglés que significa *eXtensible Markup Language*, es un lenguaje de marcas que permite la definición de la estructura sobre la cual queremos que se nos muestren los datos. Esto hace que el contenido y su estructura sea totalmente adaptable y permita la gestión, recuperación y tratamiento automático utilizando programas informáticos. Cuando hablamos de marcas nos referimos a cada uno de los elementos que definimos en el documento para

identificar los contenidos del mismo que han sido expresados de forma ordenada (véase el Cuadro 6.1). Este formato permite recoger prácticamente cualquier tipo de información textual que es, de hecho, el estándar de facto para el intercambio de información estructurada entre distintas plataformas (bases de datos, editores de texto, hojas de cálculo, etc.).

- PDF: formato abierto que ha sido desarrollado por una empresa privada (Adobe Systems). Su principal característica es que presenta el documento a reutilizar en la forma en la que va a ser impreso no ofreciendo de forma fácil (de hecho a veces es imposible) la posibilidad de extraer información del documento para su reutilización. Tiene la ventaja de ser un formato multiplataforma, es decir, casi cualquier sistema operativo, ordenador o dispositivo móvil puede visualizar su contenido. Además, en un mismo documento permite la integración de textos, imágenes o vídeos. Cuando el objetivo es obtener un documento para su impresión, este formato es el ideal ya que contiene toda la información que es necesaria para esta tarea (márgenes, tamaños de letra, sangrías, tipos de fuente, etc.).
- XLS: este formato no es abierto sino que es propietario (ha sido desarrollado por la compañía Microsoft), aunque su especificación es conocida con lo cual desde otras aplicaciones distintas a las que se creó se puede acceder y consultar la información que contenga el fichero. Pese a todo, es uno de los formatos de datos numéricos sobre los que ejecutar cálculos y agregaciones más conocidos y usados.
- ZIP: este no es un formato de representación de datos como el resto, sino que se trata de un formato de compresión. Es decir, es un formato que define la forma por la cual un conjunto de documentos que están incluidos dentro del fichero, son comprimidos (se reduce su tamaño) con el fin de enviarlos y almacenarlos de forma más sencilla. Los archivos que contenga este fichero en formato ZIP se encontrarán codificados en cualquiera de los otros formatos que hemos descrito.

Cuadro 6.1. Ejemplo de XML.

```

<?xml version="1.0" encoding="UTF8" ?>
<Opendata>
  <Documento>
    <Origen>
      <Nombre>Ministerio de Industria, Energía y Turismo</Nombre>
      <Catalogo_origen>Nombre del catálogo donde se ubica el original del
documento</Catalogo_origen>
      <Autor> Nombre del autor </Autor>
      <Mail> Correo de contacto </Mail>
    </Origen>
    <Datos_reutilización>
      <Condiciones> Condiciones de reutilización </Condiciones>
      <Categoria> Nombre de la categoria </Categoria>
      <Fechaact>Fecha última actualización</Fechaact>
      <Frecuenciaact> Frecuencia de actualización </Frecuenciaact>
      <Cobertura> Cobertura geográfica aplicable </Cobertura>
      <Tarifa> Tarifa aplicable si la hubiera </Tarifa>
      <Idiomas>
        <Idioma1> Idioma1 </Idioma1>
        <Idioma2> Idioma2 </Idioma2>
        <Idioma3> Idioma3 </Idioma3>
      </Idiomas>
    </Datos_reutilizacion>
    <Licencia>
      <Tipo_licencia> Tipo de licencia aplicable </Tipo_licencia>
      <Referencia> Dirección donde consultar el texto de la licencia</Referencia>
      <TextoLicencia> Texto de la licencia </TextoLicencia>
    </Licencia>
    <Contenido>
      <Registro>
        <Numero> 1 </Numero>
        <Idioma> ES </Idioma>
        <Titulo> Título del registro </Titulo>
        <Sinopsis> Breve resumen del contenido </Sinopsis>
        <Textocompleto>
          Texto completo del documento/registro
        </Textocompleto>
      </Registro>
      <<Registro>
        <Numero> 2 </Numero>
        <Idioma> EN </Idioma>
        <Titulo> Título del registro </Titulo>
        <Sinopsis> Breve resumen del contenido </Sinopsis>
        <Textocompleto>
          Texto completo del documento/registro
        </Textocompleto>
      </Registro>
    </Contenido>
  </Documento>
</Opendata>

```

Fuente: Elaboración propia.

- CSV: este es uno de los formatos más simples que existe, se trata de un acrónimo para el término *Comma-Separated Values* (en inglés, valores separados por comas). Son documentos sencillos, normalmente usados para representar datos (numéricos o no) en forma de tabla pero que se encuentran separados por comas o puntos y coma (para evitar el confundir las comas textuales de las comas de los valores numéricos con decimales en aquellos países que lo soportan). Conociendo la estructura de la tabla, a partir de los valores separados por comas es muy sencillo recuperar la información para su tratamiento y reutilización.
- XBRL: estándar basado en XML desarrollado por Charles Hoffman, experto contable y auditor. Este estándar auspiciado por la organización sin ánimo de lucro XBRL Incorporated International fue diseñado para simplificar la automatización del intercambio de información financiera. De esta forma, se consigue estandarizar el formato con el que esa información se distribuye entre los distintos agentes. Alguno de los datos que se muestran en este formato en el catálogo de datos públicos datos.gov.es son las cuentas anuales de la entidad pública empresarial Red.es.
- Texto plano: documento de texto sin ningún tipo de información de formato, lo único que es importante en el fichero es el contenido, pero este se encuentra disponible directamente sin ningún tipo de marca que permita acceder a él de forma estructurada. Por ello, la reutilización de la información que contiene (cuando lo que se requiera es el análisis del contenido) es muy complicada.

Pese a haber descrito algunos de los formatos de documentos disponibles en el catálogo datos.gov.es, existen muchos otros formatos abiertos susceptibles de ser utilizados en la reutilización. Completamos los anteriores formatos, extrayendo de la página de Wikipedia un listado más completo de formatos abiertos, sin entrar en detalle de los mismos, que pueden ser utilizados en un proceso de reutilización.

Multimedia

Se refiere a formatos abiertos para la gestión de ficheros de audio, imágenes y vídeo:

- ALAC: formato para archivos de audio. Antiguamente, formato propietario de la empresa Apple Inc.
- CMML: metadatos de tiempo y subtítulos.
- DAISY, *Digital Talking Book*: formato abierto para libros hablados.
- FLAC: *lossless audio codec*.
- JPEG 2000: formato de imagen estandarizado a través del ISO/IEC.
- Matroska (mkv): formato contenedor para ficheros multimedia de todo tipo: audio, vídeo, imágenes y subtítulos.
- MNG: imágenes con movimiento.
- Musepack: formato abierto de ficheros de audio.
- Ogg: contenedor de ficheros de vídeo.
- PNG: formato estándar de fichero de imágenes.
- SMIL: formato abierto para ficheros multimedia y listas de reproducción.
- SVG: formato estándar para imágenes vectoriales estandarizado por W3C.
- VRML/X3D: formato abierto para la representación de datos en 3D en tiempo real.
- WavPack: formato para ficheros de audio.
- WebM: formato de audio y de vídeo.
- XSPF: lista de reproducción de ficheros de audio.

Texto

- Texto plano.
- Formatos de transformación Unicode (UTF-8 y UTF-16): se trata de formatos de texto que son soportados por la mayor parte de lenguajes y programas.
- DVI: formato de texto independiente del dispositivo.
- ePub: formato de libros digitales definido por el International Digital Publishing Forum (IDPF).
- Office Open XML: formato para documentos de textos basado en ISO/IEC 29500:2008, se encuentra sujeto a licencias de uso.

- OpenDocument: formato de almacenamiento de texto formateado (ISO/IEC 26300:2006).
- OpenXPS: formato abierto ofrecido por la compañía Microsoft, se trata de la versión abierta del formato XLS.
- PostScript: al igual que el anterior, comenzó como un formato propietario para la descripción de páginas y lenguajes de programación y hoy en día es un sistema abierto y estándar.

Archivado y compresión

- 7z: formato de archivado y compresión.
- bzip2: formato de compresión.
- gzip: formato de compresión.
- MAFF: formato de archivado de páginas web basado en el formato abierto ZIP.
- PAQ: formato de compresión.
- SQX: formato de archivado y compresión.
- tar: formato de archivado.
- xz: formato de compresión.

Otros

- CSS: formato abierto que permite definir hojas de estilo, se usa habitualmente junto con HTML y sus derivados.
- DjVu: formato de fichero para imágenes o documentos escaneados.
- EAS3: formato binario para el almacenamiento de números decimales.
- ELF: formato ejecutable y enlazable.
- FreeOTFE: se trata de un formato que recoge un contenedor para datos encriptados.
- iCalendar: formato de datos para calendarios.
- JSON: notación de objetos reutilizables en lenguaje de programación Javascript.
- LTFS: *Linear Tape File System*.
- NetCDF: formato para el almacenamiento de datos numéricos en formato científico.
- NZB: formato usado para almacenar ficheros multi-partes y su transmisión a red Usenet.

- RSS: formato utilizado para la sindicación de noticias.
- SDXF: *Structured Data eXchange Format*, formato de intercambio de datos estructurado.
- SFV: formato checksum, habitualmente se utiliza para validar la integridad de otros ficheros (por ejemplo, ficheros comprimidos).
- TrueCrypt: contenedor para datos encriptados.
- YAML: formato de serialización de datos legible por los humanos.

Estos son solo unos cuantos de los formatos abiertos existentes, formatos que requieren un esfuerzo de elaboración por parte de las entidades públicas de ahí que muchas de las empresas reutilizadoras opten por solicitar los datos sin un formato concreto para facilitar la puesta a disposición de los datos por parte de las mismas, lo que se trata en el siguiente punto.

6_3 Raw data

Aunque he hablado de formatos para fomentar la reutilización no hay que olvidarse, tal y como comentaba, que no solo se reutilizan ficheros que se encuentran en formatos concretos (abiertos o propietarios). Hay veces que los usuarios prefieren disponer de servicios que proporcionan consultas puntuales que pueden ser integradas en su página web o ser utilizadas en sus operaciones diarias. Por ejemplo, recientemente en una conferencia se comentaba cómo la Compañía Cervecera de Canarias utiliza información del sector público para mejorar sus procesos productivos. Accediendo a la información de los datos de ocupación hotelera y a los porcentajes de extranjeros alojados en las Islas Canarias, es capaz de ajustar su producción adelantando las necesidades extra de producto final e, incluso, limitando los excesos de stock que deben ser almacenados. Es un ejemplo claro de cómo los estándares abiertos no siempre son la solución más adecuada y cómo a través de servicios que ofrecen datos puntuales se puede obtener la información necesaria para optimizar los procesos productivos.

No obstante, para realizar una reutilización exhaustiva y crear un servicio completo y fiable, esto no es suficiente. Tal y como hemos comentado, existe en algunas Administraciones Públicas una cierta confusión sobre lo que es un formato de información reutilizable. Ciertamente es que si publicamos una información en Internet en formato jpg (es decir una fotografía de un documento o un mapa) esta información se puede reutilizar. Pero es más cierto que no es muy operativo poder sacar partido a esa información (en el caso de poder hacerlo, que no siempre es posible) ya que se deben utilizar herramientas de software costosas y de limitada eficacia para extraer la información o bien transcribir de nuevo esa información de forma manual. Cualquiera de las dos opciones implica un tratamiento exhaustivo de la información, que, en el caso de que estemos hablando de un documento de longitud limitada es factible, pero en el caso de que queramos manejar, por ejemplo, datos estadísticos a nivel nacional compuestos por miles de registros, hace que el esfuerzo necesario para acometer esa tarea sea inabordable. Esto toma más peso todavía cuando se trata de información que periódicamente es actualizada, con lo que deberá repetirse el proceso una y otra vez.

Es por ello que los reutilizadores defienden el uso de la publicación de *Raw data* o datos en crudo. Este tipo de formato está compuesto exclusivamente por datos sin tratar ni catalogar. Se trata de los archivos en bruto que han sido producidos por las entidades locales y que se ofrecen a los reutilizadores para que sean ellos los que se encarguen de su gestión, tratamiento y manipulación para generar servicios de valor añadido.

Este modelo es, sin duda, el más barato para las entidades públicas pero hay que ser prudente en su uso ya que debe ajustarse a la ley vigente. Así, este fichero no puede contener datos de carácter personal, que deben de haber sido eliminados del mismo (lo que implica que el fichero ya ha sido tratado, con lo cual ya no se trata de *Raw data* sino de lo que se conoce como *cooked data*).

Es cierto que estos datos en crudo requieren un mayor tratamiento por parte de los reutilizadores pero ya en el

año 2007 Rufus Pollock²⁰ escribió un post en el blog de la Open Knowledge Foundation²¹ donde describía los motivos por los cuales las administraciones deberían proporcionar su información como Raw data.

En este post, que sigue estando totalmente vigente cinco años después, el autor describe un conjunto de acertados motivos por los cuales el Raw data es lo más acertado para gestión masiva de datos.

Con la situación económica actual (o incluso sin ella) puede ser más adecuado que las entidades públicas no dediquen recursos humanos y económicos a desarrollar un portal y servicios para ofertar unos datos que deben ser tratados de forma masiva. Lo que no quita que ciertos servicios especialmente útiles para los ciudadanos sí deben ser presentados de forma adecuada para que sean accesibles a toda la ciudadanía (hemos visto varios ejemplos a lo largo de este libro). Estos desarrollos tienen un conjunto de inconvenientes que Rufus Pollock describe de forma acertada:

- La primera limitación corresponde a que los datos en crudo permiten una reutilización completa por parte de las empresas que lo gestionan, por lo que una entidad pública no puede conocer con detalle cuál va a ser la forma en la que la información va a ser reutilizada. Aunque, como comentábamos en capítulos anteriores, sí puede establecer unas condiciones de utilización (citar fuente, no tergiversar la información, etc.) que limiten su responsabilidad en el uso de la información, sin interferir en las actividades que las empresas quieran realizar. De otra forma, el no ofrecer toda la información sino el tratarla para presentarla de una forma más ordenada hace que parte de estos datos se pierdan o, cuanto menos, que su tratamiento sea mucho más complejo.
- Es indudable que este tratamiento de la información tiene un coste, que, además, cuando se trata de información que es actualizada de forma periódica, supone un coste recurrente. Por supuesto que la publicación Raw data también tiene un coste, pero creo que todos estaremos de acuerdo en que es mucho más económico entregar 50 kilogramos de carne que preparar 2.000 hamburguesas con esa carne antes de entregarla.

²⁰ Economista cofundador de la Open Knowledge Foundation, organización sin ánimo de lucro que promueve la distribución gratuita del conocimiento, los contenidos abiertos y, por supuesto, el Open data.

²¹ <http://blog.okfn.org/2007/11/07/give-us-the-data-raw-and-give-it-to-us-now>

- Adicionalmente, hay que tener en cuenta el coste de desarrollar el portal para ofertar servicios con estos datos tratados. Cualquier portal tiene un coste de mantenimiento alto, en una media de tres años debe ser reconstruido totalmente para adaptarse a la tecnología que en el mundo de Internet evoluciona diez veces más rápido que en cualquier otro segmento empresarial. La publicación de datos en crudo (Raw data) a través de una simple página web y de un enlace de descarga tiene un coste mínimo.
- Además, el desarrollo de este tipo de portales hace que la interfaz gane más importancia que los datos que ofrece, lo cual es un error que puede hacer que perdamos el foco sobre la información que es importante. Además hay que tener en cuenta que esta interfaz no va a cubrir las necesidades de todos los usuarios, como anteriormente citábamos.
- Por último, una de las consideraciones más importantes de Pollock es el tema de la inmediatez. Un conjunto de datos que debe ser tratado (y sabemos que en algunos casos por temas de protección de datos es necesario), implica un trabajo que, dependiendo del volumen, puede suponer semanas. Durante este tiempo no se dispone de la información y, por lo tanto, no se saca partido a disponer de datos recientes en el momento justo.

Por todo ello desde la Open Knowledge Foundation se defiende la necesidad del Raw data.

Independientemente de las acertadas opiniones de Pollock sobre la necesidad de ofertar datos en crudo de forma inmediata, esto no quita para que adicionalmente (insisto, adicionalmente) se realice un tratamiento de la información para poder ofertar servicios a empresas y ciudadanos. Aunque tengan un coste adicional, los servicios a disposición de la sociedad suponen un ahorro para las Administraciones Públicas y un beneficio para los ciudadanos.

La oficina virtual de la Dirección General del Catastro es un claro ejemplo ya que lleva años ofertando certificados catastrales y otros servicios. Antiguamente, cuando una entidad pública requería de un ciudadano la presentación de un certificado catastral para algún tipo de trámite administrativo, en las oficinas de esta

Administración se formaban colas que daban la vuelta al edificio. Esto suponía una carga de trabajo muy alta en momentos puntuales, lo cual hacía muy difícil no solo conseguir la satisfacción de los usuarios, sino la organización de personal que en esos momentos tenía una carga inmensa de trabajo y en otros esta carga era residual. En esas circunstancias se podían tramitar unos 250.000 certificados catastrales anuales. Con la puesta en funcionamiento de la oficina virtual que permite, mediante el uso de certificados digitales, la generación on-line de certificados catastrales con validez jurídica, las colas y las cargas para los empleados desaparecieron y, adicionalmente, pasaron a tramitarse más de un millón de certificados al año. De no existir este servicio on-line, seguramente no se hubieran emitido este millón de certificados (no se hubiera desplazado ese volumen de ciudadanos, no pretendemos decir que las colas hubieran sido tan grandes), pero lo cierto es que la antigua situación se hubiera mantenido.

No obstante, desde la Dirección General del Catastro son conscientes de las necesidades de las empresas infomediarias, con lo que desde hace algún tiempo, adicionalmente a los servicios personalizados para ciudadanos, ofrecen la descarga masiva de toda la información catastral de España en un único fichero que puede ser tratado y gestionado por dichas empresas como Raw data. Pero al mismo tiempo se da servicio a otras iniciativas como Urbanismo en Red que realiza consultas puntuales a través de servicios web interoperables. Ambos servicios coexisten y son un éxito en sí mismos tanto individualmente como en conjunto, ofreciendo un servicio global.

Tal y como nos comenta Tim Berners-Lee respecto a este tema, es necesario desarrollar un clima adecuado de colaboración entre los productores de la información y los reutilizadores buscando un equilibrio entre las necesidades de todas las partes. De esta forma, conseguiremos una relación *win-win*, es decir una situación beneficiosa y satisfactoria para todas las partes. Todo ello pasa por que exista un entendimiento y una colaboración mutua en la que ambas partes comprendan y asuman la situación y necesidades de la otra parte para llegar a un equilibrio entre el funcionamiento de la Administración y las necesidades del mercado.



7_Marco jurídico y licencias



7_1 Marco jurídico y licencias

Aunque a lo largo del documento se ha ido describiendo la distinta normativa referente a la reutilización de información del sector público, es importante dedicar un espacio en este libro a recorrer la distinta legislación aplicable en nuestro país.

En el presente capítulo vamos a repasar los textos específicos aplicables a la reutilización de la información del sector público. Pero es importante destacar que, independientemente de la aplicabilidad de esta normativa, existen otras normativas que de forma indirecta están relacionadas con algunos de los aspectos descritos en las primeras. Estas normativas cubren aspectos tan diversos como el control de acceso a la información de carácter personal, la definición de precios públicos y el desarrollo de la Sociedad de la Información. No obstante, dado que no es objeto el entrar en tanta profundidad, estos textos adicionales solo serán mencionados. Los textos son los siguientes:

- Ley 56/2007, de 28 de diciembre, de Medidas de Impulso de la Sociedad de la Información.
- Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999.
- Ley 30/2007, de 30 de octubre, de Contratos del Sector Público.
- Ley 11/2007, de 22 de junio, de Acceso electrónico de los Ciudadanos a los Servicios Públicos.
- Ley 27/2006, de 18 de julio, por la que se regulan los Derechos de Acceso a la Información, de Participación Pública y de Acceso a la Justicia en Materia de Medioambiente.
- Ley 23/2006, de 7 de julio, por la que se modifica el texto refundido de la Ley de Propiedad Intelectual.
- Ley 59/2003, de 19 de diciembre, de Firma electrónica.
- Ley 34/2002, de 11 de julio, de Servicios de la Sociedad de la Información y de Comercio Electrónico.
- Ley 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

- Ley 30/1992, de 26 de noviembre, de Régimen Jurídico de las Administraciones Públicas y Procedimiento Administrativo Común.
- Ley 8/1989, de 13 de abril, de Tasas y Precios Públicos, modificada por Ley 58/2003, de 17 de diciembre.
- Ley 16/1985, de 25 de junio, del Patrimonio Histórico Español.
- Ley 26/1984, de 19 de julio, General para la Defensa de Consumidores y Usuarios, y su legislación complementaria.

Respecto a los textos normativos directamente relacionados con la reutilización y que afectan directamente a todas las actuaciones relacionadas con esta actividad es importante destacar en primer lugar el *Libro Verde sobre la Información del Sector Público en la Sociedad de la Información* publicado por la Unión Europea en 1998. En él, sin ser una normativa de obligado cumplimiento, se ponen de manifiesto la necesidad e importancia de la reutilización de la información del sector público, las oportunidades de negocio que supone, así como la necesidad de transparencia en las administraciones para garantizar el acceso democrático al conocimiento y a la información. El estudio hace un breve recorrido por la situación en los Estados Miembros y, en definitiva, sienta las bases para la redacción de la directiva de reutilización que regule este aspecto.

No obstante, cuando hablamos de normativa tenemos que referirnos a los siguientes textos:

- Directiva 2003/98/CE del Parlamento Europeo y del Consejo: este documento destaca la necesidad de consensuar las experiencias y la normativa entre los Estados Miembros en lo que respecta a la explotación de la información del sector público. De esta forma se fomenta la reutilización de este tipo de información y el desarrollo del sector de los contenidos digitales a partir de la información del sector público.
- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público: esta ley transpone la Directiva comunitaria 2003/98/CE al ordenamiento jurídico español. Se centra, al igual que la directiva, en definir los principios básicos para

que las Administraciones Públicas puedan ofrecer la información que producen o custodian y sobre la que tienen los respectivos derechos.

- Real Decreto 1495/2011, de 24 de octubre, que desarrolla la Ley 37/2007: incorpora importantes mejoras que buscan unificar la normativa entre los países miembros e impulsar el desarrollo de la reutilización del sector público a través de la adaptación del marco legislativo.
- Revisión de la Directiva 2003/98 (en proceso de aprobación).

Directiva 2003/98/CE, de 17 de noviembre, relativa a la reutilización de la información del sector público

La Directiva 2003/98/CE fue el resultado de la consolidación normativa de las políticas comunitarias para el establecimiento de un marco legal que fomente la reutilización de la información del sector público y su explotación comercial.

En la directiva se persigue el cumplimiento de una serie de objetivos, tal y como se describe en el desarrollo de la misma. Estos objetivos son:

- Definir el marco de condiciones para que las Administraciones Públicas pongan a disposición de los agentes reutilizadores y de los ciudadanos la información susceptible de ser reutilizada, incluyendo detalles como la definición de formatos, procedimientos de obtención de información, etc.
- Establecer las indicaciones para definir una adecuada política de tarifas y precios. En esta primera directiva la Unión Europea recomienda el establecer una política de precios gratuita o, en su caso, que la tarifa permita cubrir los costes de producción y facturación.
- Evitar cualquier tipo de acuerdo de carácter exclusivo entre una Administración Pública y la empresa que esté reutilizando su información, evitando así situaciones de monopolio.
- Crear mecanismos para el impulso a la reutilización de cara a hacer posible el aprovechamiento de la misma por parte de terceros.

A través de cinco capítulos, la directiva define el marco de actuación donde establece el concepto de reutilización y de documento reutilizable ya en el primer capítulo. En el Capítulo II se establecen los requisitos y procedimientos para el tratamiento de solicitudes de información reutilizable. El Capítulo III destaca las condiciones relativas a diversos temas necesarios para la reutilización: más concretamente, formatos bajo los que la información reutilizable debe ser presentada, la política tarifaria que define el importe que se les cobrará a los agentes reutilizadores, el establecimiento de licencias tipo a las cuales estará sujeto el documento reutilizable y, por último, la necesidad de crear dispositivos de búsqueda. No obstante este punto no ofrece la profundidad que debiera ya que no define en qué consisten los dispositivos de búsqueda pudiendo limitarse estos a un mero listado de los documentos reutilizables. En el Capítulo IV se destaca un elemento que intenta fomentar la libre competencia y que ha supuesto muchas actuaciones por parte de la Comunidad Europea para asegurarlo y es la no discriminación y el aseguramiento de ausencia de acuerdos exclusivos. Por último, en las disposiciones finales, Capítulo V, establece la necesidad de trasponer esta directiva al ordenamiento jurídico de cada uno de los Estados Miembros y de revisar la aplicación de la misma.

El inconveniente (importante) de esta directiva es que, pese a ser de obligado cumplimiento, establece en su articulado que solo las propias administraciones deciden qué información desean poner a disposición del público y cuál no. De esta forma, cualquier iniciativa depende de la buena voluntad de las entidades públicas para tener éxito. Este punto ha evolucionado a lo largo del tiempo como veremos más adelante.

Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público

La Ley 37/2007, del 16 de noviembre, sobre reutilización de la información del sector público, traspone fielmente la Directiva 2003/98/CE regulando el marco general en el cual se definen las condiciones mínimas de reutilización de la información del sector público en España. A tal

efecto define reutilización como el uso de documentos (concepto heredado de la Directiva 2003/98/CE) que han sido previamente elaborados y custodiados por las Administraciones Públicas en el ejercicio de sus funciones, por parte de personas físicas y jurídicas, con fines comerciales o no comerciales que no tienen que coincidir con el propósito inicial para el cual estos documentos fueron elaborados. En este aspecto es importante indicar que el intercambio de información entre administraciones no se considera reutilización.

Los aspectos más importantes de la transposición de la directiva son los siguientes:

- La presente ley, tal y como se describe en su preámbulo, tiene como objeto la armonización de la explotación de la información del sector público, especialmente aquella que se encuentre en soporte digital. De esta forma, de todos los documentos de libre disposición que posee el sector público, se fomentará la publicidad así como la creación de productos y servicios de información.
- En esta ley se establece que cada entidad pública tiene la potestad de autorizar o no la reutilización de sus documentos con fines comerciales y no comerciales. Del mismo modo, de la administración depende el establecimiento de las condiciones a las que esa reutilización estará sujeta en lo que se refiera al establecimiento de licencias de uso o a la reutilización libre de condiciones.
- Además, la ley establece la posibilidad de que la Administración solicite una contraprestación económica por la descarga/uso del documento, no obstante se establece que el importe de esta contraprestación no podrá superar los costes totales de recogida, producción, reproducción y difusión de los documentos.

En esta ley se estipula que corresponde a los poderes públicos definir el punto en el que se establece el límite de los ciudadanos a acceder y reutilizar información del sector público, no obstante con el tiempo, como veremos más adelante, se ha detectado que este límite supone un obstáculo para el crecimiento económico de la Unión Europea.

Respecto a la estructura de la ley, está compuesta por tres títulos que desarrollan once artículos. No entraremos en profundidad en los mismos ya que su contenido ha sido aplicado en los distintos capítulos de este libro, no obstante se puede resumir indicando que el Título I recoge el objeto y el ámbito de aplicación, definiendo “documento” como el elemento jurídico a reutilizar (independientemente de que técnicamente se trate de un documento en un formato abierto o un dato obtenido a través de un servicio en línea). Del mismo modo se establece, tal y como describimos en el Capítulo 3, las limitaciones en la reutilización de documentos relativas a aspectos como la intimidad, propiedad intelectual, etc.

En el Título II se detallan los aspectos básicos del régimen jurídico de la reutilización en lo que se refiere a la expedición de licencias, junto con sus condiciones. Este punto también ha evolucionado a lo largo del tiempo, como veremos en las normativas aprobadas con posterioridad. En este título se indica la forma en la que se deben reutilizar documentos de carácter personal (remitiendo a la Ley 15/1999 de Protección de Datos de Carácter Personal). Adicionalmente, se indica que los documentos deben ser reutilizados a través de medios electrónicos y en los formatos en los que existan. Por último, se habla sobre la posibilidad de aplicación de tasas por el suministro de la información indicando que, de una forma o de otra, estas tasas y precios públicos deben estar publicados y accesibles a los reutilizadores.

Por último, el Título III establece el procedimiento para arbitrar las solicitudes de reutilización, estableciendo los plazos de respuesta y, en su caso, el régimen sancionador. No obstante la aplicación de este título de la ley ha resultado inútil teniendo en cuenta que es la administración la que tiene potestad de decidir qué información se reutiliza y cuál no. Afortunadamente la evolución de la normativa ha resuelto este tema.

Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007

Esta normativa publicada el 8 de noviembre de 2011 desarrolla la Ley 37/2007, que describíamos anteriormente, sobre la reutilización de la información del sector público.

En ella se establecen recomendaciones y acciones de obligado cumplimiento para que las Administraciones Públicas puedan adaptar, publicar y mantener su información reutilizable. De esta forma, sus puntos más destacables son los siguientes:

- El Real Decreto define con precisión cuáles son las obligaciones de las entidades públicas en aras de promover la reutilización. Uno de los puntos clave es que se establece que todos los documentos del sector público pasan a ser reutilizables (a no ser que estén sujetos a alguna condición de exclusión), con lo cual la potestad de reutilizar ya no recae en una administración concreta. Además se deberá publicar un catálogo de documentos reutilizables definiendo también un responsable de coordinar las tareas de reutilización dentro del organismo.
- Se establece como catálogo oficial de información pública el catálogo creado en el portal datos.gov, que hemos descrito exhaustivamente en el Capítulo 6. Además, se establecen obligaciones de actualización del mismo por parte de administraciones y organismos. Hay que indicar que actualmente esta actualización no se está realizando de forma adecuada, no obstante no ha terminado el plazo para realizar estas tareas.
- El Real Decreto regula las modalidades de reutilización. Se establece como norma general la puesta a disposición de la información sin condiciones específicas, pero permitiendo a los organismos establecer condiciones cuando sí sean necesarias.
- Por último, establece la regulación del régimen aplicable a documentos reutilizables sujetos a derechos de propiedad intelectual, industrial o que contengan datos personales.

Respecto a su estructura, el Real Decreto está compuesto por cuatro capítulos con once artículos. El Capítulo I establece el objeto y ámbito de aplicación, que coincide con el objeto y ámbito de aplicación de la ley a la que desarrolla, acotando el ámbito de aplicación al sector público estatal. En el Capítulo II se define el régimen jurídico y organizativo indicando que todos los documentos serán por defecto reutilizables, la forma en

la que serán puestos a disposición vía medios electrónicos y que, cuando esto no sea posible, será necesaria la puesta a disposición bajo demanda. Se define la necesidad de creación de un órgano para la coordinación de las actividades de reutilización definiendo las labores que le corresponden y se establece el catálogo datos.gob como el catálogo oficial de información pública indicando las responsabilidades de actualización y mantenimiento. En el Capítulo III se establecen las modalidades de reutilización definiendo las condiciones generales para la puesta a disposición de la información (estas modalidades las describiremos un poco más adelante). Por último, el Capítulo IV detalla el régimen aplicable a documentos sujetos a derechos de propiedad intelectual o que contengan datos personales, especificando la forma en la que deben ser puestos a disposición por parte de las entidades públicas.

Las mejoras que incorpora este Real Decreto son realmente diferenciales, ahora todo es reutilizable y además se tiene que disponer de una persona física identificada a quien acudir cuando haya necesidad de solventar cualquier duda, petición o reclamación relacionada con la reutilización de información en esa entidad local. La implicación de estos cambios para la gestión de una empresa es inmensa.

Adicionalmente se indica, y esto es importante, que en el momento que se pone el documento a disposición de los reutilizadores para su uso, se están cediendo los derechos necesarios para desarrollar cualquier actividad de reutilización autorizada de forma gratuita, y no exclusiva.

Se establecen plazos concretos para aplicar las indicaciones definidas en el Real Decreto. Así, las entidades públicas deben tener aprobado, en el plazo de un año desde la firma del Real Decreto, su plan de reutilización, que define la realización de todos los cambios organizativos y procedimentales para fomentar la reutilización, y adicionalmente deberán poner en marcha las tareas y medios necesarios para la publicación de, al menos, cuatro conjuntos de datos importantes en los seis meses siguientes al anterior plazo.

Revisión de la Directiva 2003/98 (pendiente de aprobación)

Aunque esta revisión no está definitivamente aprobada en el momento de publicar este libro, hemos considerado necesario destacarla por el avanzado estado de tramitación y los cambios que implica.

Esta revisión introduce una serie de modificaciones a los artículos presentados en la Directiva 2003/98 de las cuales destacaremos las más importantes.

En primer lugar, y siguiendo la estela de lo que ya se había legislado en España a través del Real Decreto 1495/2011, establece la obligación para todos los Estados Miembros de autorizar la reutilización de todos los documentos de libre disposición, siempre y cuando el fin de estos documentos sea la reutilización de la información en los términos que se establece en la directiva original. Del mismo modo, el organismo público puede establecer condiciones a esta reutilización como el reconocimiento de la fuente, pero nunca serán un freno al impulso de la reutilización de la información del sector público.

El alcance de los documentos susceptibles de ser reutilizados se amplía a museos, archivos y bibliotecas. Además, tanto para estas fuentes como para el resto de las que se han definido en la directiva original, la entidad pública debe poner a disposición la información en formatos legibles por una máquina junto con sus metadatos que faciliten su interoperabilidad. Asimismo podrá establecer una tasa por el servicio, pero es obligación del organismo el justificar la cuantificación de la tasa establecida.

Por último, un punto que es muy importante, ya que elimina una posible barrera a la reutilización, es la existencia de acuerdos exclusivos entre una entidad pública y una determinada empresa, prohibiendo que se produzcan y estableciendo un plazo de expiración para los existentes.

Como se puede observar, poco a poco se avanza hacia una liberalización estableciendo condiciones más ventajosas para la reutilización, la evolución a lo largo de estos años es palpable a través de las nuevas normativas aprobadas. Hoy en día, las oportunidades son inmensas,

pero aún quedan temas pendientes sobre los que hay que definir una legislación adecuada como veremos en el siguiente apartado.

7_2 Modalidades de reutilización y licencias aplicables

Como anteriormente comentábamos, estas modalidades de reutilización están definidas en el Capítulo III del Real Decreto 1495/2011. Ahí se definen unas condiciones de uso general para cualquier modalidad de puesta a disposición de los documentos reutilizables, estas condiciones son:

- No desnaturalizar el sentido de la información.
- Citar la fuente de los documentos.
- Mencionar la fecha de la última actualización, siempre que esté incluida en el documento original (es responsabilidad del organismo el mantener la información actualizada).
- No vincular o sugerir que los organismos públicos titulares de la información participan, patrocinan o apoyan la reutilización que se lleve a cabo con ella.
- Conservar y no alterar ni suprimir los metadatos relativos a la fecha de actualización y a las condiciones aplicables a la dicha reutilización.

Estas condiciones estarán accesibles por medios electrónicos y vincularán a cualquier agente reutilizador por el mero hecho de hacer uso de los documentos.

Las modalidades de reutilización han sido descritas en el artículo 8, Modalidades de puesta a disposición de los documentos reutilizables, no obstante esta normativa ha sido analizada y ordenada por el Decano de la Facultad de Documentación de la Universidad Complutense de Madrid D. Fernando Ramos Simón en su artículo “Licencias abiertas y de pago en el nuevo contexto RISP” (todavía no publicado):

- 1. Modalidad general básica:** esta modalidad es la impulsada por defecto en el Real Decreto y únicamente requiere el cumplimiento de las condiciones de reutilización descritas anteriormente debiendo cumplir

adicionalmente los principios de transparencia y no discriminación. A través del uso de este tipo de licencias más conocidas como “licencias abiertas” se permite al agente reutilizador el acceso, uso, reutilización y distribución de la misma. Este tipo de licencia fomenta los principios del Open data en su estado puro fomentando la puesta a disposición de contenidos de forma libre. Sin duda, es el modelo que de forma más clara fomenta la competitividad frente al proteccionismo.

2. Licencias-tipo: para esta modalidad el Real Decreto se ajusta a lo previsto en la Ley 37/2007, en la cual se define que las licencias deben estar disponibles en formato digital. Se definen un conjunto de modalidades de licencias-tipo que se describen a continuación:

a. Licencias-tipo “libres”: entre este modelo de licencias la más recomendada y utilizada dentro del sector público es la licencia EUPL (Licencia Pública de la Unión Europea), básicamente el modelo es el mismo que el de las licencias de modalidad general básica, pero se incluyen adicionalmente restricciones referentes a la creación de versiones modificadas o derivadas. Estas restricciones suelen consistir simplemente en que la obra derivada que se cree esté sujeta a las mismas condiciones que la obra original, básicamente consiste en no limitar el alcance de la licencia a la obra reutilizada sino a cualquier producto que se construya a partir de él, lo cual tiene todo el sentido del mundo. Por ejemplo, si se está reutilizando una gráfica con información económica, debería citarse la fuente en cualquier obra derivada y, además, no se debería desnaturalizar el sentido de la información.

b. Licencias-tipo específicas: en este modelo el Real Decreto hace referencia a la Ley 37/2007 en la que se establece que las licencias que se elaboren en un organismo sean uniformes en su tipología. Además, en su artículo 9 especifica que la licencia debe incorporar, al menos, la información relativa a la finalidad concreta de la licencia, indicando si el uso va a ser comercial o no comercial, los motivos para los que se concede la reutilización así como la duración de la misma y las obligaciones de todas las partes (tanto

del reutilizador como del organismo que concede la licencia). Se definirán las responsabilidades de uso y las modalidades financieras, indicándose el carácter gratuito de la misma o, en su caso, la tasa o precio público aplicable.

3. Reutilización previa solicitud: es la modalidad que se aplica a todos aquellos casos en los que, para poder reutilizar una información, es necesario una gestión y procesamiento de la solicitud por parte de la entidad pública (por ejemplo, en la copia de planos en papel en el Instituto Geográfico Nacional). En general, se aplica a todos los documentos que no se encuentren en formato digital. No obstante es obligatorio que el organismo público tenga disponible en su página web la información referente a las tasas o precios públicos aplicables para los documentos reutilizables bajo esta modalidad.

La pregunta que surge a cualquier gestor de un organismo público es cuál es la modalidad de licencia más adecuada. ¿Se debe cobrar o no cobrar por el servicio? Tenemos ejemplos de todos los tipos, el INM (Instituto Nacional de Meteorología) lleva años llevando una política de servicios basada en tasas y precios públicos que resulta muy beneficiosa para este organismo ya que su información es muy valiosa, ningún otro la ofrece y actualiza como ellos y supone una importante fuente de ingresos. Por otro lado, la Dirección General del Catastro, ofreciendo la información de forma totalmente gratuita, ha fomentado la creación de empleo y negocios (alguno de los cuales ha sido nombrado en este libro), y al mismo tiempo ha generado un ahorro en costes de personal al reducirse el número de personas que acuden de forma presencial a realizar gestiones.

A nivel europeo tenemos un ejemplo claro, la *Danish Enterprise and Construction Authority* (DECA) redujo las tasas de los servicios que ofrecía de forma telemática de forma drástica lo que supuso un incremento en el número de reutilizadores de un 10.000 % (no es una errata, realmente hablamos de un diez mil por cien), fomentando el crecimiento del mercado de reutilización en un 1.000 % a lo largo de ocho años. El retorno en

impuestos para el Gobierno se ha estimado que es cuatro veces mayor que el volumen de ingresos antiguo proveniente de tasas y precios públicos por estos servicios ahora gratuitos. El ejemplo es extremo, pero es muy claro, las oportunidades están ahí.

Hemos visto que el Real Decreto, incluso la revisión de la directiva, incorpora cambios y mejoras en las licencias, no obstante somos de la opinión que para fomentar la reutilización sería necesario crear un único modelo de licencia a nivel europeo. De esta forma se podría facilitar e impulsar la reutilización entre Estados Miembros mezclando datos que de otra forma estarían sujetos a las licencias específicas de cada uno de ellos. Esto no significaría que la información tuviera que ser gratuita, sino que puede estar sujeta al pago de tasas o a un precio público pero con unas condiciones generales uniformes que fomenten la interoperabilidad entre todos los agentes.

8_Reutilización e innovación



A lo largo del libro hemos descrito todos los aspectos importantes relativos a la reutilización desde el punto de vista del emprendedor, la cuestión es: ¿cómo se puede acometer una empresa de estas características?, ¿cuál es el procedimiento que se debe seguir para conseguir lanzar una iniciativa empresarial de éxito? Desafortunadamente no hay una metodología claramente definida para este tipo de negocio y, aunque la hubiera, es obvio que no tiene garantías de éxito.

La Open Data Foundation establece entre sus principios para el Open data que debería ser sencillo realizar tareas como:

- Descubrir la existencia de los datos (en nuestro caso disponemos del repositorio de datos del portal <http://datos.gob.es>).
- Acceder a los datos para realizar investigación, análisis y reutilización.
- Encontrar información detallada que describa los datos y sus procesos de producción.
- Mantener una comunicación efectiva con las agencias involucradas en la producción, almacenamiento y distribución de los datos.
- Compartir conocimiento con otros usuarios.

Las necesidades derivadas de estos principios están cubiertas por la información que se puede obtener a partir de los datos que ofrecen las administraciones públicas para reutilizar (y hemos destacado en varias ocasiones que, en caso de necesitar información adicional por parte de un organismo, existen procedimientos para solicitarla). Así, la realización de estas tareas nos puede proporcionar la información necesaria para poder acometer la definición de un modelo de negocio.

Tal y como hemos comentado a lo largo de la obra, el infomediario es un sector que basa su modelo de negocio en la reutilización de información. Obtener información y aportarle valor creando algún tipo de producto o servicio a partir del que obtener un beneficio es, a grandes rasgos, su operativa básica.

Por ello, lo que es un hecho es que las empresas infomediarias presentan un perfil tecnológico elevado, una de las características de estas empresas es el uso de la

innovación de cara a realizar un eficiente análisis, gestión y tratamiento y presentación de la información reutilizada.

Algunos de los puntos que debiéramos tener en cuenta de cara a diseñar una empresa infomediaria son los siguientes:

- Definir la actividad de la empresa infomediaria.
- Establecer la organización y ubicación de la empresa e identificar socios estratégicos.
- Realizar una caracterización de los clientes y sus necesidades.
- Identificar las fuentes de los datos.
- Definir estrategias de comercialización.
- Realizar un plan de negocio definiendo recursos iniciales, retos y plazos para alcanzarlos y beneficios esperados.

Algunos de estos puntos no son específicos de una empresa infomediaria, por lo que considero que ya hay infinidad de títulos en el mercado que ofrecen el conocimiento necesario para acometer estas ideas. Por ello vamos a centrarnos en los puntos más relacionados con la reutilización de datos *per se*.

8_1 Definir la actividad de una empresa infomediaria

No importa si somos una empresa de nueva creación o si se trata de una empresa consolidada que quiere acometer nuevas iniciativas de negocio. El primer paso consiste en poder definir cuál va a ser la actividad de la empresa infomediaria. Es obvio, tenemos que saber qué queremos hacer.

De esta forma tenemos varias alternativas:

1. Identificar una necesidad que pueda solventarse con datos reutilizables.
2. Investigar los datos buscando una oportunidad de negocio.
3. Trasladar un modelo de negocio consolidado en otro país a la realidad del nuestro.

Salta a la vista que la primera de ellas es la más complicada de acometer a no ser que de una forma u otra te enfrentes directamente a esa necesidad. Sería encontrar

lo que W. Chan Kim²² llama un “océano azul”, un modelo de empresa en la que la competencia no tenga demasiada importancia, porque está claro que si hay una necesidad es porque no hay una solución que la satisfaga, o bien porque la solución que se ofrece no abarca las necesidades de todos los clientes y esto supone una nueva oportunidad de negocio.

La segunda opción implica un importante trabajo de campo no exento de dificultad, ya que tenemos que partir de catálogos de datos para identificar de qué forma estos datos nos pueden beneficiar. Esto puede resultar útil cuando ya tenemos una empresa o trabajamos en un sector que se dedica a una área comercial concreta que conocemos y, analizando los datos reutilizables en esa área, podemos identificar oportunidades de negocio.

Vamos a ver cómo tendríamos que acometer la tercera opción. Pongamos el caso de un ejemplo que se ha citado en este libro como un modelo de negocio en Reino Unido pionero en reutilización de datos del sector público (aunque hoy en día haya sido adquirida por otro portal): la web www.upmystreet.com.

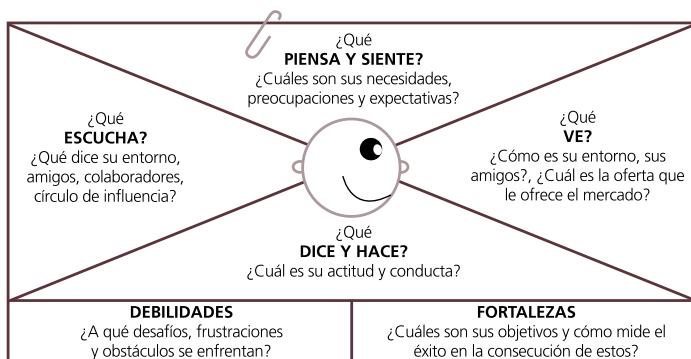
Esta página web recoge información del sector público que posteriormente georreferencia en GoogleMaps de forma que, a través de un clic, un usuario puede obtener información sobre colegios, parques, tasa de criminalidad, hospitales, nota media de selectividad, etc. De esta forma una persona que quiere adquirir un inmueble puede obtener toda esa información de su posible futura dirección. De manera natural, las empresas inmobiliarias comenzaron a utilizar la web como una herramienta de fuerza de ventas que les permitía ofrecer a sus clientes información de valor añadido que podía ayudarles a tomar la decisión sobre su compra. Así, este modelo de negocio consolidado en Reino Unido puede ser trasladado a otro país como, por ejemplo, España.

8_2 Realizar la caracterización de los clientes

Siguiendo con el ejemplo anterior, en este momento conocemos qué es lo que queremos hacer, pero puede ser que no tengamos muy claro cuál es la orientación que

²²Chan Kim, W. y Mauborgne, R. (2005).

Figura 8.1. El mapa de empatía. Construyendo nuestro modelo de negocio desde la visión del cliente.



Fuente: Adaptado de Xplane.

le queremos dar. No hay que olvidar que se trata de un negocio creado para un mercado anglosajón y no tiene por qué encajar en el modelo español.

Necesitamos definir un modelo de negocio diferente. El mapa de empatía es una herramienta desarrollada por la compañía Xplane (ahora Dachis Group), que nos ayuda a entender mejor a nuestro futuro cliente. A través de un conocimiento más profundo de su entorno y la visión de sus necesidades y problemas, nos ofrece una visión única del mundo y de sus propias necesidades.

A través de esta herramienta de diseño creativo se tiene que responder a seis preguntas.

1. ¿Qué escucha el cliente en su entorno referente?
 - “Qué difícil es encontrar colegio para los niños...”.
 - “Me gustaría saber qué hospitales y centros de salud hay cerca”.
 - “Ha habido un montón de atracos en mi barrio”.
 - ...
2. ¿Qué piensa y qué siente referente al problema que nos ocupa?
 - “Necesito saber si los colegios para mis hijos tienen la calidad necesaria”.

- “Tengo asma, ¿cómo es el aire en mi zona?”.
- ...
- 3. ¿Qué ve el cliente?
 - “Me gusta ir al cine los sábados”.
 - “Nos gusta salir en bicicleta los fines de semana”.
 - “La gasolina es muy cara cerca de mi casa”.
 - ...
- 4. ¿Qué dice y qué hace el cliente?
 - “Tengo que ir a comprar medicinas y es tarde”.
 - “Me cuesta pagar los gastos escolares de mis hijos”.
 - ...
- 5. Debilidades:
 - “El tráfico es horrible los lunes por la mañana”.
 - ...
- 6. Fortalezas:
 - “Me gusta que haya parques cerca de mi casa”.
 - ...

8_3 Definir las fuentes de los datos

A partir de las preguntas anteriores podemos detectar las necesidades de los posibles clientes e identificar las fuentes de información que nos den respuesta a ellas.

Para el ejemplo que nos ocupa vamos a escoger alguna de las preguntas, habría que resolverla para todas las ideas y necesidades que se han identificado en el modelo de empatía:

1. “Qué difícil es encontrar colegio para los niños...”, requiere el listado de los colegios en tu localidad (pongamos Madrid).
2. “Me gustaría saber qué hospitales y centros de salud hay cerca”, requiere acceso al listado de hospitales y centros de salud en Madrid.
3. “Ha habido un montón de atracos en mi barrio”, requiere información de índices de criminalidad.
4. “Necesito saber si los colegios para mis hijos tienen la calidad necesaria”, requiere información sobre la nota de selectividad y otras calificaciones como los resultados de las pruebas del CDI, informe PISA, etc.

5. “La gasolina es muy cara cerca de mi casa”, necesita los precios de las gasolineras cercanas al domicilio.
6. “Tengo que ir a comprar medicinas y es tarde”, se necesita saber cuáles son las farmacias de guardia en la zona.
7. “El tráfico es horrible los lunes por la mañana”, se necesita la información del tráfico actualizada.

Es importante resaltar que este proceso no es trivial, requiere de la participación de varias personas trabajando en equipo para que sea efectivo y, además, debe ser un proceso iterativo más que secuencial.

Identificadas nuestras oportunidades en este momento tenemos que buscar esas fuentes de datos. Así, a continuación exponemos una fuente de datos para cada necesidad formulada:

1. Respecto a la primera, encontramos una web (http://gestiona.madrid.org/wpad_pub/run/j/Inicio.icm) en la que se pueden consultar los centros educativos en toda la comunidad de Madrid e incluso se puede hacer una consulta por proximidad al domicilio.
2. Haciendo una búsqueda en Internet, encontramos <http://centrossanitarios.sanidadmadrid.org/RedAsistencial/CentrosAsistenciales.aspx>, donde tenemos acceso a todos los hospitales y centros de salud de la comunidad de Madrid.
3. Tras hacer una búsqueda por Internet no encontramos la información que buscamos, estos datos pertenecen al Ministerio del Interior, decidimos enviar una solicitud de información de reutilización.
4. Localizamos el portal www.madrid.org. En el buscador que se nos muestra a la derecha buscamos “resultados CDI” y el primer resultado de la búsqueda nos lleva a una página donde podemos descargar la información que necesitamos en formato pdf.
5. Accedemos al portal del Ministerio de Industria, Energía y Turismo, donde encontramos <http://geoportal.mityc.es/hidrocarburos/eess> que nos permite obtener la información de las gasolineras cercanas a una dirección concreta y sus precios.

6. Accedemos al Colegio Oficial de Farmacéuticos el cual ofrece el servicio de búsqueda de farmacias (www.cofm.es/informacion-corporativa/buscador-farmacias).
7. Desde la DGT podemos observar la información del tráfico actual junto con las incidencias que se vayan produciendo (<http://infocar.dgt.es/etraffic>).

Para todos estos datos, debemos identificar aquellos que nos resultan útiles, aquí nos encontramos con varias casuísticas:

1. Tendremos que identificar la forma en que los necesitamos: si accesibles a través de un servicio o integrados en nuestra base de datos. En el caso de que tan solo necesitemos el acceso a través de un servicio, deberemos conocerlo así como tenerlo identificado, documentado y establecer procedimientos para comprobar su funcionamiento de forma periódica.
2. En aquellos casos en que necesitemos descargar el documento para su gestión y tratamiento, se tratará de automatizar el proceso. Es el ejemplo de la descarga en pdf de los resultados del CDI (Conocimientos y Destrezas Indispensables).
3. Para aquellos datos que no conozcamos (como puede ser el índice de criminalidad), deberemos reclamar la información reutilizable al organismo correspondiente (en este caso al Ministerio del Interior).
Un posible modelo de carta podría ser el que aparece en el Cuadro 8.1. Este modelo de carta sería aplicable a cualquiera de las solicitudes.

8_4 Definir las estrategias de comercialización

A partir de los datos que tenemos, debemos ofrecer un servicio al cliente que le resulte de interés.

Necesitamos desarrollar un visor que nos permita georreferenciar datos geográficos. En este visor debemos mostrar un mapa o capa base para lo cual podemos utilizar los servicios de pago de Google o utilizar el PNOA (Plan Nacional de Ortofotografía Aérea, dependiente del Ministerio de Fomento).

Cuadro 8.1. Modelo de carta de solicitud de reutilización de datos.

Asunto: Solicitud reutilización datos – [*departamento responsable de los datos*]

[*fecha*]

Estimado Sr/Sra. [*nombre de la persona responsable*]:

Nos ponemos en contacto con usted como emprendedores que estamos constituyendo una empresa que ofrece servicios a partir de información pública. Para poder ejercer nuestra actividad requerimos alguna información que está accesible en su web y que debemos descargar/tener acceso a la misma de forma remota.

Es por ello que nos ponemos en contacto con [*departamento responsable de los datos*], con el fin de solicitar la obtención de un acuerdo de licencia no exclusivo que nos permita procesar con nuestra tecnología la información solicitada para su reutilización/obtener el acceso al servicio de publicación de su web, con el fin de agregarla de forma automatizada a nuestros sistemas ampliando los contenidos que ofrece nuestra empresa y así mejorar el servicio que ofrecemos a nuestros clientes.

A estos efectos, nuestras solicitudes de licencia se basan en las disposiciones de la Directiva 2003/98/CE del Parlamento y del Consejo de 17 de noviembre de 2003 relativa a la reutilización de la información del sector público, que, en el caso español, ha sido transpuesta a la legislación española mediante la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, y refrendada con el Real Decreto 1495/2011 que establece en el Artículo 2 la autorización general para la reutilización de los documentos del sector público y puesta a disposición por medios electrónicos.

Quedo a su disposición para cualquier duda o información necesaria acerca de la naturaleza de nuestra solicitud o cualquier otro asunto que requiera de nuestra atención.

A la espera de sus indicaciones acerca del modo de iniciar la tramitación de dicha licencia, aprovecho para enviarle mi más cordial saludo,

Fdo D. XXXXX

A través de servicios web, podremos geolocalizar sobre este mapa toda la información que hemos recuperado y así podremos ofrecer al ciudadano toda la información de la que disponemos, que hemos identificado antes y que responde a sus preguntas y necesidades, identificadas en el apartado anterior.

Adicionalmente, se puede establecer un servicio para particulares registrados en el que puedan consultar la nota de corte para entrar en un determinado colegio concertado en su localidad el año anterior y un formulario para averiguar cuál sería su nota este año y así comprobar sus posibilidades. Teniendo a los usuarios registrados, se les puede ofertar la información publicada segmentada de tal manera que nos sirva como fuente de ingresos.

Además, a empresas de servicios como las inmobiliarias se les puede ofertar la posibilidad de generar informes personalizados y maquetados con la información disponible que elijan para poder adjuntarlo como documentación a la oferta de venta de un inmueble que le hagan a un determinado usuario.

El modelo de negocio está definido, no hemos explotado todas sus posibilidades ya que somos conscientes de que hay muchas más preguntas y respuestas y mucha más información susceptible de ser utilizada. Pero lo que está claro es que a través de esta vía se puede definir un negocio creativo y diferente.

8_5 Realizar un plan de negocio

Disponemos de la información, tenemos acceso a las necesidades del cliente y al modelo de comercialización, lo que queda pendiente es analizar la viabilidad del negocio. Para ello es necesario realizar un plan de negocio al uso. Se puede utilizar una nueva herramienta de desarrollo de negocio llamada *Business Model Canvas* desarrollada por Alexander Osterwalder en su libro *Business Model Generation* (2010) en la cual se identifican todos los aspectos a tener en cuenta en nuestro negocio (véase la Figura 8.2).

Figura 8.2. Business Model Canvas.

Socios clave ¿Quiénes son nuestros socios clave? ¿Quiénes son nuestros suministradores clave? ¿Qué recursos clave vamos a adquirir de nuestros socios? ¿Qué actividades clave realizan los socios?	Actividades clave ¿Qué actividades clave requiere nuestra propuesta de valor? ¿Nuestros canales de distribución? ¿Nuestras relaciones con clientes? ¿Nuestras fuentes de ingresos?	Propuestas de valor ¿Qué valor entregamos al cliente? ¿Cuál de los problemas de nuestro cliente vamos a ayudarle a resolver? ¿Qué paquetes de productos y servicios ofrecemos a cada segmento de cliente? ¿Qué necesidades del cliente estamos satisfaciendo?	Relaciones con clientes ¿Qué tipo de relación espera que establezcamos y mantengamos cada uno de nuestros segmentos de cliente? ¿Cuáles hemos establecido? ¿Cómo están integrados con el resto de nuestro modelo de negocio? ¿Cuánto cuestan?	Segmentos de cliente ¿Para quién estamos creando valor? ¿Quiénes son nuestros clientes más importantes?
Recursos clave ¿Qué recursos clave requiere nuestra propuesta de valor? ¿Nuestros canales de distribución? ¿Relaciones con clientes? ¿Fuentes de ingresos?			Canales ¿A través de qué canales quieren ser contactados nuestros segmentos de clientes? ¿Cómo les contactamos ahora? ¿Cuáles funcionan mejor? ¿Cuáles son más eficaces en costes? ¿Cómo los integramos con las rutinas de cliente?	
Estructura de costes ¿Cuáles son los costes más importantes inherentes a nuestro modelo de negocio? ¿Qué recursos clave son los más caros? ¿Qué actividades clave son las más caras?		Fuentes de ingresos ¿Para qué valor están realmente dispuestos a pagar nuestros clientes? ¿Para qué pagan actualmente? ¿Cómo están pagando ahora? ¿Cómo preferirían pagar? ¿Cuánto contribuye cada fuente de ingresos a los ingresos totales?		

Fuente: Business Model Generation (2010).

Tendremos que rellenar cada uno de los apartados descritos en la plantilla para tener una visión definitiva de nuestro negocio. Aun así este autor considera que es necesario definir adicionalmente un plan de negocio completo en el que se establezca un plan financiero detallado, un plan de marketing completo, etc. No entraremos en la forma o estructura de estas tareas ya que no es objeto de este libro.

Lo que es importante destacar es que las organizaciones deberán tener perfiles con suficiente conocimiento para realizar una recuperación y gestión de datos, así como la capacidad de presentarlos de forma correcta y completa. Es necesario asimismo no centrarnos exclusivamente en las oportunidades, que son muchas, sino tener además visibles las amenazas (como podría ser la no disponibilidad de los datos por parte de la entidad que nos los proporciona).

Por último, no debemos olvidar que se trata de una compañía orientada a la gestión de datos, de ahí que requiera de los perfiles adecuados y de los procesos claramente definidos para poder capturar y aportar valor a los datos reutilizados.

Estamos ante un nuevo modelo de negocio que todavía ofrece muchas oportunidades. El Open data y la información que ofrece requiere de líderes que dirijan la compañía liderando en valores y orientando el talento hacia la consecución de los objetivos teniendo en cuenta las barreras tecnológicas, operativas, funcionales y legales que presenta la consecución de esos objetivos. Pero sobre todo es urgente que los gestores comiencen a identificar oportunidades de negocio, a definir el rol de la información reutilizable como un elemento más de gestión y a identificar los puntos en los cuales puede aportar valor y generar negocio.

Bibliografía



- Asimov, I. (1980), *Momentos estelares de la ciencia*. Alianza Editorial, Madrid.
- Chan Kim, W. y Mauborgne R. (2005), *la estrategia del océano azul*. Harvard Business School Press, Boston.
- Comisión Europea (1998), *Libro Verde sobre la información del Sector Público en la Sociedad de la Información* ([ftp://ftp.cordis.lu/pub/econtent/docs/gp_es.pdf](http://ftp.cordis.lu/pub/econtent/docs/gp_es.pdf)).
- Dekkers, M.; Polman, F.; te Velde, R. y de Vries, M. (2006), Estudio MEPSIR "Measuring European Public Sector Information Resources" (http://ec.europa.eu/information_society/policy/psi/mepsir/index_en.htm).
- Everis (2006), *Estudio sobre la situación en España de la Reutilización de la Información del Sector Público*. Everis, Madrid.
- García Manjón, J. V. (2010), *Gestión de la innovación empresarial*. Netbiblo, A Coruña.
- Hagel III, J. y Singer, M. (1999), *Net Worth*. Harvard Business School Press, Boston.
- Halliday, S. (2010), "Florence Nightingale", en *BBC History Magazine*, nº 8, agosto.
- Ministerio de Industria, Energía y Turismo y Ministerio de Hacienda y Administraciones Públicas (2009), *Guía Aporta sobre reutilización de la información del sector público*. Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información, Madrid.
- Ministerio de Industria, Energía y Turismo y Ministerio de Hacienda y Administraciones (2012), *Estudio de caracterización del Sector Infomediario*, Madrid (www.ontsi.red.es/ontsi/es/estudios-informes/estudio-de-caracterizaci%C3%B3n-del-sector-infomediario-en-espa%C3%B1a-edici%C3%B3n-2012).
- Osterwalder, A. y Peigner, Y. (2010), *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. John Wiley & Sons Ltd.
- Pira International (2000), *Commercial exploitation of Europe's public sector information-Informe Pira*. Surrey, England (http://ec.europa.eu/information_society/policy/psi/docs/pdfs/pira_study/commercial_final_report.pdf).
- Ramos Simón, F. (2012), "Licencias abiertas y de pago en el nuevo contexto RISP" (pendiente de publicación).

Normativa

Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, de julio de 2009, sobre Revisión de la Directiva 2003/98/CE de reutilización de la información del sector público (http://ec.europa.eu/information_society/policy/psi/docs/pdfs/openda-ta2012/open_data_communication/es.pdf).

Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público (www.boe.es/aeboe/consultas/bases_datos/doc.php?id=DOUE-L-2003-82244).

Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. Transpone la Directiva 2003/98/CE, de 17 de noviembre de 2003. (www.boe.es/aeboe/consultas/bases_datos/act.php?id=BOE-A-2007-19814).

Ministerio de Industria, Energía y Turismo y Ministerio de Hacienda y Administraciones (2011), *Estudio de caracterización del Sector Infomediario*. Madrid (www.ontsi.red.es/ontsi/es/estudios-informes/estudio-de-caracterizaci%C3%B3n-del-sector-infomediario-junio-2011).

Ministerio de Industria, Energía y Turismo y Ministerio de Hacienda y Administraciones Públicas (2012), *Guía de aplicación del Real Decreto 1495/2011*, Madrid (<http://datos.gob.es/datos/sites/default/files/PLAN-CISP-GRD-07.3.4.doc>).

Propuesta de la Directiva 2003/98/CE de reutilización de la información del sector público, de diciembre de 2011 (http://ec.europa.eu/information_society/policy/psi/docs/pdfs/directive_proposal/2012/es.pdf).

Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal (www.boe.es/aeboe/consultas/bases_datos/doc.php?id=BOE-A-2011-17560).

Pocket Innova

Gestión de la innovación

AUTOR: Juan Vicente García Manjón
ISBN: 978-84-9745-477-3

Gestión del conocimiento

AUTOR: Monserrat Santillán de la Peña
ISBN: 978-84-9745-481-0

Web 2.0

AUTOR: José Luis Marín de la Iglesia
ISBN: 978-84-9745-483-4

Biotechnología

AUTOR: Juan P. Duque
ISBN: 978-84-9745-485-8

Proveedores de conocimiento

AUTOR: Javier González Sabater
ISBN: 978-84-9745-489-6

Hazlo distinto

AUTOR: Santiago Sousa Carreira
ISBN: 978-84-9745-484-1

El ABC de la innovación

AUTORES: Juan Vicente García Manjón
Javier Alfonso Rodríguez Escobar
ISBN: 978-84-9745-492-6

Mercado Alternativo Bursátil

AUTORES: David Carro Meana
Paula Veloso Pereira
ISBN: 978-84-9745-559-6

La actitud innovadora

AUTOR: Antonio Flores
ISBN: 978-84-9745-360-8

Business Angels

AUTORES: Pablo Martínez García
José Gabriel García Ortega
ISBN: 978-84-9745-494-0

Equipos innovadores

AUTOR: Pablo Villanueva Alonso
ISBN: 978-84-9745-558-9

Los activos intangibles y sus retos

AUTOR: Borja Barrutieta
ISBN: 978-84-9745-486-5

Fiscalidad de la I+D+i

AUTOR: Felipe Alonso Murillo
ISBN: 978-84-9745-488-9

I+D+i en Europa

AUTOR: Manon van Leeuwen
ISBN: 978-84-9745-490-2

Start-ups

AUTOR: Isidre March
ISBN: 978-84-9745-458-2

Clusters

AUTOR: Javier García Díez
ISBN: 978-84-9745-572-5

Las 3R de su negocio

AUTOR: Fernando Sáenz-Marrero
ISBN: 978-84-9745-570-2

La empresa conectada

AUTOR: David Ruiz Uceta
ISBN: 978-84-9745-561-9

i-Economía

AUTORES: Javier García
Paco Prieto
Pablo Priesca
ISBN: 978-84-9745-582-4

Eco-innovación

AUTORES: Javier Carrillo
Pablo del Río
Totti Könnölä
ISBN: 978-84-9745-576-3

Web Semántica

AUTOR: Jose Emilio Labra Gayo
ISBN: 978-84-9745-571-8

Internet de los objetos

AUTORES: Juan Manuel Cueva Lovelle
Jordán Pascual Espada
Óscar Sanjuán Martínez
B. Cristina Pelayo García-Bustelo
ISBN: 978-84-9745-476-6

Liderazgo e innovación 2.0

AUTOR: Virgilio Gallardo (Coord.)
ISBN: 978-84-9745-577-0

Para más información visite:
www.netbiblo.com

