





JAVIER GARCÍA-ARÁEZ MARTÍN-MONTALVO / JUAN JOSÉ GENERELO LANASPA

La protección de datos del usuario de archivos por Internet: del expediente de investigador a la analítica web

Digamos algo obvio: Internet es un potente factor de cambio y ha supuesto incontables novedades en los planteamientos personales, de ocio y laborales de todos los ciudadanos que lo incorporan como herramienta de uso habitual. Su utilización es a día de hoy tan habitual que difícilmente podríamos etiquetar a la Red de redes como “nueva tecnología de la información”. Para una mayoría creciente de la población Internet es, simplemente, información. Una herramienta de uso diario.

Afortunadamente, servicios de información y documentación tradicionalmente presenciales –como es el caso de archivos– han sabido, paulatinamente, fabricarse un nicho en el que desarrollarse y tener una presencia visible. Con gran frecuencia aparecen versiones *online* o herramientas paralelas a estos servicios de información que hacen de Internet y sus posibilidades casi ilimitadas de acceso un poderoso aliado para el cumplimiento de sus objetivos institucionales.

La naturaleza del usuario de archivo, por tanto, ha cambiado. Si hasta fines del siglo pasado cabía hablar de dos tipos básicos de *clientes* de los archivos, el usuario-investigador –el que realiza un estudio más o menos académico– y del usuario-ciudadano –el que acredita defender un derecho a partir de la información de los archivos ante la administración o ante terceros– ahora podemos

hablar cada vez más del usuario-internauta, que podemos definir como aquel que reclama información de cualquier fuente del tipo que sea –y en ellas se incluyen los archivos– para todo tipo de fines que, a su vez, no está obligado a revelar.

Del uso de las nuevas tecnologías surge una nueva realidad, cada día más presente: el que fuera usuario presencial de archivo –previa presentación en su día de la tarjeta nacional de investigador o actualmente del DNI– puede convertirse ahora en un elemento anónimo que, cómodamente instalado en su casa, puesto de trabajo o incluso sentado en un parque, accede a la documentación albergada en cualquier archivo, consulta un expediente o realiza incluso una reproducción del mismo.

Usuarios presenciales y usuarios-internautas

Ahora bien, en el archivo “tradicional” o “físico”, en oposición al archivo virtual que sale a Internet, el usuario realiza sus consultas, normalmente de documentos originales, a través del contacto directo y personal con sus trabajadores. Su visita es controlada desde su inicio hasta su fin –como vimos, a través de la tarjeta de investigador o a través de su DNI– lo que permite saber qué consulta, cuándo lo consulta e incluso en muchas ocasiones por qué consulta una determinada documentación.

En el ámbito *online* no se da esta circunstancia. Existen diferencias muy importantes entre el usuario presencial del archivo y el usuario-internauta que accede a su faceta online. Podemos indicar, brevemente, las características principales de ambos perfiles:

- a. El usuario presencial será siempre una persona física de la que tendremos datos confiables. Se sabrá con seguridad su nombre y DNI y es perfectamente posible conocer su profesión o actividad diaria. También es posible saber cuál es su motivación para acceder a un servicio de archivo determinado.

Este tipo de usuarios tiene un contacto directo con el personal de archivo y con las herramientas habituales de recuperación de información que se ponen a su disposición, los instrumentos de descripción y las bases de datos. Sus visitas a la institución están, por razones obvias, limitadas por variables espacio-temporales. Además, en este caso es absolutamente necesario que el usuario tenga un conocimiento previo del servicio de archivo del que está haciendo uso, como mínimo debe conocer su existencia y llegar hasta él.

El archivero tiene, por tanto, un contacto directo con el usuario, conoce quién le visita y en muchos casos por qué le visita. También es posible saber con exactitud qué uso hace el usuario del archivo, qué documentación consulta, qué reproducciones solicita, etc. De hecho, en el ámbito de los archivos estatales, desde hace mucho tiempo está regulado que esta información se conserve en un *expediente de investigador*¹ donde habrían de constar “como datos indispensables el nombre y apellidos del titular, el número del documento nacional de identidad, el pasaporte o la tarjeta de identificación personal del país de procedencia y las fechas de expedición (...) así como las razones que justifiquen la investigación o consulta”². La aplicación de la legislación sobre protección de datos, a partir de finales de los años 90 del siglo pasado, obligó a registrar los ficheros informatizados donde se recogía esa información sin afectar a la esencia del sistema. No obstante, a la vista de la nueva legislación y jurisprudencia sobre protección de datos, consideramos necesaria una revisión a fondo del concepto de expediente de investigador para que siga sirviendo, como hasta ahora, para garantizar especialmente la seguridad en la conservación de los documentos y, perseguir, en su caso, posibles robos o deterioros de los documentos originales³.

- b. El usuario-internauta, sin embargo, presenta una serie de características derivadas de la naturaleza anónima y telemática de su visita. De este tipo de usuario no se tiene, a priori, información personal de ningún tipo y puede acceder al archivo desde cualquier lugar y a cualquier hora. Es un usuario anónimo, que puede consultar el archivo como consulta cualquier otra web de uso habitual.

Como este tipo de usuario no tiene contacto con el personal es el propio usuario el que debe “auto-servirse” la documentación que desea consultar, a través de las herramientas y funcionalidades de recuperación de información que se hayan diseñado para el archivo online.

Además, el usuario internauta puede acceder al archivo en Internet sin poseer un conocimiento previo de su existencia. Es posible imaginar el escenario habitual de un usuario consultando en Google (o cualquier otro buscador) y obteniendo resultados que le conducen al archivo en Internet, o clicando en un enlace hallado en una web y le conduce hasta un documento digitalizado visible desde las funcionalidades del archivo online.

TIPO DE USUARIO	CARACTERÍSTICAS	NECESIDADES DE INFORMACIÓN ASOCIADAS AL TIPO DE USUARIO
Presencial	<p>Datos personales conocidos: nombre, ocupación, etc.</p> <p>Contacto directo con el personal de archivo.</p> <p>Acceso limitado por variables espacio-temporales.</p> <p>Puede realizarse un seguimiento individual de su actividad en el archivo.</p> <p>Debe conocer, como mínimo, la existencia del archivo para realizar su consulta.</p>	<p>Tipologías o categorías de usuario.</p> <p>Ocupación/Motivación para la consulta del archivo.</p> <p>Documentación consultada por el usuario.</p> <p>Documentación reproducida por el usuario.</p> <p>Satisfacción derivada del uso del servicio.</p>
Telemático	<p>Acceso muy limitado a sus datos personales, ocupacionales, etc.</p> <p>Acceso telemático al archivo, sin limitación de horarios ni ubicación.</p> <p>Acceso a la información/documentación de forma automática "self service".</p> <p>Puede acceder al archivo online sin conocer previamente su existencia, a la búsqueda de información en Internet.</p> <p>Deben utilizarse herramientas técnicas para el estudio de su comportamiento. Este estudio nunca puede ser individualizado.</p>	<p>Tipologías o segmentos de usuarios (local-foráneo, nuevo-recurrente, fidelizado-ocasional, etc.).</p> <p>Uso de las funcionalidades automatizadas del archivo online (buscadores, visualizadores de documentos digitalizados, etc.).</p> <p>Acceso a la información custodiada en el archivo online.</p> <p>Realización de reproducciones digitales de la documentación.</p> <p>Visibilidad y accesibilidad del servicio en el medio online: posicionamiento en buscadores, fuentes de tráfico.</p>

El archivero no puede prescindir de información específica tanto de sus usuarios como del uso que los mismos hacen de su servicio de archivo. Es necesario acceder a una cierta información si quiere procurar un mejor servicio y adecuarlo a las demandas de estos mismos usuarios.

Se hace patente en este momento una cierta brecha entre las ventajas ofrecidas por el modelo de archivo online y las necesidades de información del gestor de dicho archivo (convertido ahora en *webmaster*): no es posible, sin

herramientas adicionales, obtener datos sobre los usuarios, sus accesos al servicio, sus consultas, uso de la web y funcionalidades, etc.

Va a ser necesario, por tanto, obtener esa información, o gran parte de esa información porque, ¿podemos saber si acceden los usuarios a la web del archivo online? Y habiendo accedido, ¿consultan documentación digitalizada o meramente descrita?, ¿realizan reproducciones de la misma? ¿Está obteniendo, en definitiva, algún resultado el esfuerzo económico y laboral desarrollado por la institución de archivo que se ha propuesto dar el salto a la Red?

Para resolver estas y muchas otras cuestiones será necesario hacer uso de herramientas ya disponibles en el ámbito de las páginas webs (que comparten, al fin y al cabo, muchas de las necesidades de información sobre usos y usuarios que se le plantean al archivero online): las herramientas de analítica web.

Las herramientas de analítica web

Las herramientas de analítica web son aplicaciones diseñadas para obtener datos sobre la utilización que los usuarios hacen de una página web. Son muy variadas y, muchas de ellas, gratuitas. Su uso está muy extendido entre los *webmasters* de todo tipo de sitios webs y, como no podía ser menos, también en los buscadores de archivos⁴. En su formato más habitual (método utilizado, por ejemplo, por Google Analytics) su funcionamiento parte de la inclusión de un pequeño fragmento de código *javascript* que, incrustado en el código de una página web, interactuará con los navegadores de los usuarios que accedan a la misma.

Una vez activado dicho código, cada visitante que realice un acceso al sitio web recibirá en su navegador una *cookie*, que realizará un seguimiento de su actividad en el sitio, incluyendo duración de la visita, URLs a las que accede, clicks, búsquedas internas, etc.

Dichas *cookies* almacenarán, asimismo, cierta cantidad de información del usuario de la página web: proveedor de Internet, web desde la que accede al sitio sobre el que se hace el seguimiento, IP del usuario (con el fin de identificar su lugar de procedencia), navegador que utiliza e idioma del mismo.

Como una *cookie*, a no ser que el usuario especifique lo contrario, permanece en su navegador permanentemente, también será posible identificar a los nuevos usuarios y a aquellos que hayan accedido al sitio web con anterioridad. De la misma manera, será posible conocer el número de visitas recurrentes de los usuarios a la web.

Las herramientas de analítica web, una vez obtenida la información pertinente sobre los usuarios del servicio, ponen a disposición del webmaster o gestor del servicio aplicaciones webs que muestran las diferentes estadísticas de uso de sus sitios webs. Las herramientas más sencillas mostrarán datos puramente estadísticos

como número de visitas (o de visitantes únicos), lugares de procedencia de las mismas o direcciones webs visitadas por los usuarios.

Algunas de las herramientas de analítica web más potentes, sin embargo, ponen a disposición del gestor web diversas mecánicas de configuración de los datos obtenidos. De esta forma será posible categorizar a los usuarios en función de variables personalizadas, que incluyen fuentes de las que procede el tráfico online, frecuencia de acceso de los visitantes recurrentes o uso de las diversas funcionalidades de la web. Serán este tipo de herramientas de analítica web –entre las que destacan entre las gratuitas *Clicky Analytics*, *Open Web Analytics* y *Google Analytics*– las que permitan realizar estudios detallados de los usuarios de un sitio web y de la utilización que hacen del mismo.

Así, mediante una configuración adecuada, las herramientas de analítica web pueden trascender su interpretación habitual de contadores de visitas para convertirse en herramientas para el estudio de usuarios. Será posible identificar tipologías de usuarios en función de diversas variables y estudiar individualmente el uso que cada una de ellas hace de la web y de sus funcionalidades.

Esta utilización avanzada de las herramientas de analítica web es la *segmentación de los datos obtenidos*. Los ejemplos mencionados de herramientas de analítica web permiten al gestor del archivo configurar, filtrar y mostrar la información obtenida mediante el uso de *cookies* de tal forma que es posible aislar información sobre usos o usuarios del servicio web de forma individual.

Utilización de las herramientas de analítica web

El análisis de los usuarios, de sus necesidades y de su interacción con el archivo, tanto presencial como online, tiene como objetivo permitir al archivero procurar un mejor servicio, adecuarlo a las necesidades de sus “clientes” y favorecer el proceso de búsqueda y hallazgo de la información albergada en la institución.

Desde este punto de vista –el estudio de los usuarios como herramienta de *mejora* del servicio de archivo– las herramientas de analítica web, aun con sus limitaciones respecto a la naturaleza de los datos que son capaces de obtener, ofrecen la oportunidad al gestor del archivo online de conocer en profundidad el uso que se está realizando sobre el servicio que provee.

Tradicionalmente, los estudios de usuarios de archivos –archivos físicos o presenciales– han venido categorizando a los usuarios en función de su actividad o su motivación para acceder a la información⁵. De esta forma, se analiza el comportamiento de los usuarios dividiéndolos en grandes bloques: usuarios internos y externos y, dentro de estos, investigadores, “ciudadanos”, estudiantes, etc.

Establecidas estas categorías como ejes rectores del comportamiento de los usuarios, se procede a realizar el análisis de la interacción de estos usuarios con el

archivo y con la documentación consultada. Los datos obtenidos –fundamentalmente cualitativos– deben aplicarse a la mejora del servicio, de la atención al usuario, mecanismos archivísticos y consecución de objetivos definidos del servicio.

De esta forma, un estudio tipo de usuarios establece unas categorías de usuarios en función de sus motivaciones para acceder a la información y en función de la actividad específica que les ha conducido a realizar una consulta.

Es aquí donde aparece la principal brecha entre los estudios de usuarios de archivos presenciales y de archivos online: las herramientas de analítica web, cuyo uso y configuración son necesarios para realizar este tipo de informes, no permiten obtener los datos de los que parten los estudios de usuarios tradicionales. Por su naturaleza online y de uso anónimo, *no es posible conocer la ocupación, actividad profesional, motivación o datos personales de los usuarios*. No puede categorizarse, como se venía haciendo, a los usuarios como investigadores o estudiantes, sencillamente porque no se conoce esta información. Es la aplicación al archivo online del conocido tópico del anonimato en Internet.

Por el contrario, el uso de herramientas de analítica web bien configuradas va a permitir categorizar a los usuarios en función de criterios diferentes pero de gran potencial: se puede agrupar a los usuarios en función del *uso que hayan dado a la web, de sus hábitos de utilización o reutilización del servicio, del tipo de información que desean o del uso y calidad de uso de las funcionalidades del servicio de información*.

De esta forma, el estudio de usuarios de archivo se convierte en estudio de usuarios-internautas. El usuario permanece ahora anónimo y las categorías que se acostumbraban a manejar en la profesión constituyen una incógnita, pero es posible acceder a otro tipo de datos, desglosarlos y compararlos. Varía tanto la naturaleza de los usuarios como la forma de aproximarse a su estudio, pero permanece el objetivo, y en definitiva la finalidad de su realización.

Es decir, las categorías que se manejarán en el archivo online se basarán en datos empíricos de uso del servicio de información: las categorías de usuarios no se definen por bloques preestablecidos, sino por resultados en el uso del propio servicio de información.

Será posible agrupar y seccionar los datos obtenidos para su estudio individual o su estudio comparado. Establecida una categoría de usuarios, se podrá estudiar el uso específico que dicha categoría ha hecho del archivo online. Por poner varios ejemplos tangibles: Será posible conocer el porcentaje de usuarios que llevan a cabo búsquedas en las herramientas de búsqueda interna de la web. Podrá conocerse también si los nuevos usuarios o los usuarios casuales son capaces de acceder a la información deseada, independientemente del formato en que se encuentre.

El estudio de los usuarios del archivo en Internet estará limitado, por tanto, por dos circunstancias:

- El usuario es anónimo y accede de forma telemática.
- Las herramientas de analítica web tienen unas opciones limitadas de obtención de datos de usos y usuarios de un servicio online.

Como ya se ha mencionado, para realizar estudios de usuarios-internautas será necesario configurar convenientemente las herramientas de analítica. Al servicio del gestor del archivo se encuentran las diferentes herramientas de filtrado e identificación de uso de la web y de datos de los usuarios proporcionadas por dichas herramientas.

Por lo tanto, será posible obtener *categorías de usuarios* o *segmentos de la información de uso y usuarios* del servicio en función de determinadas variables. Una vez establecidas estas categorías o segmentos, será posible realizar un análisis extensivo de su utilización del servicio, acceso a la información provista por el mismo, etc.

A partir de la experiencia acumulada con la aplicación de herramientas de analítica web en el sitio web de *DARA, Documentos y Archivos de Aragón*⁶, podemos establecer el siguiente desglose de categorías de usuarios telemáticos:

- Segmentos según datos de los usuarios:
 - Usuario interno/externo.
 - Usuario local/foráneo.
- Segmento-idioma utilizado.
- Segmentos según datos de acceso al servicio:
 - Usuario ocasional/recurrente/habitual.
 - Nuevo usuario/Usuario retornante.
 - Usuario implicado/Usuario casual.
- Segmentos según datos de uso de las herramientas y funcionalidades del servicio:
 - Usuarios con utilización de herramientas de búsqueda internas.
 - Usuarios con acceso a documentación descrita.
 - Usuarios con acceso a documentación digitalizada.
 - Usuarios que realizan reproducciones de documentos.

Mediante la segmentación de la información obtenida por la herramienta de analítica web se supera la tendencia a su utilización como un mero contador de visitas (dato, por otra parte, de vital importancia). Con el establecimiento de estas y otras categorías de usuarios del archivo online no solo será posible realizar un estudio pormenorizado del comportamiento de los usuarios en el archivo online, sino también del impacto sobre dicho comportamiento de eventuales cambios en

la web, mejora del posicionamiento en buscadores de Internet, desarrollo de nuevas herramientas y funcionalidades en el archivo online, etc.

La naturaleza no presencial del usuario de archivo online, propia de cualquier otro servicio web, provoca que su estudio deba atender a una serie de condiciones específicas, no aplicables al usuario presencial.

De esta forma, el estudio de usuarios en Internet deberá atender a una serie de datos específicos y de importancia capital en el funcionamiento y éxito de un archivo online como son las *fuentes de tráfico*, el *posicionamiento en buscadores* y la puesta en funcionamiento de una *red de enlaces funcional*.

El archivero online debe tener en cuenta el supuesto de que el usuario no tiene conocimiento previo de la existencia de su servicio online. Para poder conectar al usuario (y su potencial necesidad de información) con la información provista por el archivo deben utilizarse mecanismos propios del mundo de la Web.

Un estudio de usuarios-internautas debe, por tanto, además de analizar el comportamiento de los usuarios en su web, el uso de las herramientas y funcionalidades puestas a disposición de los mismos y el acceso a la información almacenada en su seno, hacer un seguimiento de las fuentes de tráfico que han propiciado la llegada de los mismos al servicio online. Es conveniente estudiar las diferentes webs que enlazan hasta el archivo online, la presencia en buscadores de Internet e incluso los términos o cadenas de términos utilizados por los usuarios en dichos buscadores. El objetivo de este tipo de análisis será la mejora de la visibilidad y accesibilidad del servicio online.

Aplicación de la normativa de protección de datos a la analítica web

Cuando se habla de archivos y protección de datos se da por hecho que nos estamos refiriendo a los datos contenidos en los documentos que se consultan. La privacidad del usuario de un servicio web también es un bien que debe protegerse y que, al contrario que el anterior, apenas ha despertado interés en la literatura profesional. A continuación vamos a plantear el problema y la forma en que se está acometiendo desde *DARA*.

El principal problema de la aplicación de analítica web relacionado con la protección de datos personales tiene que ver con la naturaleza de los datos de los usuarios. Como se ha indicado anteriormente, para la realización de estudios analíticos del funcionamiento del archivo online es necesario recabar una cierta información sobre los usuarios que acceden al mismo. En su formato más habitual, las herramientas utilizadas interactúan con los navegadores de los usuarios, a priori anónimos, mediante la inclusión en dichos programas de una cookie, un “ítem” individual identificable por una aplicación externa.

Estas *cookies* serán capaces de obtener información de una naturaleza un tanto ambigua: en ningún caso serán capaces de obtener el nombre personal o dirección exacta de los usuarios de un servicio de archivo online, pero sí que se accede a otro tipo de información, digital, que está relacionada con la normativa actual de protección de datos de carácter personal, en concreto sus direcciones IP.

Una dirección IP es un código numérico que identifica un dispositivo conectado a Internet. Este código, que puede ser dinámico (y en cuyo caso el problema normativo estaría resuelto), puede ser –y en muchos casos es– estático, por lo que sería posible, en principio, asociar una dirección IP con una persona o conjunto de personas físicas. A través de esta dirección sería posible identificar de forma personal al usuario que maneja un dispositivo determinado.

Para la Agencia Española de Protección de Datos, además, no hay dudas. En el informe 327/2003 emitido por esta institución, se aborda esta problemática concreta. ¿Deben ser consideradas las direcciones IP como datos de carácter personal, en relación con la Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal? Según dicho informe, las direcciones IP en combinación con terceras informaciones –por ejemplo las que obran en posesión de las empresas proveedoras de servicios de Internet– pueden ser utilizadas para identificar a un usuario de la Red. Si bien se especifica que no siempre puede realizarse esta identificación, la posibilidad real de utilización de la dirección IP de un usuario para obtener su identidad civil provoca que este tipo de datos sean considerados como datos de carácter personal.

Así pues, los datos obtenidos mediante la utilización de herramientas de analítica web son considerados como datos personales, y como tales están sujetos a la normativa vigente. Como norma general, el gestor de un archivo online o webmaster no puede visualizar estos datos considerados de carácter personal. En su lugar, los proveedores de las herramientas de analítica web se encargan de almacenar los datos obtenidos, mostrando en sus aplicaciones webs aquellos de naturaleza estadística.

Sin embargo, se están almacenando de hecho datos de carácter personal. Será responsabilidad del gestor del archivo online seleccionar una herramienta de analítica web que no solo responda a sus necesidades de información, sino que respete también el derecho de privacidad de los usuarios. Entre toda la oferta del mercado disponible, en *DARA Documentos y Archivos de Aragón* se ha escogido la herramienta *Google Analytics*, entre otras razones, porque se ha considerado que cumple con los requisitos exigibles en materia de protección de datos.

Google y la posesión de los datos de carácter personal

Al utilizar el servicio de analítica web de Google, automáticamente el gestor de un sitio web y, en este caso, de un archivo online está permitiendo a Google Inc. que capte y almacene datos que, en España, son considerados de carácter personal.

Sin embargo, se da la circunstancia de que el gestor del archivo en ningún caso, sin la utilización de software de terceros, almacena ni puede acceder a las direcciones IP de los usuarios de su servicio. Estos datos son captados por Google y almacenados por Google, y el gestor del archivo online se convierte en un intermediario entre el usuario y sus datos personales con una empresa que además está ubicada fuera de las fronteras de España.

Según la normativa española, el gestor del archivo online que hace uso de *Google Analytics* se convierte en *exportador de datos de carácter personal*, estableciéndose la empresa americana como *importador de datos de carácter personal*. La salida de los datos al extranjero, por descontado, no exime a los responsables del cumplimiento de la legislación vigente.

Ahora bien, este modelo de captación y exportación de datos de carácter personal está contemplado en la normativa española, concretamente en el artículo 67 del Reglamento de Desarrollo de la Ley Orgánica de Protección de Datos de Carácter Personal (RD 1720/2007): existe una relación de países –y en algunos casos empresas– que, como importadores de datos de carácter personal, son considerados equiparables en su nivel de protección al establecido en la normativa española.

En el caso de las empresas domiciliadas en Estados Unidos se considera equiparables a todas aquellas que se hayan adherido a la condición de Puerto Seguro (“Safe Harbor”), ratificada por el Departamento de Comercio⁷.

La adhesión a la condición de Puerto Seguro es opcional: las empresas interesadas deben dirigirse al Departamento de Comercio de los Estados Unidos, y acreditar el cumplimiento de las medidas de seguridad y salvaguarda de datos de carácter personal establecida. Puede accederse a un listado completo de las empresas que garantizan el cumplimiento de dichas condiciones en <http://export.gov/safeharbor/>. Es el caso de Google Inc. De esta forma, Google es considerada compatible como receptora de datos de carácter personal con la legislación española.

Según se indica en el mencionado Reglamento de Desarrollo, no será necesaria la autorización de la Agencia Española de Protección de Datos para la exportación de los mismos a países extracomunitarios que sean considerados equiparables en su nivel de protección a la normativa española.

Por lo tanto:

- Los datos recabados por las herramientas de analítica web que utilicen para ello un fragmento de código *javascript* son consideradas, al menos en España, como de naturaleza personal y por tanto están sujetos a la normativa vigente.
- En el caso de *Google Analytics* –herramienta de uso mayoritario y de utilización en *DARA, Documentos y Archivos de Aragón*– estos datos de carácter

personal son responsabilidad del gestor del archivo online. Los datos son exportados a un país fuera de la Unión Europea, concretamente a los Estados Unidos de América. El gestor del archivo online sólo puede acceder a una parte de estos datos –excluyendo las direcciones IP de sus usuarios– y no los almacena personalmente.

- Según la normativa vigente, para exportar datos de carácter personal a los Estados Unidos sin la autorización expresa de la AEPD es necesario que la empresa o entidad receptora de los mismos sea considerada Puerto Seguro o Safe Harbor, condición que en el caso de Google Inc. se cumple⁸.

Interacción del gestor del archivo con Google Inc. para salvaguardar la normativa vigente

El uso en cualquier página web de la herramienta de analítica web de Google exige, eso sí, una serie de condiciones que deben cumplirse. Estas condiciones vienen reflejadas en el contrato de uso cuya aceptación es necesaria para la implementación de este tipo de herramientas en un archivo online.

En primer lugar, deberá verificarse que el usuario de Google Analytics no permite la utilización de los datos de carácter personal recabados en el uso de su página web. Permitir el uso de estos datos vulneraría la normativa española a no ser que el webmaster o gestor de archivo online contase con la aprobación expresa de cada uno de los usuarios de su servicio. Al no permitir la utilización por Google de estos datos, se asegura el cumplimiento de la política de privacidad específica de la compañía americana y, por tanto, de la legislación, como vimos, española.

Además, el gestor del archivo deberá incluir de forma explícita, visible y localizable un aviso legal o política de privacidad propia en el que se especifiquen los siguientes aspectos:

- Se está utilizando una herramienta de Google Inc. para el estudio estadístico del uso y usuarios de la página web. Se incluirán los datos pertinentes de dicha empresa.
- Los datos recabados, mediante el uso de cookies, en concreto las direcciones IP, son considerados como de carácter personal. Estos serán almacenados por Google Inc. en condiciones de seguridad y privacidad compatibles con la normativa española.
- Que dichos datos no serán tratados por terceros a no ser que así lo requiera la legislación y dichos terceros traten los datos por cuenta de Google.
- Finalmente, debe especificarse que el usuario siempre puede desactivar el seguimiento de sus datos a consecuencia del uso del servicio web mediante una configuración adecuada de sus navegadores.

Alternativas al almacenamiento de direcciones IP: el caso alemán

Durante el año 2011 Google Analytics estuvo en el punto de mira por la obtención y tratamiento de datos de carácter personal a raíz de la prohibición expresa de su utilización en páginas y servicios webs alemanes.

Según la legislación alemana, cualquier gestor que utilizase los servicios de Google Analytics vulneraba el derecho a la privacidad de sus usuarios. El hecho de que sus direcciones IP fuesen almacenadas y exportadas era considerado ilegal en el país centroeuropeo.

Esta problemática llevó a las autoridades alemanas, incapaces de alcanzar un acuerdo con Google Inc., a prohibir la utilización de su herramienta de analítica web en webs alemanas. Esta situación se mantuvo durante la mayor parte del pasado año.

En septiembre de 2011, sin embargo, Google Inc. desarrolló una opción en la configuración de su herramienta que permitiría su uso en Alemania y, por tanto, en países con legislación de similares restricciones. Mediante la modificación, por parte del webmaster o gestor de servicio online del código *javascript* de seguimiento de usuarios a través de cookies, se incluía la posibilidad de mantener anónimas las direcciones IP de los usuarios, eliminando con ello el carácter personal de los datos recabados.

Esta opción está, por descontado, disponible para los usuarios de Google Analytics en España. Mediante la anonimización de las direcciones IP de sus usuarios el gestor procederá a eliminar el componente personal de los datos almacenados por Google Analytics. Como contrapartida, la calidad de los datos –específicamente de los informes de procedencia geográfica de los visitantes– se verán perjudicados, perdiendo precisión en sus resultados.

La utilización de este código modificado es una opción muy interesante, siempre y cuando se tengan en cuenta las limitaciones impuestas a la calidad de los informes obtenidos. En DARA se ha optado no seguir, de momento, el modelo alemán aunque no se descarta aplicar una solución semejante en caso de que la legislación española o futuras normas del Gobierno de Aragón evolucionen en un sentido más proteccionista de la privacidad aunque ello suponga una limitación de las posibilidades de la analítica web que se aplique a nuestro sitio.

Reflexiones finales

Lo expuesto hasta aquí da pie a algunas reflexiones que consideramos útiles pero que son necesariamente provisionales y revisables, tanto más en el caso que estamos tratando, que sufre continuas modificaciones por el vertiginoso avance de la tecnología:

- El usuario-internauta es una realidad cada vez más presente en el mundo de los archivos. El crecimiento en su número y en la variedad de sus aspiraciones hace necesario un replanteamiento de los objetivos y los sistemas de organización de los archivos. El trabajo del archivero debe orientarse a atender estas demandas al menos tanto como a las de los usuarios presenciales.
- La analítica web proporciona herramientas de análisis y, por tanto, de gestión, más que útiles que necesarias para satisfacer esas demandas.
- Es necesario también buscar puntos de encuentro entre la analítica web y la gestión de usuarios presenciales. Partiendo de metodologías de análisis diferentes hay que buscar que sean realmente complementarios.
- El uso de analítica web ha de ser plenamente respetuosa con la protección de datos de carácter personal de acuerdo con la normativa española y comunitaria. La elección de la herramienta adecuada, además del análisis coste/beneficio, debe tener en cuenta esta variable.
- Debido a lo cambiante del panorama tecnológico, las decisiones en este aspecto deben ser revisadas continuamente para mantener los preceptos anteriores.

Notas

¹ La regulación de la tarjeta arranca de la Orden del Ministerio de Educación Nacional de 4 de marzo de 1959, por la que se dictan normas para la obtención de copias y fotocopias. La Orden fue sustituida por el RD 1969/1999, de 23 de diciembre, por el que se regula la expedición de la tarjeta nacional de investigador para la consulta en los archivos de titularidad estatal y en los adheridos al sistema archivístico español. El recientemente promulgado RD 1708/2011, de 18 de noviembre por el que se establece el Sistema Español de Archivos (...) y su régimen de acceso, aunque dedica su capítulo IV al acceso a los archivos y sus documentos no deja muy claro si han de pervivir los expedientes de investigador y en qué condiciones.

² RD 1969/1999 de 23 de diciembre, art. 3.4 y 4.1.

³ Valga como ejemplo el lamentable caso reciente del Archivo Histórico Provincial de Teruel donde, gracias a la actuación de su director que investigó en el histórico de consultas del Archivo, se pudo llegar a identificar a un ladrón de documentos para ser detenido posteriormente por la Policía.

⁴ Un estudio comparado de las más habituales puede verse en GARCÍA-ARÁEZ MARTÍN-MONTALVO, Javier: *Analítica web al servicio del archivo online: indicadores de uso y propuesta de configuración de herramientas de análisis web en servicios de archivo en Internet*. Tesis de Maestría. Máster en Gestión de Unidades de Información y Documentación. Universidad de Zaragoza, 2011.

⁵ RUBIO HERNÁNDEZ, Alfonso. Estudios de usuarios en archivos municipales. *Scire*, 10 (1), 2004, p. 220-221. Puede verse también el trabajo ya clásico de TARRAUBELLA, X., Tarraubella, X. "Los archivos y sus usuarios", *Actas de la V Conferencia Europea de Archivos*, Lligall, 1997, pp. 39-43. Para el ámbito de Castilla y León, sigue siendo válido en parte *Archivos de Castilla y León: gestión de usuarios*, Valladolid: Consejería de Educación y Cultura, 1996.

⁶ El buscador DARA está accesible en Internet <<http://www.sipca.es/sipca/dara>> desde junio de 2010. Forma parte de un proyecto de colaboración abierto a todos los archivos aragoneses que quieran integrarse. Actualmente incluye un buscador general, dos temáticos y 12 buscadores de archivos. Entre estos últimos se incluyen los archivos históricos provinciales, de la Administración Local, de fundaciones y asociaciones, nobiliarios y religiosos, La analítica web se aplica desde principios de 2011.

⁷ En cumplimiento de la Decisión de la Comisión Europea 2000/520/CE de 26 de julio de 2000.

⁸ Puede consultarse la política de privacidad de Google Inc. en <http://www.google.com/intl/es/policies/privacy/>