Análisis de la privacidad en Internet

Autor: Ayoze Fernández Afonso
Consultor: Cristina Pérez Solà

Fecha: 2 de Enero de 2015

Estudio sobre la privacidad en Internet, en el que se obtiene una visión sobre cuál es la actualidad en este tema. Las redes sociales son el principal tema de controversia cuando se habla sobre privacidad, pero no son el único, existiendo otros temas de bastante interés como el tracking que se está llevando a cabo hoy en día a través de Internet. Con el paso de los años se han ido creando diferentes leyes que afectan a la privacidad en la red, como la Patriot Act que permite a las autoridades americanas la capacidad de monitorizar las comunicaciones con el objetivo de prevenir el terrorismo, violando así la privacidad tanto de ciudadanos americanos como internacionales. Se han descubierto programas de espionaje llevado a cabo principalmente por el Gobierno de EEUU y UK, algunos de los cuáles se mencionan en este documento. El crimen se ha trasladado al mundo cibernético, teniendo como uno de los principales campos de aplicación el negocio del tráfico de información. En este proyecto se comentarán cuáles son los vectores de ataque en los que el cibercrimen y la privacidad están relacionados.

This project is about Internet privacy, analyzing the current situation. When people speak about privacy, the topic of most concern is the use of social networks, but it's not the only one, there are others like online tracking, which is covered in this paper as well. Over the years, many laws have been created that affect online privacy, like the Patriot Act which allows US authorities to monitor both american and international citizens under the premise of preventing terrorism. These past years some spying programs have been uncovered, showing that the US and UK have been monitoring the rest of the world for some time. The crime has move to the cyber world, where one of the main business is the exchange of information. This project will introduce some of the attack vectors which are directly related to the privacy of Internet users.

Índice

1 - Introducción

Desde hace relativamente poco tiempo, la sociedad ha empezado a darse cuenta de los problemas que implican las nuevas tecnologías y el estar conectado. Comienzan a salir a la luz casos de vigilancia a nivel mundial, las noticias sobre robo de datos cada vez son más frecuentes, etc.

Se ha visto como Internet es un arma de doble filo. Es una gran fuente de información que se puede usar tanto para bien como para mal, y es en esto último donde radica el serio problema de privacidad que experimentan los usuarios. Muchas de las aplicaciones y servicios que se utilizan hoy en día requieren una gran cantidad de datos personales para "mejorar" la experiencia de usuario. Otras veces, es el propio usuario el que "regala" información personal, debido a una falta de conocimiento sobre cómo de importante son esos datos. Sea cuál sea el caso, con Internet hay un problema de privacidad evidente.

El objetivo de este proyecto consiste en realizar un repaso a los últimos años, analizando la situación actual de la privacidad y cómo se ha llegado hasta ella. Para esto, el proyecto se estructurará de la siguiente manera:

1. Estudio general sobre el interés en la privacidad, donde se observa entre otras cosas ¿qué regiones están más interesadas en la privacidad? o ¿qué temas despiertan mayor interés cuando se habla de privacidad?
2. Listado de tecnologías que han aparecido en los últimos años y que afectan a la privacidad.
3. Leyes y programas esponsorizados por el Gobierno (de EEUU mayormente) en los que la privacidad de los ciudadanos está implicada de algún modo.
4. Análisis de la relación existente entre los vectores de cibercrimen actuales y la privacidad.

El método que se utilizará para la elaboración del proyecto será:

1. Búsqueda de información en fuentes públicas sobre la privacidad.
2. Análisis y síntesis de la información, identificando cuáles son los aspectos de interés para el objetivo del proyecto.
3. Redacción de la memoria utilizando los datos obtenidos.

2 - Herramientas utilizadas durante el estudio

Para la realización del estudio se utilizarán fundamentalmente dos herramientas de análisis de información a gran escala, lo que se conoce hoy en día como *Big Data, c*ada una ofreciendo un punto de vista diferente sobre un mismo tema, además del buscador de Google para ampliar información sobre un tema en concreto.

2.1 - MediaCloud

MediaCloud[1] es un proyecto de código abierto llevado a cabo conjuntamente entre el Berkman Center for Internet & Society de la Universidad de Harvard[2] y el Center for Civic Media del MIT[3].

Esta plataforma ofrece a investigadores la posiblidad de realizar un análisis cuantitativo sobre un tema en concreto en Internet. Cuando se busca una palabra se devuelve una gráfica que muestra la frecuencia de esa palabra en el rango fijado, esto quiere decir, la cantidad de veces que aparece la palabra en el texto. Además, se devuelve otro listado con las palabras que más aparecen en el mismo texto, de modo que el investigador pueda hacerse una idea sobre que temas están relacionados entre un término y otro.

Para poder realizar dicho análisis, la herramienta se encarga de recolectar información de una gran cantidad de sitios de todo tipo, tales como blogs, periódicos, media, etc.

2.2 - Google Trends

Google Trends[4] es un sistema de analíticas de Google que ofrece una visión relativa del interés de un tema en Internet.

Los datos que utiliza son las búsquedas realizadas a través del propio buscador de Google, por lo que la información que se puede extraer de este servicio está directamente ligada con las búsquedas realizadas en Google, el cuál es el buscador más usado hoy en día[5].

Esta herramienta ofrece una perspectiva relativa desde el punto de vista de que los datos devueltos están basados en el total de búsquedas realizadas en el buscador. Esto quiere decir que, que un punto se encuentre más elevado que otro en la gráfica no significa que el término haya perdido popularidad, sino que del total de búsquedas realizadas, la búsqueda del término ha tenido menos impacto. Esto puede ser debido a multitud de factores, como por ejemplo la búsqueda de otro término más popular, el aumento del número de búsquedas a nivel general, etc. También ofrece la posibilidad de realizar un análisis en función de la popularidad de un término en concreto, en vez de a nivel general, de modo que se puede analizar cómo de popular es una cosa en función de la otra.

Finalmente, que en la gráfica aparezca un 100 no significa que sólo se ha buscado ese término en ese periodo de tiempo, sino que al menos el 10% de las búsquedas han sido por ese término.

3 - Interés general en la privacidad

La primera sección del estudio consistirá en un análisis a nivel general del concepto de privacidad a lo largo de estos últimos años. Hay que tener en cuenta que los datos obtenidos con MediaCloud datan a partir de mediados del 2010, ya que es cuando la plataforma empezó su funcionamiento.

Este estudio estará centrado en los lenguajes español e inglés, ya que son los lenguajes que domina su autor y ocupan la segunda y tercera posición respectivamente del ranking mundial[6], estando por detrás del Chino Mandarín.

Antes de comenzar el estudio, una breve introducción al concepto de privacidad. La privacidad, según la Real Academia Española, es el *"Ámbito de la vida privada que se tiene derecho a proteger de cualquier intromisión"*. Traducido a las nuevas tecnologías, esto significa que un usuario tiene derecho a decidir qué datos quiere compartir y quién será el receptor de esta información.

Teniendo en cuenta lo dicho anteriormente, en la siguiente sección se analiza la popularidad de las palabras *privacy* y *privacidad*. ¿Por qué no se utilizan las palabras *private* y *privado*? La explicación es que el concepto de privacidad **siempre** hace referencia al tema de objeto de este estudio, es decir, a los datos privados de un usuario. Sin embargo, hablar de privado hace referencia a otros ámbitos, como por ejemplo una universidad privada o el sector privado. A continuación una comparación que representa esta diferencia.

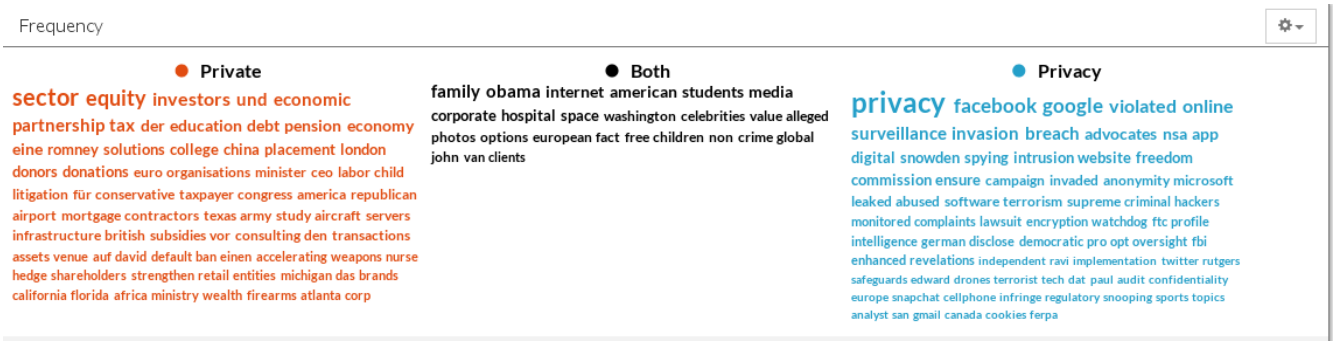

Figura 1 – Comparativa entre "private" y "privacy"

Se observa una clara diferencia en el ámbito de aplicación de cada término, siendo claramente *privacy* el que está relacionado con Internet. Un dato a destacar es que en los textos que contienen la palabra *private*, ni si quiera es ésta la palabra con mayor frecuencia, mientras que en los textos con *privacy* sí que lo es.

3.1 - Privacy

Primero se realiza el análisis a nivel global buscando por el término inglés, *privacy*. MediaCloud devuelve la siguiente gráfica al realizar la búsqueda desde Junio del 2010 al 1 de Noviembre de 2014.

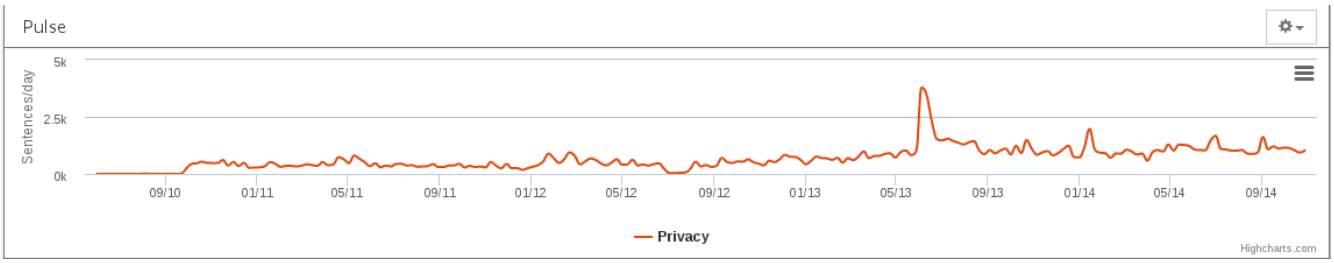

Figura 2 – Frecuencia del término "privacy"

Se observa que más o menos ha habido cierta estabilidad. Destacar la caída que hubo a mediados del 2012, cuando prácticamente no se publicó nada que contuviera la palabra privacy. A partir de ese punto se comienza a experimentar un ligero incremento en contenido con esa palabra, hasta que en Junio del 2013 se dispara la gráfica y a partir de este instante la media de contenido referente a privacy aumenta el doble. Este pico hace referencia al famoso caso de Edward Snowden, del que se entrará en detalles más adelante.

La caída de Julio del 2012 se ha analizado pero no se han encontrado evidencias de por qué hubo tal caída, con un mínimo de 32 historias devueltas el día 7 de Julio. De hecho, buscando por noticias de ese mes se puede contemplar como efectivamente habían publicaciones con *privacy* en su texto, entre otros en sitios de los que se encuentran en las fuentes de MediaCloud como el NY Times, el Wall Street Journal o la BBC.

A continuación se muestra una imagen con los términos que más aparecen junto a *privacy:*

Figura 3 – Términos que aparecen junto a "privacy"

El principal interés está centrado en la privacidad de Google y de Facebook, donde preocupan las invasiones y violaciones de privacidad. Además, la NSA y sus programas de vigilancia son otros de los grandes destacados. Analizando esta nube de palabras también se observa que las nuevas tecnologías como smartphones y drones están en el punto de mira de la privacidad.

Pasando ahora a Google Trends, se analiza como han variado las búsquedas del término *privacy* en la última década. A nivel mundial, esta es la gráfica de las estadísticas de la búsqueda:

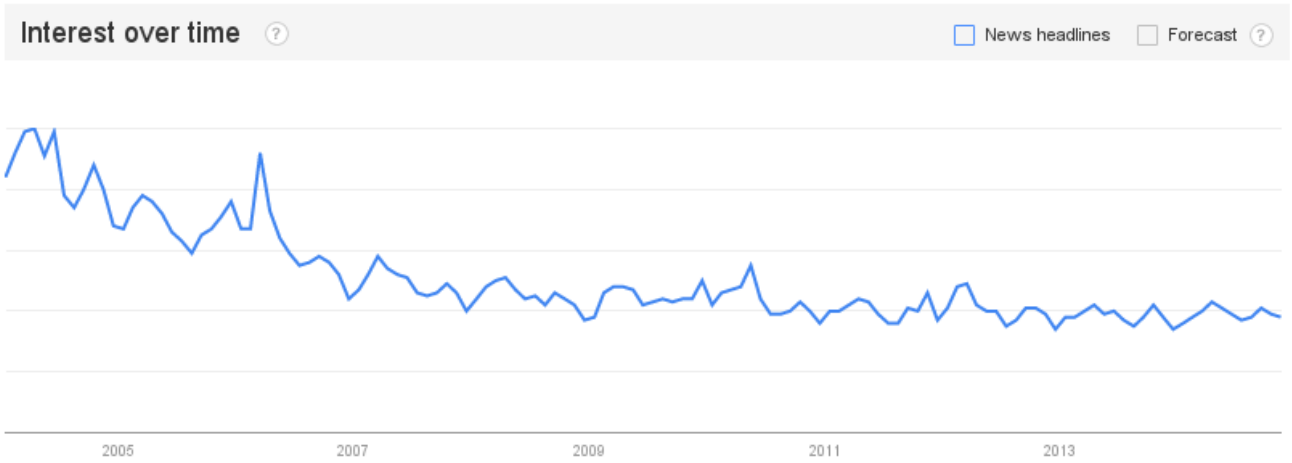

Figura 4 – Búsquedas realizadas en Google por “privacy”

Según esta gráfica parece ser que la sociedad no está interesada en la privacidad, al contrario de lo que parecía indicar MediaCloud, por lo tanto, ¿qué puede haber pasado? El truco está en **qué se le pregunta a Google**. Cuando alguien está interesado en buscar temas de privacidad, generalmente no escribe simplemente *privacy* como sería nuestro caso, sino que añade algún término más a la búsqueda. Así por ejemplo, si en vez de privacy se busca por *privacy policy* se obtiene:

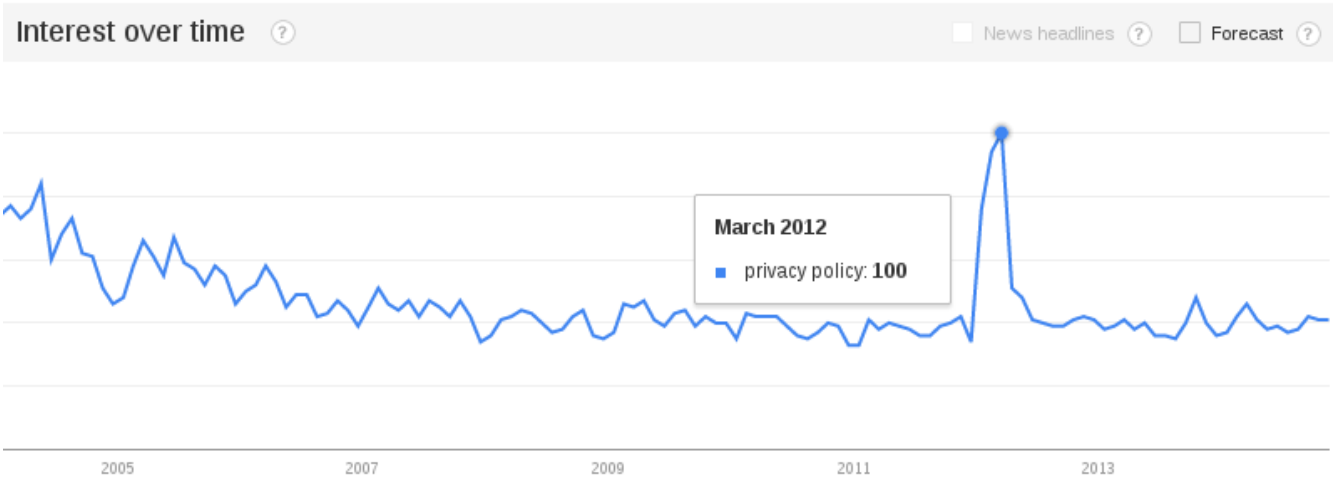

Figura 5 – Búsquedas realizadas en Google por “privacy policy”

En esta gráfica se muestra que el interés por politícas de privacidad ha sido más o menos constante, salvo en Marzo 2012 cuando Google implementó una nueva política de privacidad[7].

Volviendo al caso inicial del término *privacy*, se contempla un dato bastante curioso: el país con mayor interés parece ser Italia. Es un dato curioso porque buscando por una palabra en inglés, lo más común es que uno se espere ver a un país angloparlante encabezando la lista, y sin embargo, aparece Italia. El segundo dato importante que se saca de esta gráfica es que, ante todo pronóstico, no es Estados Unidos el país que lidera la clasificación de los paises angloparlantes, sino Australia.

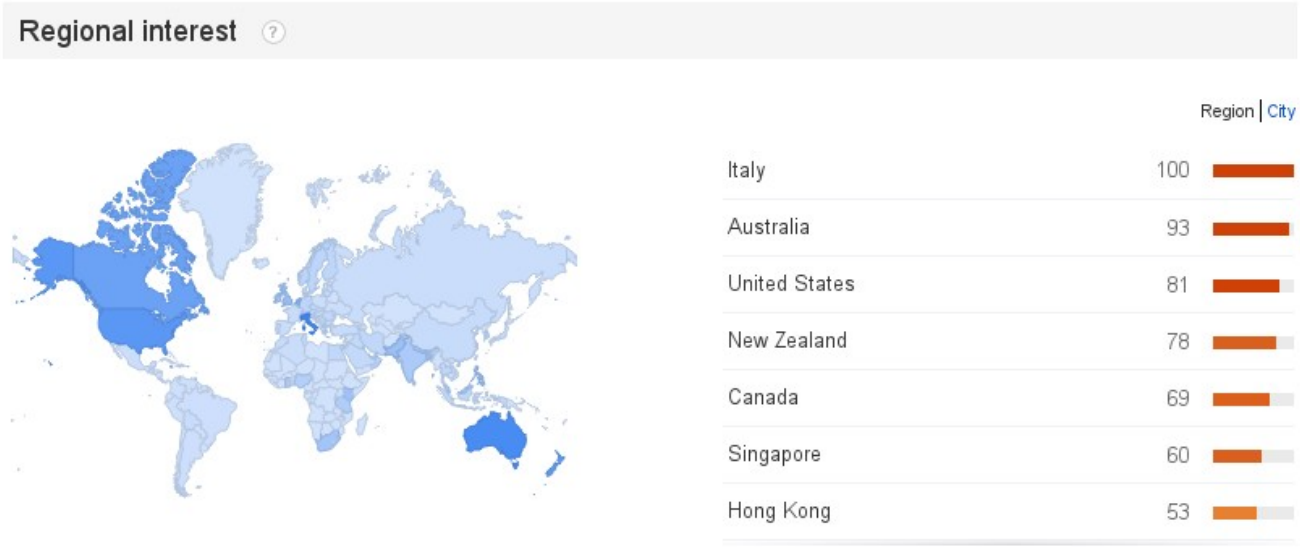

Figura 6 – Regiones que más han buscado “privacy policy”

Un dato a tener en cuenta de estas puntuaciones es que no son datos absolutos, es decir, esto no significa que la cantidad de italianos interesados por la privacidad es mayor que la cantidad de estadounidenses, sino que de dentro de la población italiana hay más interés que dentro de la población estadounidense. Obviamente si se traduce esto a valores absolutos, la cantidad de estadounidense sería mucho mayor, ya que según los datos públicos de Google[8], la población total de Italia es de 59.83 millones de habitantes mientras que la de EEUU es de 316.1 millones de habitantes.

Finalmente se confirma como gran parte del interés recae específicamente tanto en la privacidad de Google como la de Facebook.

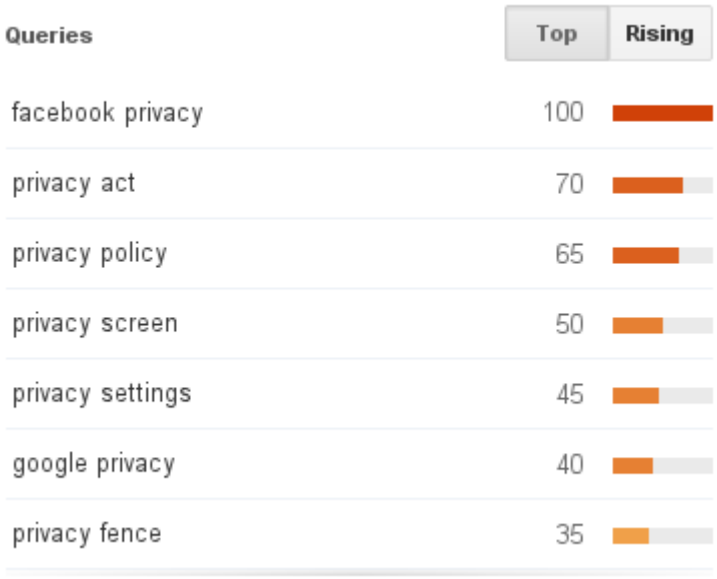

Figura 7 – Búsquedas realizadas con “privacy”

3.2 - Privacidad

Veamos ahora como se ve este mismo análisis desde el punto de vista de los países hispanoparlantes. La siguiente imagen muestra la gráfica de contenido que tenga la palabra *privacidad* utilizando MediaCloud.

Figura 8 – Frecuencia del término “privacidad”

A simple vista se pueden determinar dos hechos bastante importantes. Primero, es que el contenido con la palabra *privacidad* es prácticamente nulo hasta Abril de 2013, y segundo, es que una vez que sucede el caso de Snowden y comienza a haber más contenido con la palabra *privacidad*, éste es ínfimo comparado con el término *privacy*.

Utilizando Google Trends se obtiene la siguiente gráfica:

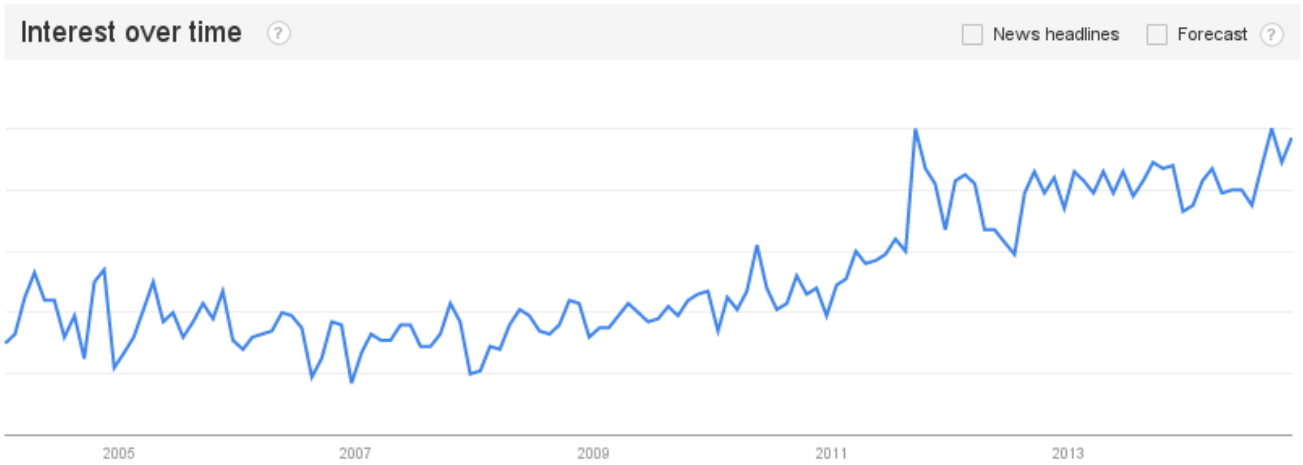

Figura 9 – Búsquedas realizadas por “privacidad”

El interés por este tema va en aumento, cada vez siendo más los hispanohablantes que buscan información en Google por *privacidad*. En el caso de los países se observa un claro predominio por parte de los paises americanos, liderando México la búsqueda.

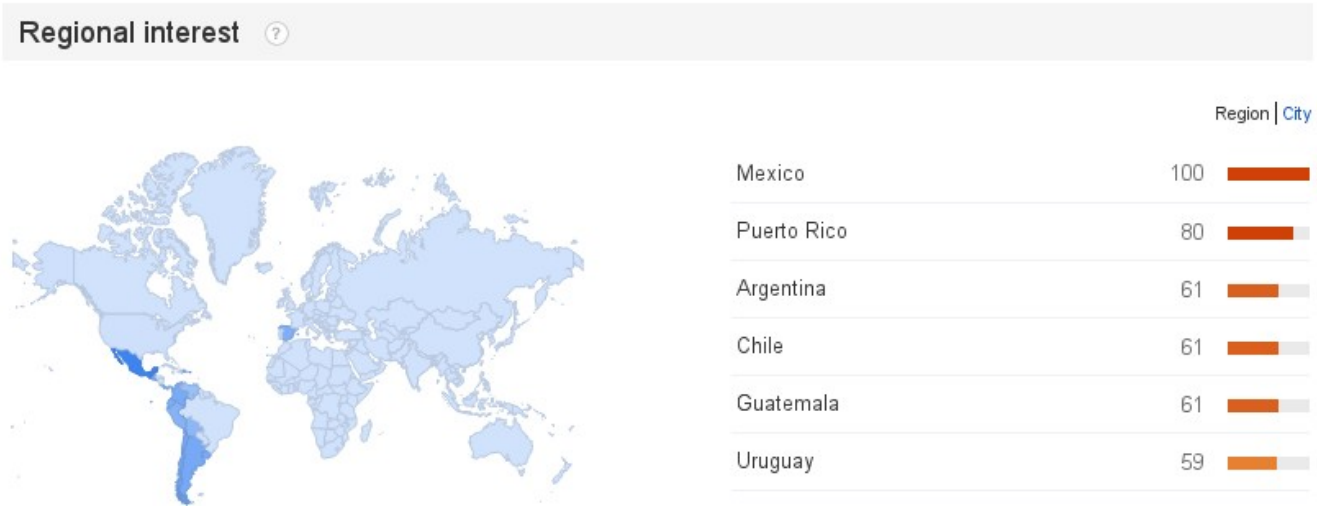

Figura 10 – Regiones que más han buscado por “privacidad”

En la lista de las consultas más realizadas se observa como Google ni si quiera llega a aparecer, lo cuál es bastante desconcertarte. Posiblemente esto es debido a que en español se utilizan distintos artículos y preposiciones, los cuáles añaden diferentes formas de realizar una misma consulta. En este caso vemos como de siete búsquedas distintas, en verdad sólo se están tratando cuatro temas: privacidad, privacidad en facebook, aviso de privacidad y privacidad de datos. En la versión en inglés los siete resultados implicaban siete temas distintos.

Figura 11 – Búsquedas realizadas con “privacidad”

3.3 - Privacy vs Privacidad en España

Para finalizar esta sección se estudia la diferencia entre *privacy* y *privacidad* en España.

Analizando la situación se puede comprobar como aquí si que hay cierto interés por la privacidad de Google, aunque no tiene ni comparación con la preocupación que existe por la privacidad de Facebook.

Figura 12 – "Privacidad" en España

De este análisis se puede inferir también que las regiones que no están interesadas en la privacidad de Google son la mayoría de los países americanos, ya que si en la búsqueda global no aparecía *privacidad google* y sin embargo en la búsqueda de España sí, significa que en general los países americanos tienen otras prioridades en lo que a privacidad se refiere, lo cuál no significa que haya algún país concreto en el que si que aparezca *privacidad google* entre los términos más buscados.

Si se analizan las búsquedas por *privacy* en España se obtiene una visión completamente distinta a la vista anteriormente, la cuál es bastante interesante.

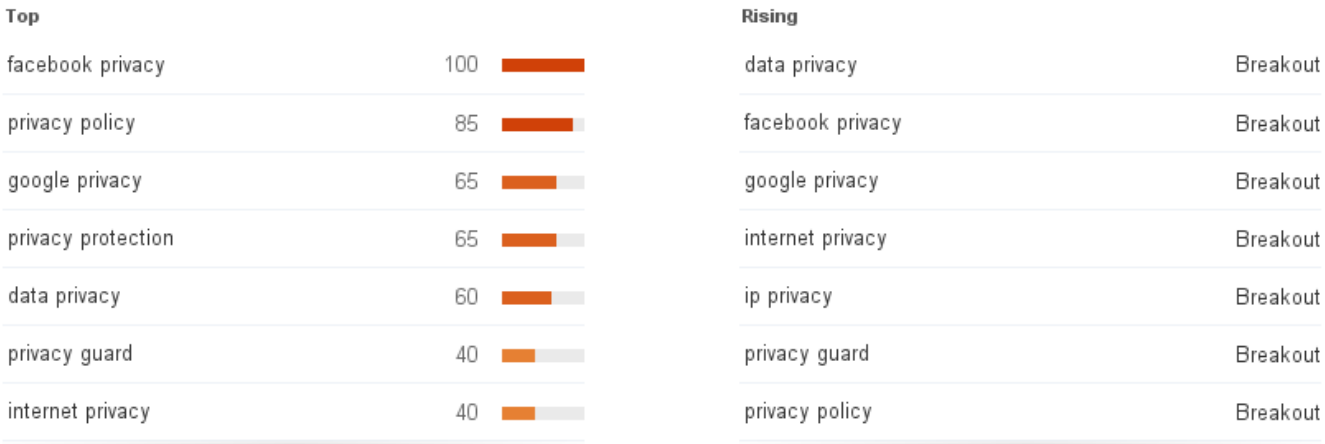

Figura 13 – "Privacy" en España

Facebook sigue en cabeza pero ya no por tanta ventaja con Google. Pero esto no es lo más importante de esta búsqueda, si observamos la imagen vemos como unas de las principales búsquedas son *privacy protection, privacy guard e internet privacy,* ninguna de las cuáles fueron realizadas con los términos en español. Además se observa como una de las búsquedas que está en aumento es *ip privacy*, lo que significa que posiblemente en España empiezan a haber más usuarios interesados en técnicas de anónimato en Internet, como por ejemplo VPNs[9] o TOR[10].

Realizar un análisis de contenido es irrelevante en este caso, ya que como es de esperar el contenido en las fuentes de información españolas estará utilizando *privacidad* en vez de la versión inglesa.

4 - Tecnologías que afectan a la privacidad

En esta sección se estudiarán diferentes tecnologías que han surgido estos últimos años y que afectan a la privacidad de los usuarios, algunas para bien y otras para mal.

4.1 - Redes sociales

Como se ha visto en el análisis anterior, la principal causa de preocupación son las redes sociales, lo cuál parece bastante obvio ya que con la introducción las redes sociales los usuarios de Internet han visto un medio en el que publicar todo tipo de datos personales, sin pensar previamente qué implicaciones tienen esos datos en su vida privada.

Las redes sociales, aunque parecen un concepto nuevo, vienen desde el año 2002, año en que se creó la primera red social (con éxito) conocida como Friendster[11]. En la gráfica siguiente se representa la popularidad basada en búsquedas de esta red, teniendo un pico en el 2009 debido a que en esa fecha le fue otorgada la patente de *Compatibilty Scoring*[12], una característica similar al *Personas que quizás conozcas* que vemos hoy en día en otras redes como Facebook o LinkedIn.

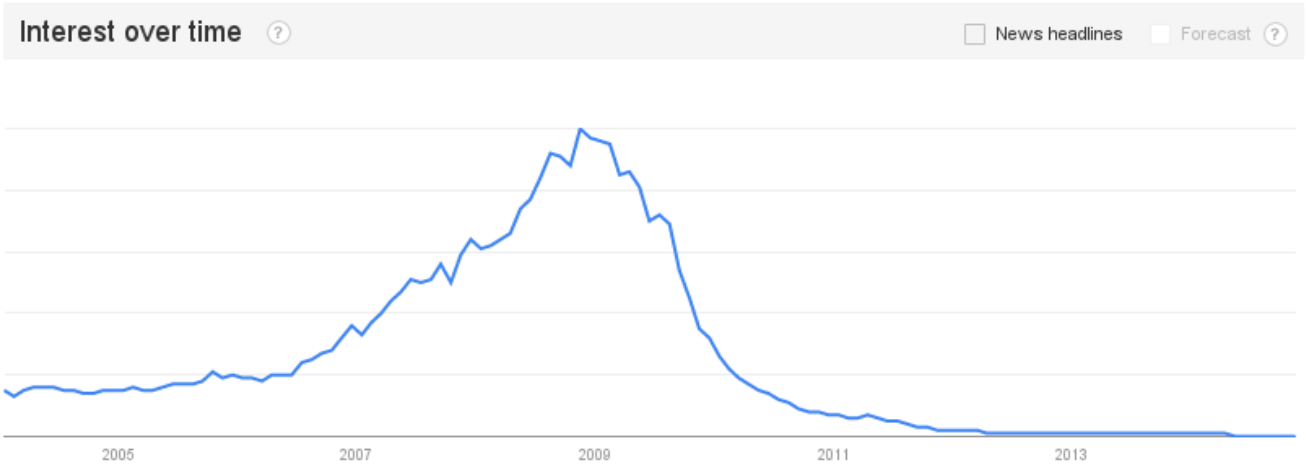

Figura 14 – Búsquedas realizadas por “Friendster”

La pérdida de interés en esta red está directamente ligada al auge de otras redes, las que conocemos y utilizamos hoy en día. Más adelante se mostrará una comparativa sobre este hecho.

Un dato interesante de esta primera red social es que prácticamente toda la base de usuarios era asiática, debido generalmente a que está basada en Malasya, por lo que se puede afirmar con certerza que fueron los asiáticos los pioneros en el uso de la redes sociales.

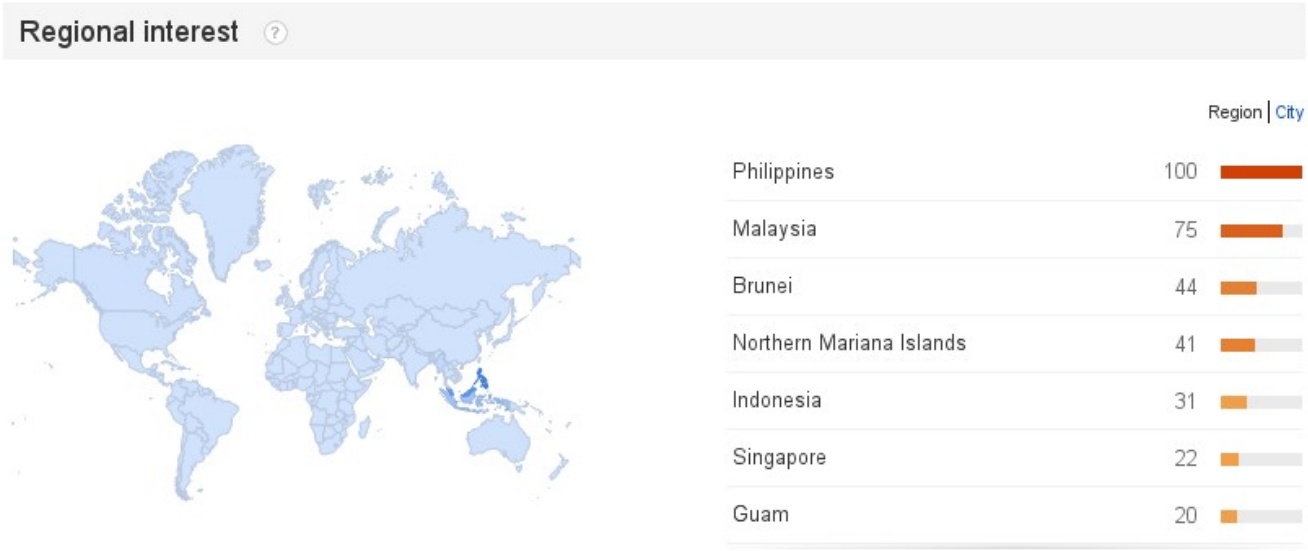

Figura 15 – Países que más han buscado Friendster en Google

Analizando *Friendster* con MediaCloud, no se obtienen grandes resultados como era de esperar, ya que MediaCloud comenzó a funcionar a mediados del 2010 y el interés en Friendster decayó en 2009. Sin embargo se identifica un pico importante el 26 de Abril de 2011, día en que la red anunció que se iba a borrar todo el contenido relacionado con la red social[13], para convertirse en lo que es hoy en día, una plataforma para juegos.

Figura 16 – Frecuencia del término “Friendster”

Al siguiente año comienza su funcionamiento la red profesional LinkedIn, y también se crea la que fue la red más usada en la era pre-Facebook: MySpace. Esta red social basada en contenido musical estuvo en auge hasta que en 2008 Facebook la sobrepasó[14]. Hoy en día sigue activa con un nuevo diseño, pero la cantidad de usuarios activos que tiene es irrelevante comparada con el resto de redes.

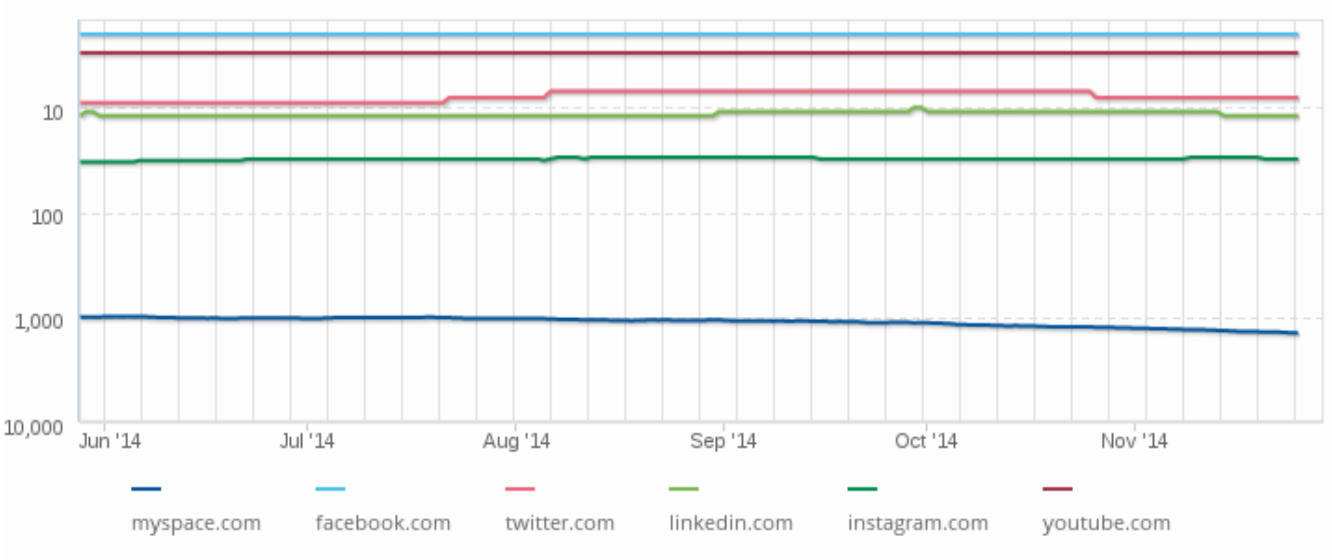

Figura 17 – Comparación del tráfico de varias redes en Alexa

El interes en MySpace se observa como decae claramente desde que Facebook se hace famoso.

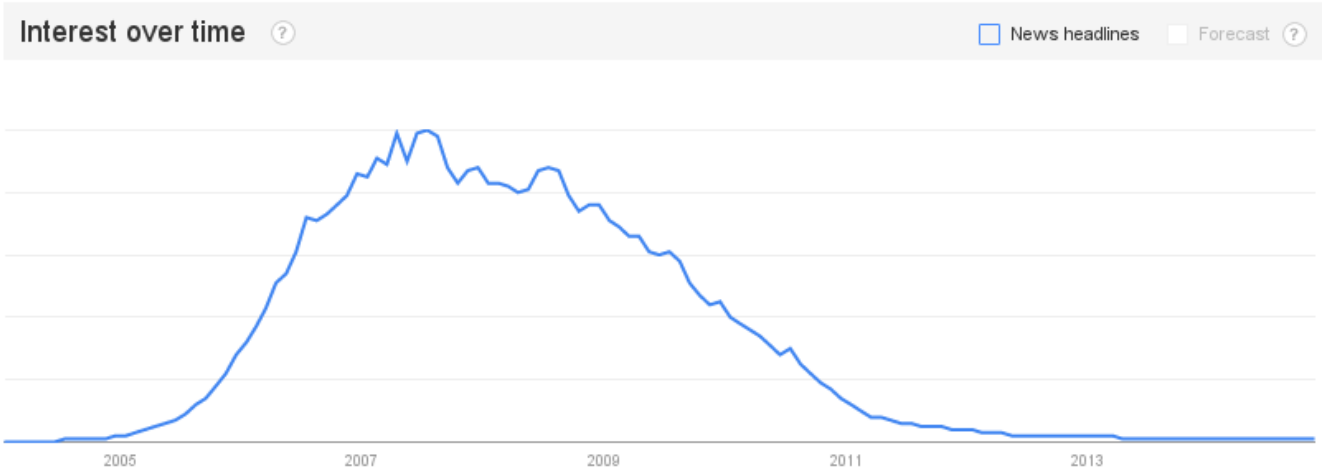

Figura 18 – MySpace en Google Trends

Una curiosidad de MySpace es las regiones que más han realizado búsquedas en Google por tal servicio, ya que no son nada comúnes. Aunque MySpace tiene una buena base en Estados Unidos, son los territorios estadounidenses del Pacífico los que más han buscado información sobre esta red.

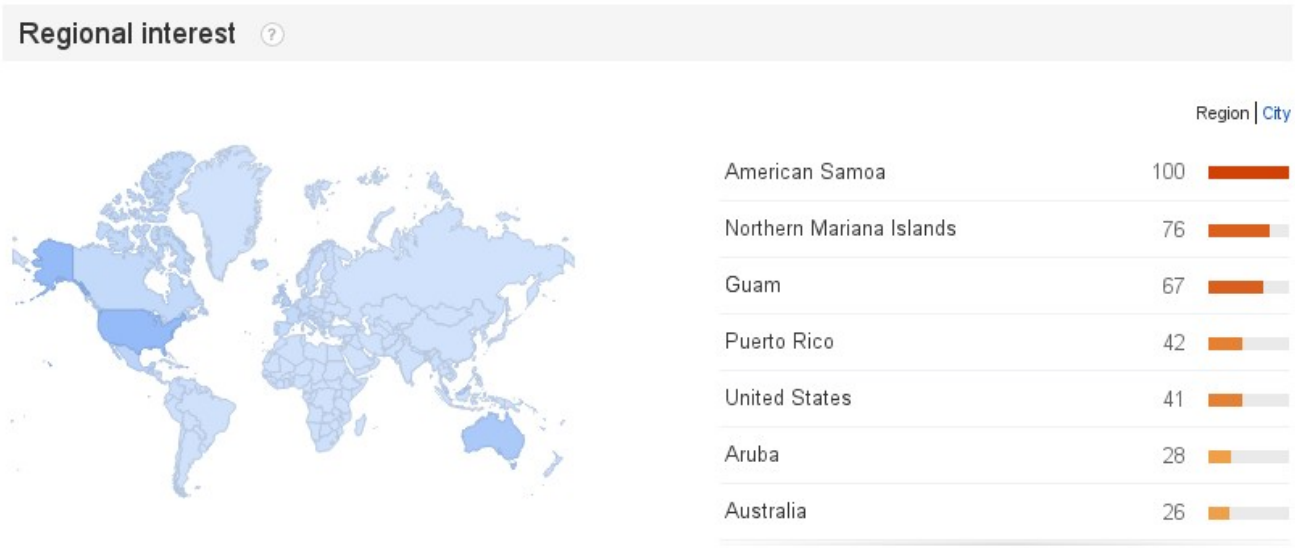

Figura 19 – Regiones que más han buscado “MySpace”

2004 es el año en el que Facebook fue creada. Al principio sólo era accesible para los estudiantes de Harvard, universidad a la que pertenecía Mark Zuckerberg, su creador. Al año siguiente fue abierta a otras universidades, pero no fue hasta Septiembre del 2006 cuando estuvo disponible para todo el público mayor de 13 años.

A partir de este momento empezaron a salir todo tipo de redes sociales, basadas en mensajes, en fotos, en videos, en deportes...

A continuación se muestran una gráfica comparativa de algunas de las más famosas: Facebook, Twitter, Instagram, YouTube y LinkedIn.

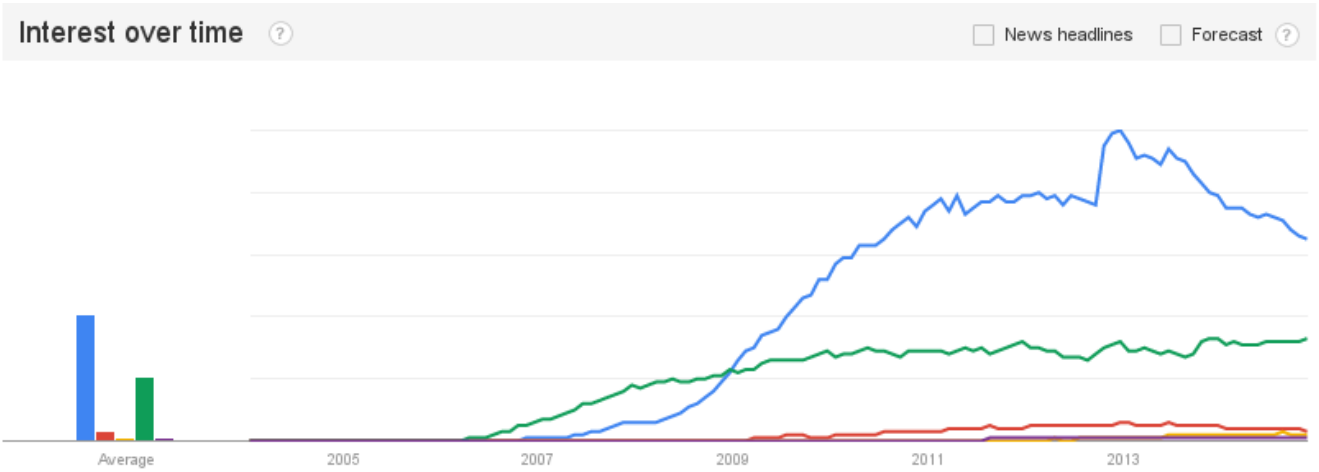

Figura 20 – Comparación de la popularidad de varias redes a nivel de búsquedas en Google

Aunque YouTube tiene una ligera ventaja al principio, desde que Facebook salta a la fama su crecimiento ha sido imparable. Hay que recordar que esta gráfica no muestra la actividad de cada red sino la relación de los términos más buscados, representándolos en función de la que haya sido más

buscada. En este caso, la gráfica muestra cuánto se buscan las otras redes en función de lo que se ha buscado *Facebook*. Para consultar la popularidad en función del tráfico se puede consultar la Figura 17 mostrada anteriormente.

Referente al tema que nos ocupa, la privacidad, ¿cómo es esta relación?

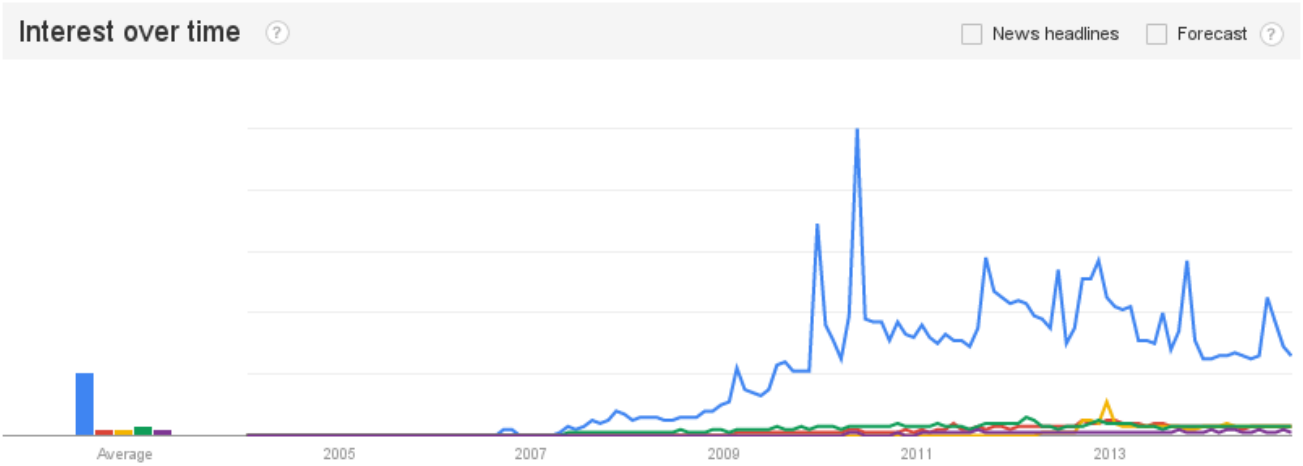

Figura 21 – Comparación de las búsquedas por la privacidad de cada una de las redes

Como era de esperar Facebook encabeza la gráfica. Hay un pico en Diciembre del 2012 que hace referencia a un cambio de política de privacidad en Instagram que causó bastante revuelo[15]. Es destacable que la privacidad de LinkedIn sea la que menos peso tenga, estando por debajo de redes como Twitter o YouTube en las que en teoría no deberían haber datos personales, al contrario que en LinkedIn donde hay información a nivel profesional de los usuarios (educación, experiencia, etc).

Por su propia naturaleza las redes sociales presentan un problema de privacidad para los usuarios, que en principio depende de ellos mismos, pero que como se ha visto también viene causado muchas veces por malas implementaciones de la seguridad del servicio.

La privacidad del usuario también viene reflejada por la red social en sí. No es lo mismo un usuario de YouTube que se dedica a ver/publicar/comentar videos, que un usuario de Endomondo por ejemplo, que es una red social para deportistas que está "subida al carro" del e-health[16], donde por defecto todos los datos de los usuarios son públicos, **datos referentes a la salud del usuario**[17].

Los datos que se comparten en una red social pueden dar lugar a los siguientes problemas:

- Robo de identidad: no es de esperar que con la cantidad de datos personales que se ofrecen, éste no sea uno de los principales problemas.

- Depredadores sexuales: una vez más, la cantidad de datos personales y la facilidad de crearse cuentas falsas (dándoles realismo con datos personales de otras personas) se convierte en un serio problema en las redes sociales. Existe por ejemplo el caso de Peter Chapman[18], quién a través de un perfil falso logró quedar con una chica de 17 años a la que violó y asesinó.

- Acoso: otro de los problemas es el acoso, o stalking en inglés. Ya sea por medio de perfiles falsos o porque la información es pública directamente, uno se puede encargar de seguir los pasos de una víctima. ¿Quién es? ¿Con quién está? ¿Cuáles son sus hobbies? ¿Cuál es su trabajo? Etc.

- Fama: por medio de las redes sociales se puede obtener fama, ahora la cuestión está en:¿Qué fama? En el mayor de los casos la respuesta conlleva aspectos negativos del usuario. Veáse por ejemplo la multitud de memes y videos virales que se aparecen a través de una red social[19].

- Trabajo: íntimamente ligado con el punto anterior pero desde otro punto de vista. A la hora de buscar un trabajo, la actividad social de un usuario influye. Si en un proceso de selección el responsable de recursos humanos se dedica a echar un vistazo a los perfiles de las redes sociales de un candidato, puede descubrir la suficiente información como para decidir si concederle una entrevista o no.

- Monitorización: la monitorización es otro de los grandes problemas ligados a las redes sociales y los datos que en éstas se comparten. Se podría decir que estamos ante un caso de stalking, pero a nivel global. Las redes sociales y las aplicaciones que se utilizan dentro de ellas, muchas veces ofrecen nuestros datos a terceras partes, información que viene detalla en los términos del servicio, texto que la mayoría de los usuarios nunca lee antes de aceptarlos[20].

4.2 - Tor

El proyecto Tor ha sido una de las mejoras, con base tecnológica, a la privacidad en esta última década. Aunque su creación data del 2002, no fue hasta el 2006 cuando el proyecto empezó a ganar más adeptos, gracias a la creación de la fundación de The Tor Project[21].

Tor funciona enrutando el tráfico de un usuario por múltiples nodos, o relays, cada uno cifrando el tráfico, hasta que llega a un nodo de salida que es el que redirige el tráfico cifrado al destino inicial del usuario.

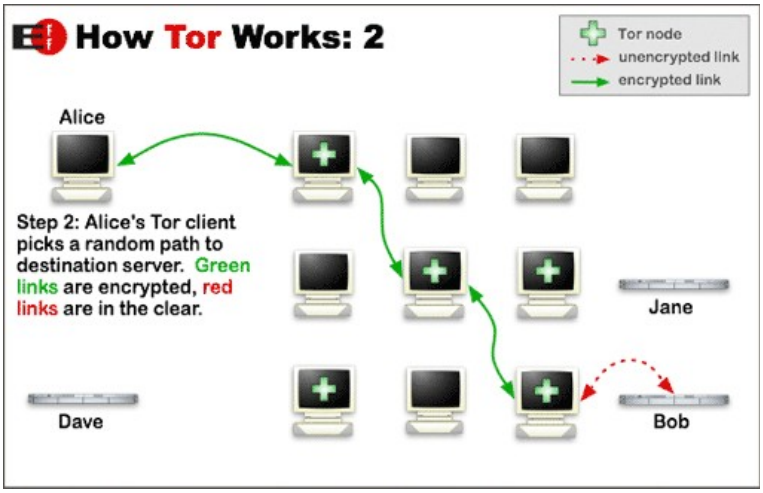

Figura 22 – Funcionamiento de Tor

Realizando este enrutado de tráfico por varios nodos, los cuáles solo conocen el camino hacia el nodo anterior y no más allá, se consigue evitar que se monitorice la actividad en Internet de un usuario, ya que a priori es imposible relacionar el origen inicial con el destino final.

Aunque la idea principal de Tor era la de ofrecer un medio por el que navegar anónimamente conservando la privacidad, obviamente esta característica atrajo a todo tipo de delincuentes que utilizan la red para ejecutar sus fechorías. Por este motivo siempre ha habido un cierto interés en romper el funcionamiento de esta red, y concretamente en Julio del 2014 se demostró como es posible revelar la identidad de un cliente[22].

En la siguiente imagen se muestra el interés en Tor en estos últimos años. Se observa como la gráfica se va incrementando con el paso de los años, lo que significa que cada vez existe un mayor interés en la navegación segura.

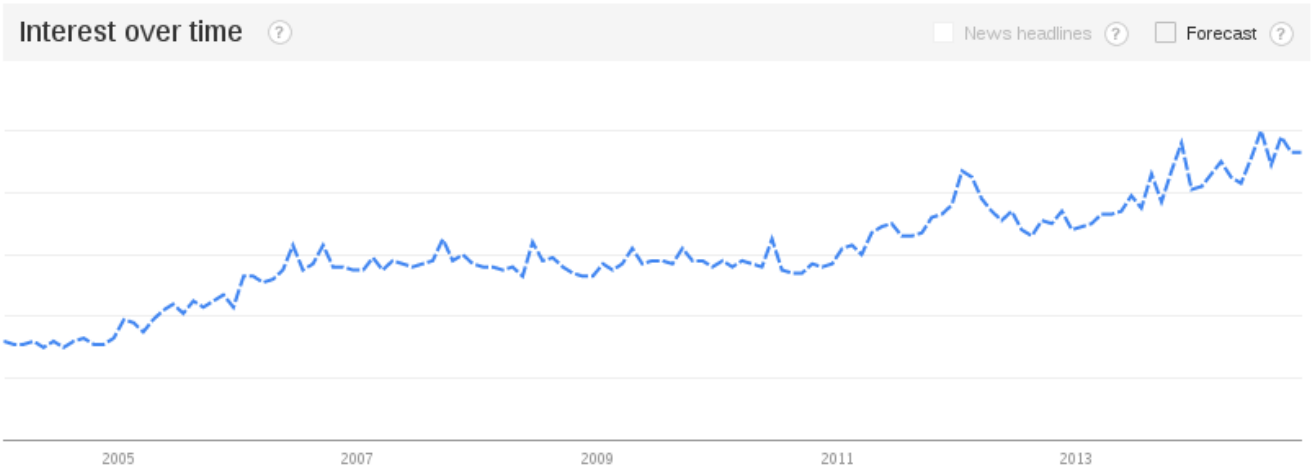

Figura 23 – Interés en Tor

Como era de esperar, entre las regiones que más interés muestran en la herramienta se encuentran varias conflictivas de Oriente Medio, como Siria, Irán o Afghanistán. Como opinión personal me parece sorprendente que sea Noruega el país que lidere la lista, ya que no me lo esperaba para nada. He buscado información sobre si existe algún tipo de censura en Internet en este país, pero no parece ser el caso[23]. Sin embargo, buscando más información compruebo que Tor es un nombre propio utilizado sobre todo en los países nórdicos, lo cuál puede explicar este resultado.

Figura 24 – Regiones más interesadas en Tor

4.3 - Servicios online

Con la facilidad de acceso a Internet de banda ancha que existe hoy en día en los países desarrollados, los negocios se han visto forzados a moverse al mundo virtual para poder seguir siendo competitivos. Además, la facilidad de comenzar un negocio online ha promovido que cada vez más hayan nuevos emprendedores ofreciendo sus servicios a través de Internet. Este boom de servicios electrónicos se debe en gran parte al auge del cloud computing, aproximadamente a partir de 2006[24][25] con la aparición del servicio Elastic Computer Cloud (EC2) de Amazon.

El cloud computing consiste en un modelo de infraestructura en el que el cliente no necesita disponer de equipos físicos, ahorrándose así el coste de mantener un lugar en el que colocar sus equipos, el coste del equipo en sí o el coste del personal que mantenga esos equipos. Se ofrecen tres tipos de servicios, los cuáles se pueden definir a grosso modo como:

- Infrastructure as a Service: se ofrecen recursos y es el cliente el que decide qué hacer con ellos.
- Platform as a Service: se ofrece un sistema funcional que el cliente usa según sus necesidades.
- Software as a Service: se ofrece software, librando al cliente de mantener la infraestructura.

Independientemente del tipo de servicio utilizado, todos tienen el mismo problema de cara a la privacidad: la pérdida de control físico de la información. Utilizando servicios cloud, es el proveedor del servicio el que se encarga de garantizar que no hayan fugas de información.

El modelo de negocio que más afecta a los usuarios es el de SaaS. Un ejemplo son los servicios de almacenamiento en la nube que existen hoy en día, siendo los más populares Dropbox, Google Drive o Apple iCloud.

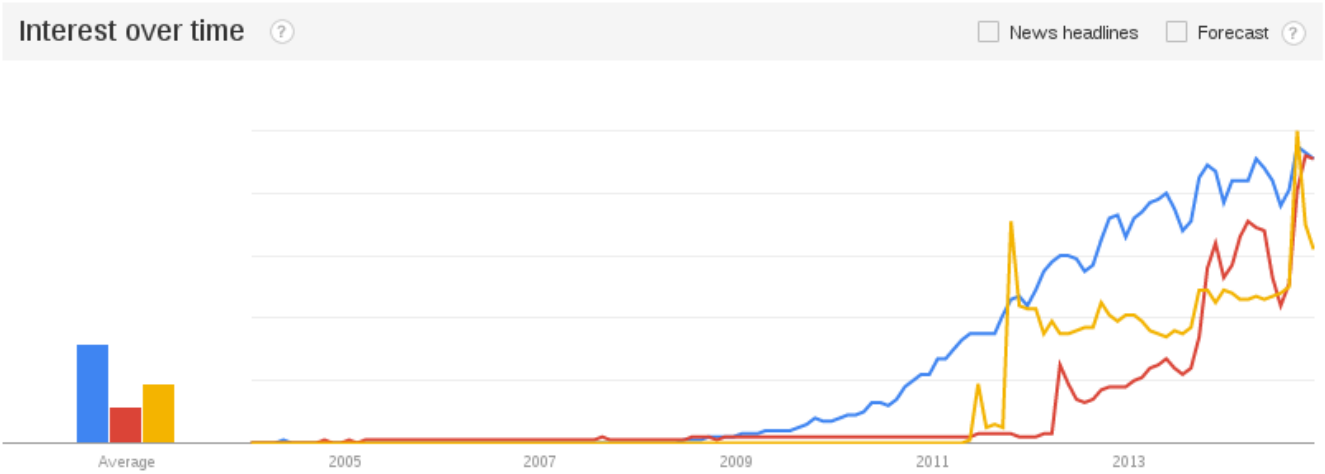

Figura 25 – Comparación entre Dropbox, Google Drive e iCloud

El problema de estos servicios radica en el uso que hacen los usuarios de ellos. No dispongo de datos contrastados, pero por experiencia puedo asegurar que un gran número de usuarios utiliza estos servicios como su método de copia de seguridad, subiendo a estas plataformas todos sus datos privados. Esto tiene un beneficio, y es que si por algún motivo se pierden los datos originales se tiene una copia para restaurarlos. Sin embargo, existen dos problemas evidentes:

- El proveedor del servicio es el encargado de la seguridad de los datos. Una mala gestión de la seguridad por parte del proveedor, y los datos del usuario pueden quedar expuestos. Además, el proveedor del servicio al tener el control de la información, puede acceder a los datos.

- Seguir una mala política de contraseñas o caer en un ataque de ingeniería social puede significar acceso no autorizado a los datos privados de un usuario.

Analizando el interés en la privacidad de cada uno de los servicios anteriores, se observa como el interés en Dropbox está por encima de las otras dos, lo que en parte es debido a que como se vió en la gráfica anterior, ha sido el servicio que más interés ha tenido de los de este tipo. Un dato a extraer de esta gráfica es que aunque el uso de los servicios en la nube está en auge desde el 2009, no es hasta el 2011 cuando empiezan a realizarse búsquedas sobre la privacidad de éstos.

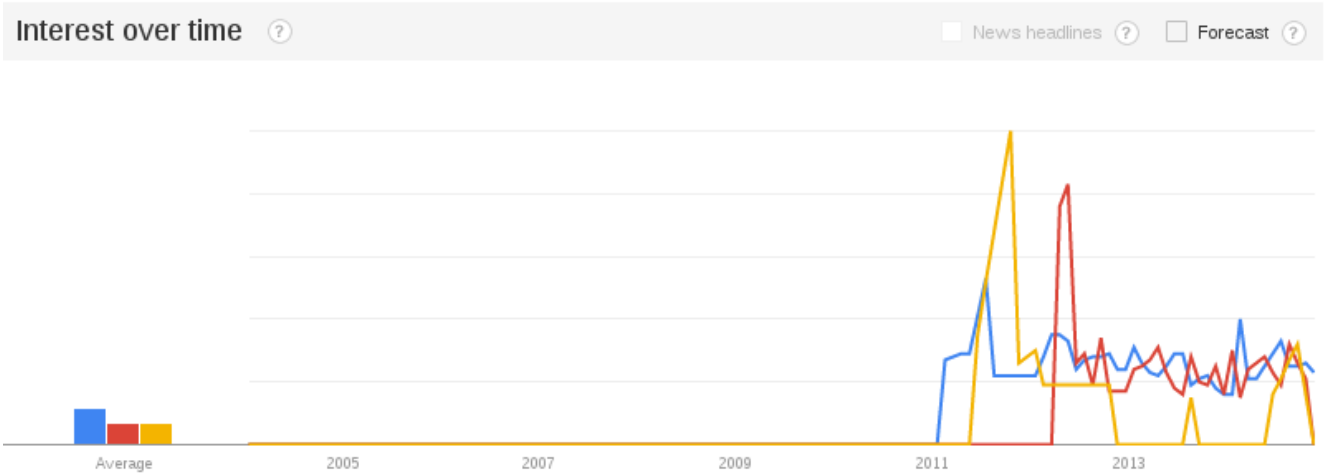

Figura 26 – Comparación de Dropbox privacy, Google Drive privacy e iCloud privacy

Sea por necesidad o por hobby, el uso de servicios online se volvió necesario para todos los usuarios, por lo que se necesitaba de tecnologías que permitieran estar conectados a estos servicios siempre que el usuario lo necesitara, y así aparecieron los smartphones. Un smartphone, o teléfono inteligente, es un dispositivo que va más allá de un teléfono móvil convencional, incorporando gran cantidad de características que lo convierten en un ordenador personal de bolsillo. La historia de un smartphone como tal se remonta a finales de los años 90, pero el verdadero auge de la era de la conexión a Internet 24/7 comienza en 2007 y 2008 con los lanzamientos de iPhone y Android[26].

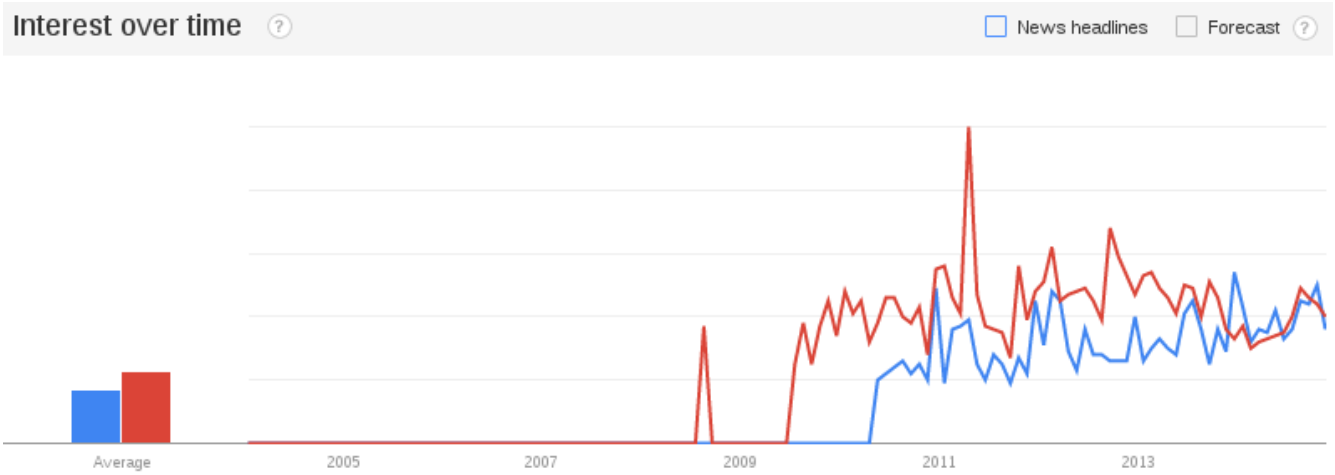

Figura 27 – Android privacy vs iPhone privacy

Desde que Android comenzó a ganar popularidad se observa como más o menos el interés en la privacidad de cada sistema ha estado equilibrado, generalmente estando un poco por encima el sistema de Apple. Destaca el pico originado en Abril de 2011 en el término de búsqueda *iPhone privacy*. Este pico tiene su explicación y es que en ese mes se descubrió que Apple estaba utilizando la geolocalización para controlar en todo momento donde se encontraban sus usuarios[27].

Los smartphones proporcionan prácticamente todo lo necesario para un usuario medio: Internet, redes sociales, GPS, fotografía y videos, juegos, etc, por lo que se han convertido es una herramienta primordial para nosotros, tal y como indica este informe de Nielsen[28]. Debido a este uso, cada vez más suelen haber más cantidad de datos personales almacenados en el dispositivo, datos que son de interés para agencias de inteligencia, delincuentes y para el negocio de la publicidad. Algunos de estos datos son[29]:

- Fotos y videos personales.
- Emails y SMS enviados y recibidos.
- Llamadas realizadas y recibidas.
- Contactos.
- Contraseñas.
- Datos financieros.
- Contenido del calendario.
- Localización, edad y sexo.

Estos datos se pueden extraer al menos a través de tres métodos distintos. El primero de ellos es por el uso de redes inseguras (wifi público), pero este no es problema exclusivamente de los smartphones, sino en general. Después tenemos el caso de los backdoors instalados en el sistema operativo por el propio propietario, que aunque todas las empresas siempre niegan la existencia de ellos, se ha demostrado[30] como sí que existen. Finalmente, las aplicaciones (apps). Un smartphone funciona a través de apps que el usuario instala para hacer uso de sus servicios, por ejemplo la aplicación de Facebook ofrece el servicio de Facebook en un smartphone sin necesidad de acceder a la página web. Cada vez que se va a instalar una aplicación se piden una serie de permisos que el usuario debe conceder, el problema es que muchos usuarios nunca leen estos permisos e instalan la aplicación sin más, y pueden estar otorgando a terceros el acceso a sus datos privados.

Un ejemplo es el uso de apps que pretenden parecerse a alguna app popular del momento. Por ejemplo, si se observan los permisos de la famosa aplicación Candy Crush Saga, se tiene:

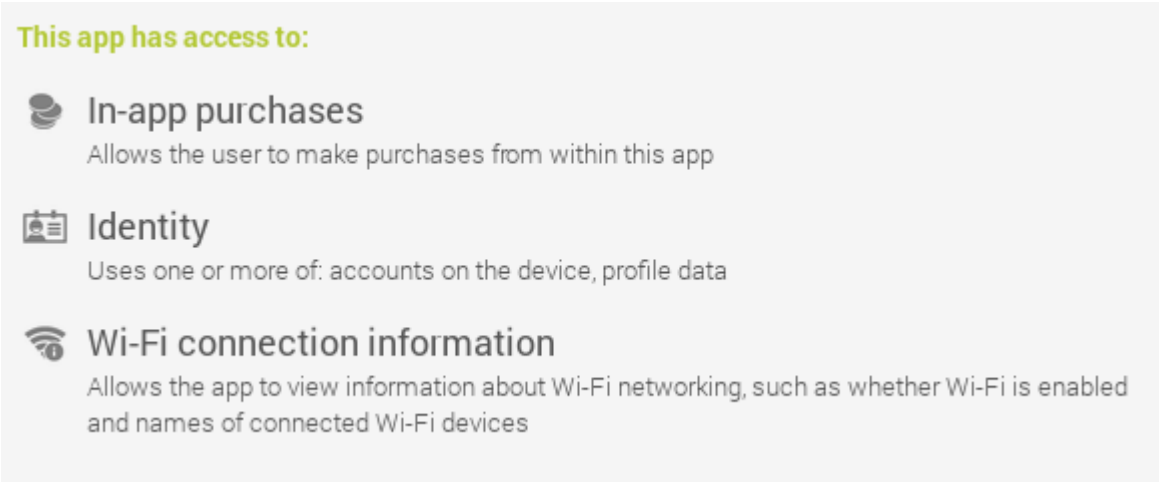

Figura 28 – Permisos Candy Crush Saga en Android.

Si ahora se miran los permisos que pide Candy Frenzy, una app que pretende ser un clon de Candy Crush Saga, se tiene que pide acceso a la localización, a las fotos y a la información de las llamadas, **todo esto para un juego!**

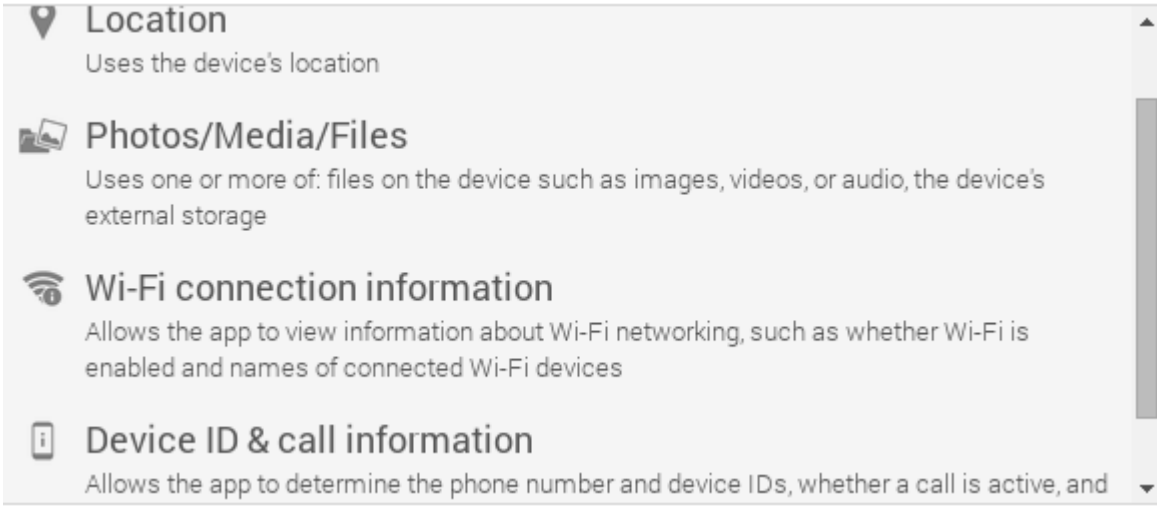

Figura 29 – Permisos de Candy Frenzy

Un usuario inteligente y consciente de los problemas de privacidad no instalaría esta aplicación. Sin embargo como se puede ver a continuación, **82406** personas han instalado esta app, ofreciéndoles sus datos privados al creador de Candy Frenzy.

Figura 30 – Usuarios que han instalado Candy Frenzy

Hemos visto que uno de los datos que interesa extraer de un smartphone es la localización. Un smartphone siempre se lleva encima, por lo que controlando donde se encuentra se tiene localizado a la persona en todo momento. Esto a dado pie a lo que se conocen como **servicios basados en localización**[31]. Estos servicios presentan un serio problema para la privacidad, ya que no sólo los proveedores de red serían capaz de localizar a una persona en cualquier momento, sino que cualquiera dueño de ese servicio tendría la posibilidad de localizar a sus usuarios. Este problema a la privacidad lo ha demostrado el MIT, realizando un estudio en el que concluyen que es posible identificar a casi cualquier persona basándose en los patrones de movimiento, concretamente al 95% de un conjunto de 1.5 millones[32].

4.4 - Tracking

Las tecnologías existentes hoy en día permiten que se pueda realizar un perfil de un usuario, llegando hasta el punto de identificar quién es la persona que se encuentra detrás de una pantalla. Como es de esperar, esta información es bastante valiosa y por eso hay una gran industria detrás, con empresas encargadas de recopilar estos datos para diferentes usos, generalmente para el marketing online.

No todo el tracking es malo. En un principio el tracking se utilizó para poder ofrece vistas personalizadas de una página a cada usuario. Por ejemplo, seleccionar el idioma en que se quiere ver la página, añadir artículos a un carrito de compras, mantener sesiones abiertas para no tener que volver a introducir credenciales... son algunos de los ejemplos de tracking bueno. Aunque todavía se le sigue dando este uso, cada vez se orienta más a lo citado anteriormente, la obtención de la mayor información posible sobre cada usuario.
Dos métodos clásicos que han existido desde siempre para llevar a cabo el tracking han sido el uso de cookies[33] y el uso de la cabecera referer de HTTP[34]. Las primeras almacenan información sobre la sesión y las preferencias del usuario, mientra que la segunda proporciona información sobre cómo ha llegado el usuario hasta el sitio.

El tracking por medio de cookies tiene una fácil solución, desactivarlas. Obviamente esto repercutirá en la interacción con el sitio (hasta el punto de no ser usable). A medida que las tecnologías avanzan, empiezan a aparecer nuevos métodos más avanzados de seguimiento de los que uno no se puede librar tan facilmente como de las cookies, algunos de los cuáles se comentan a continuación:

- Evercookies: o lo que es lo mismo, cookies permanentes. En un principio las cookies eran un simple fichero de texto que se borraba y adiós problemas. Sin embargo, para prevenir que esto ocurriera, se han identificado otros puntos en los que almacenar cookies tales como objetos de Flash, almacenamiento de Silverlight, cabeceras Etag o varias bases de datos de HTML5. Aprovechándose de esto, Samy Kamkar[35] creó Evercookies[36], un script que replica las cookies en todos estos puntos, de modo que se hace prácticamente imposible borrar las cookies.

- Browser fingerprinting: ¿Qué pensarías si te dijeran que existe un método capaz de identificarte por muchas medidas de anonimato que utilices y fueras capaz de librarte de las evercookies? Pues esta técnica existe desde hace relativamente poco. El experimento Panopticlick[37], llevado a cabo por la EFF[38], demostró como es prácticamente posible identificar inequívocamente a cualquier persona a través de varias propiedades del navegador.

Variable	Source	Remarks
User Agent	Transmitted by HTTP, logged by server	Contains Browser micro-version, OS version, language, toolbars and sometimes other info.
HTTP ACCEPT headers	Transmitted by HTTP, logged by server	
Cookies enabled?	Inferred in HTTP, logged by server	
Screen resolution	JavaScript AJAX post	
Timezone	JavaScript AJAX post	
Browser plugins, plugin versions and MIME types	JavaScript AJAX post	Sorted before collection. Microsoft Internet Explorer offers no way to enumerate plugins; we used the PluginDetect JavaScript library to check for 8 common plugins on that platform, plus extra code to estimate the Adobe Acrobat Reader version.
System fonts	Flash applet or Java applet, collected by JavaScript/AJAX	Not sorted; see Section 6.4.
Partial supercookie test	JavaScript AJAX post	We did not implement tests for Flash LSO cookies, Silverlight cookies, HTML 5 databases, or DOM globalStorage.

Figura 31 – Fuentes de información para browser fingerprinting

Para empeorar aún más las cosas, en Julio de este mismo año se publicó un paper[39] en el que se explica un nuevo método de browser fingerprinting basado en el Canvas de HTML5, el cuál se ha comprobado que está siendo utilizado ampliamente por todo Internet[40].

El tracking es una característica de bastante preocupación, y por ello se han intentado implementar medidas que lo impidan, como Do Not Track.

Do Not Track es una nueva cabecera propuesta para HTTP con el nombre DNT cuya función es indicar a la aplicación web que desactive sus opciones de seguimiento al usuario. Se encuentra disponible desde Diciembre del 2010, cuando Microsoft fue el pionero e incorporó esta característica en Internet Explorer 9. El resto de navegadores siguieron este mismo camino, y hoy en día todos tienen disponible la opción de Do Not Track.

Aunque es una gran mejora para la privacidad en Internet, se tiene el inconveniente de que por mucho que el usuario active esta opción, no existe una ley detrás que obligue a la aplicación web a desactivar sus opciones de seguimiento, por lo que simplemente pueden hacer caso omiso a esta cabecera y seguir funcionando tal cuál, monitorizando al usuario. En consecuencia, Do Not Track no se puede considerar una solución definitiva para la lucha contra la privacidad[41]. Existen empresas que si hacen honor a esta opción, como por ejemplo Twitter[42], mientras que otras como Yahoo[43] han decidido ignorar esta opción.

La falta de una ley detrás que obligue a cumplir el objetivo de Do Not Track supongo que será uno de los motivos, posiblemente el principal, por el que el interés en DNT ha caído.

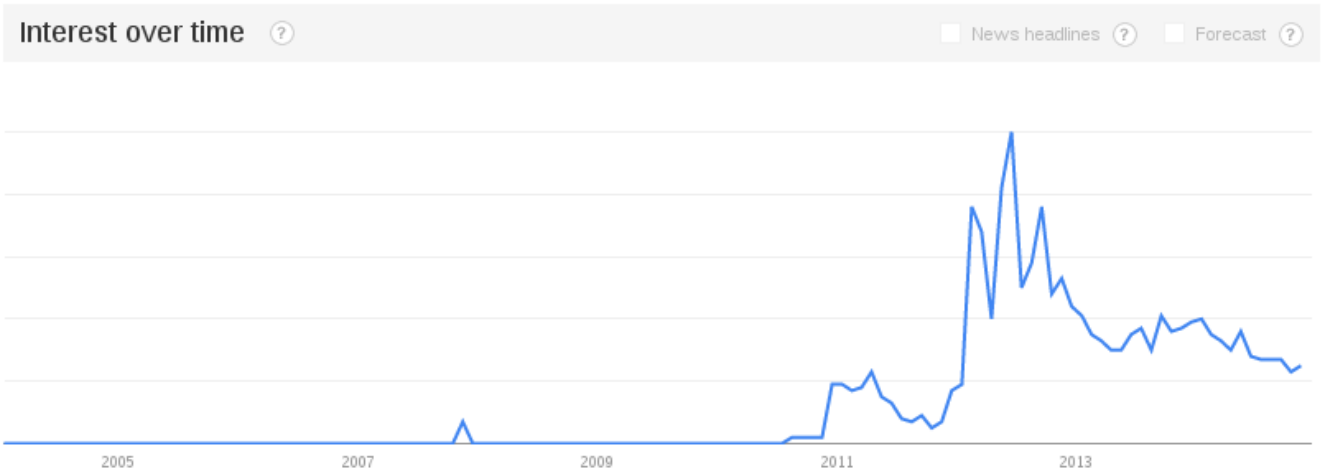

Figura 32 – Interés en Do Not Track

4.5 - Privacy by design

Privacy by design es un concepto que hace referencia a que se tenga en cuenta la privacidad desde el diseño del servicio en vez de cuando éste esté en funcionamiento. La principal precursora de este concepto ha sido Ann Cavoukian[44] que lleva desde los años 90 promoviendo esta cultura, aunque como se puede comprobar en esta gráfica, el interés general ha sido más bien reciente.

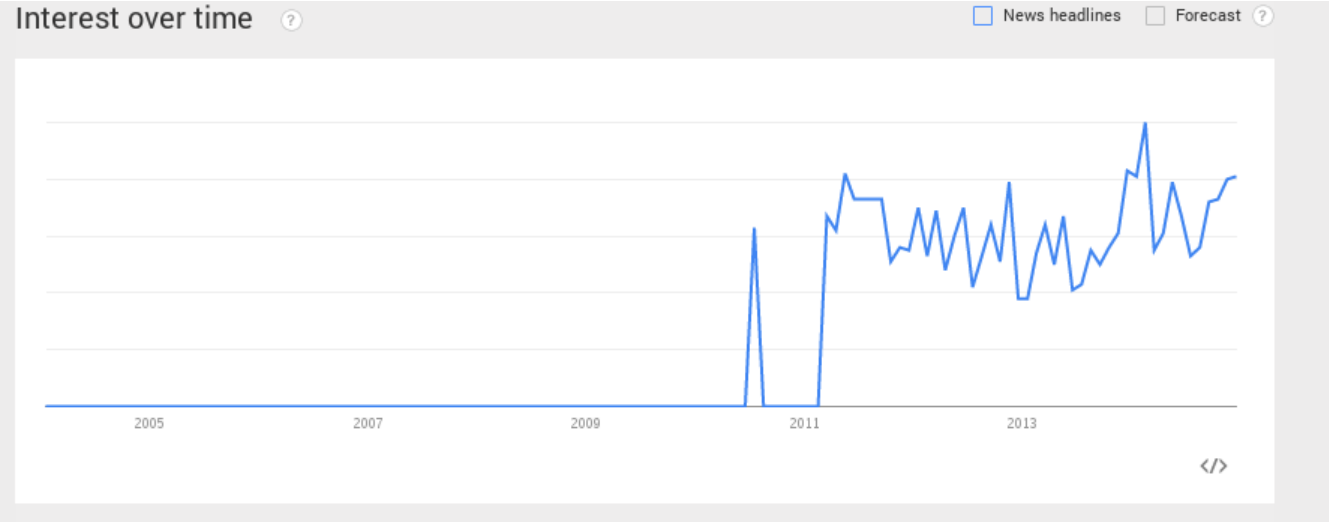

Figura 33 – Interés en “privacy by design”

La idea del privacy by design se basa en los siguiente siete principios:

- Prevenir en vez de remediar.
- Privacidad como la opción por defecto.
- Privacidad embebida con el diseño, no como un add-on.
- Ver la privacidad como un aliado no como un enemigo (caso privacidad vs seguridad por ejemplo).
- Proteger la información durante todo su ciclo de vida.
- Visibilidad y transparencia.
- Facilitar a los usuarios el acceso a las opciones de privacidad.

Algunos ejemplos en los que se tiene en cuenta este concepto son:

- Microsoft, como parte de su Security Development Cycle[45], tiene en cuenta la privacidad desde el propio diseño.
- Google+. Esta red social no fue la primera por parte de Google. Le precedió Buzz, red por la que Google tuvo problemas legales debido a violaciones de privacidad[46]. Para el desarrollo de Google+, se tomó la privacidad en cuenta desde su diseño.
- Icono de localización en smartphones. Aunque parece algo insignificante, el icono que indica cuando están los servicios de ubicación activos en un smartphone es parte de un proceso de privacy by design.

5 - Leyes y programas de espionaje

En esta siguiente sección se comentarán algunas de las leyes que se han creado en los últimos años y que afectan a la privacidad, así como programas esponsorizados por el Gobierno de EEUU y UK.

5.1 - U.S. Laws: FISA, PATRIOT Act y FREEDOM Act

Debido a que Estado Unidos ha sido la primera potencia mundial (1872-2014)[47] durante todo el desarrollo de la era digital, sus leyes afectan al resto del mundo ya sea de manera directa o indirecta. Basadas principalmente en cuestiones de guerra, se han firmado una serie de actas en las que se recogen leyes para regular la recopilación de inteligencia.

La primera de estas actas es la FISA[48], Foreign Intelligence Act, creada en 1978 para especificar los procedimientos para recolectar inteligencia extranjera por medios físicos o electrónicos. El espionaje a través de medios electrónicos es el tema que interesa en este proyecto. La FISA permitía el espionaje a potencias extranjeras, pero no a ciudadanos americanos. Sin embargo, tras los ataques del 11S y de anthrax sufridos en 2001, la cosa cambió, creándose la Patriot Act.

La USA Patriot Act[49], Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act, legaliza una serie de acciones con el objetivo de mejorar la lucha contra el terrorismo. De estas acciones, las principales relacionadas con la privacidad de los usuarios son[50]:

- Otorgar a las autoridades el privilegio de poder monitorizar y espiar a los ciudadanos siempre que sea con el objetivo de prevenir el terrorismo. El gobierno se basó en que esto mismo ya se hacía para luchar contrar el tráfico de drogas y las mafias, por lo que no veía por qué no hacerlo para el terrorismo.
- Permiso para realizar escuchas telefónicas itinerantes, lo que significa que si se tiene permiso para realizar escuchas telefónicas del teléfono de un presunto criminal, y éste se desprende del teléfono, no se requiera de una nueva orden para escuchar el nuevo dispositivo.
- Cuando se obtiene una orden judicial, se debe informar a la persona que le afecte que se va a ejecutar tal orden. Con el objetivo de prevenir que el posible terrorista se pueda deshacer de evidencias, se ha añadido una ley que permite ejecutar la orden desde el momento en que se informa.
- Permite a los agentes federales pedir órdenes judiciales para obtener los registros de los negocios.

Si bien la Patriot Act está hecha para combatir el terrorismo, también es un medio por el que legalmente se autoriza la violación de la privacidad de los ciudadanos, ya que **todo el mundo es sospechoso hasta que se demuestre lo contrario**.

En la siguiente imagen se puede observar como el interés ha decaído con el paso de los años, lo cuál puede significar que al principio hubo mucho revuelo sobre esta Ley, pero que hoy en día es algo que ya está asumido y por lo tanto no levanta interés.

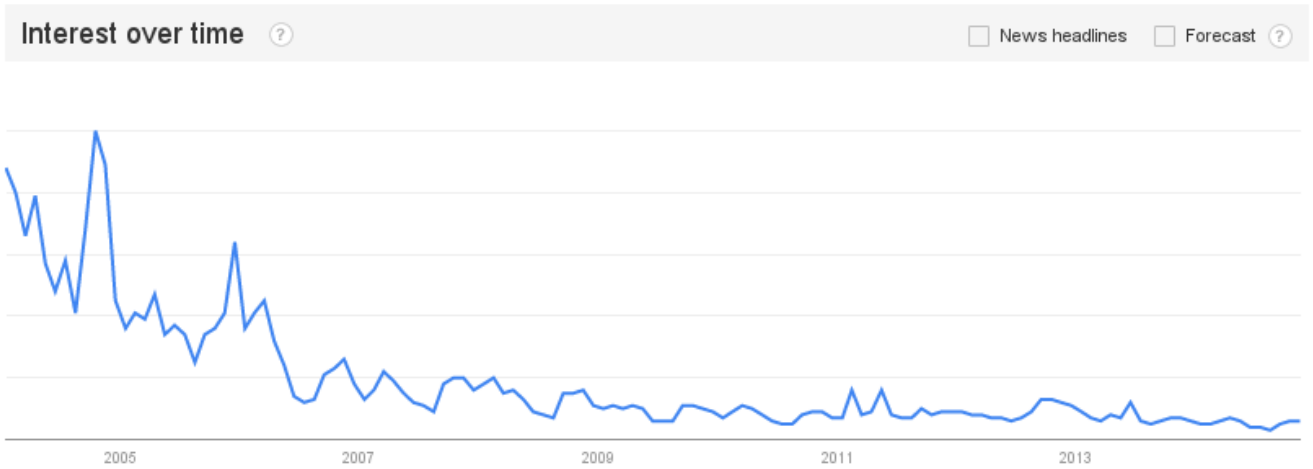

Figura 34 – Interés en “Patriot Act”

Realizando análisis cuantitativo con MediaCloud se observa como la frecuencia de los términos FISA y Patriot Act ha sido mínima, salvo los primeros meses tras las revelaciones de Snowden, lo cuál refuerza la afirmación del párrafo anterior.

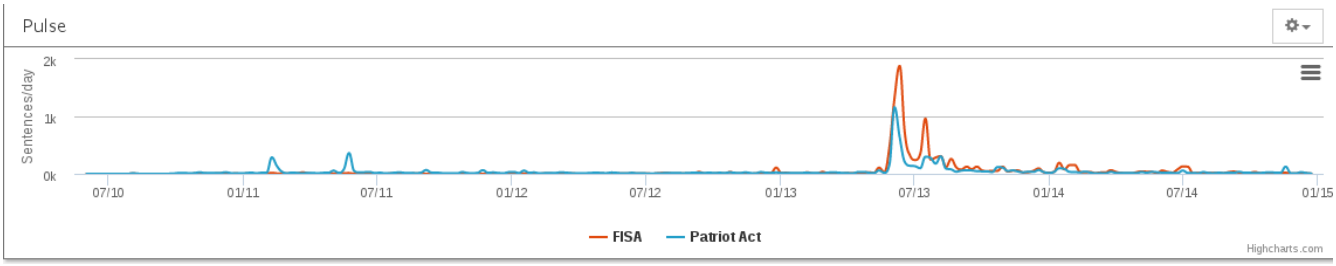

Figura 35 – Frencuencia de “FISA” y “Patriot Act”

Actualmente la Patriot Act se encuentra vigente hasta el 1 de Junio de 2015, pero ya hay grandes empresas de Internet como Google que están haciendo presión para que se realice una reforma, en beneficio de la libertad, de cara a la próxima posible extensión de la Patriot Act[51].

Hablando de reformas, una de las propuestas tras las revelaciones de Edward Snowden fue la USA Freedom Act[52], Uniting and Strengthening America by Fulfilling Rights and Ending Eavesdropping, Dragnet-collection and Online Monitoring Act. Con esta ley se pretendía limitar los poderes de la NSA, deteniendo la recopilación de datos masivos de los ciudadanos americanos por parte de esta agencia. Sin embargo, se sometió a votación el 18 de Noviembre, y de los 60 votos positivos que se requerían para que al menos se considerara establecer un debate sobre esta reforma, sólo se consiguieron 58. La mayoría de los senadores republicanos se negaron a votar a favor de la USA Freedom Act bajo la premisa de que se dejaría al país expuesto a ataques terroristas[53].

5.2 - Espionaje de la NSA

Tras los ataques del 11S, se han creado diferentes programas de espionaje llevados a cabo por la agencia de inteligencia de los Estados Unidos, la NSA. Entrar en detalles sobre este asunto resultaría en un proyecto completo, por lo que a continuación se realiza un análisis de modo general sobre como ha sido la evolución de este espionaje llevado a cabo por los EEUU desde el 11S hasta que se descubren los cuatro programas de recolección de información: MAINWAY, MARINA, NUCLEON y PRISM.

Antes de entrar a analizar este asunto, hay que introducir el caso de Edward Snowden ya que gracias a éste se conocen muchos de los datos que van a ser utilizados a continuación. Edward Snowden[54] fue un analista que trabajó para la NSA como contratista, periodo durante el cuál obtuvo información confidencial sobre los programas de espionaje llevados a cabo por la NSA junto a otras agencias de inteligencia internacionales, principalmente el GCHQ[55] de Reino Unido. El 1 de Junio de 2013, comenzó a filtrar esta información a diversos periódicos internacionales, desenmascarando así estos programas y detonando el interés en la privacidad a partir de entonces, como se ha visto en sucesivas gráficas anteriormente. En la siguiente se puede observar como desde que sale a la luz, a diario se publican textos haciéndole referencia.

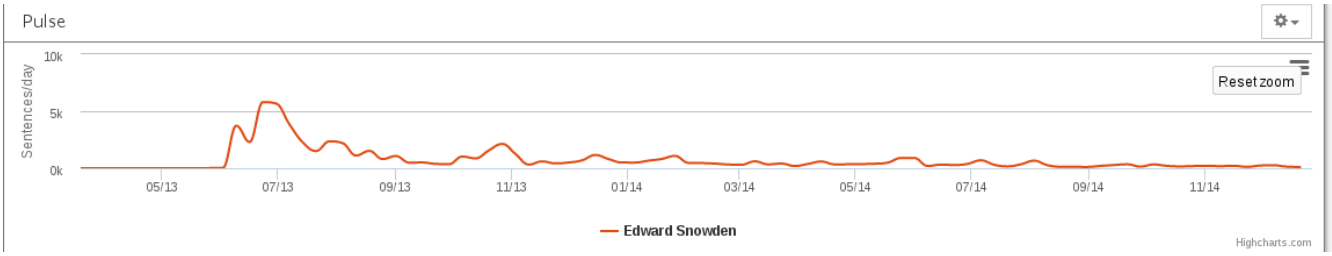

Figura 36 – Frecuencia de “Edward Snowden”

Ahora que se conoce cuál es la principal de fuente de información de los programas de espionaje de la NSA, se puede comenzar a analizar la evolución.

Todo comenzó con un artículo[56] en el New York Times el 16 de Diciembre del 2005, en el que se revelaba que George W. Bush había secretamente autorizado a la NSA a monitorizar las comunicaciones de los ciudadanos sin pasar por el FISC[57]. El mismo Presidente confirmó el día siguiente la existencia de tal programa[58], conocido como Terrorist Surveillance Program[59], el cuál forma parte de una serie de programas secretos de recolección de inteligencia con el nombre clave Stellar Wind[60]. En el 2008 se firma la FISA Amendments Act of 2008[61], legalizando las actividades llevadas a cabo entre 2001 y 2007 por el programa secreto del Presidente Bush.

El 11 de Mayo de 2006, USA Today publicó un artículo[62] en el que se revela que la NSA está almacenando datos de los registros teléfonicos de AT&T, Verizon y BellSouth. Hoy en día, gracias a Snowden, se sabe que esta base de datos se llama MAINWAY[63] y que almacena los metadatos de las

llamadas realizadas pero no su contenido, llamadas tanto nacionales como internacionales. Además se sabe que existe una base de datos similar para almacenar los datos del tráfico de Internet, llamada MARINA[64]. La monitorización de Internet se llevaba a cabo en una habitación conocida por Room 641A[65], donde se interceptaba el tráfico en el mismísimo backbone de Internet, es decir, en unos de los routers principales de Internet.

Tanto MAINWAY como MARINA se encargan de la recolección de inteligencia a través de metadatos. Para la recopilación a través del contenido se usan los programas PRISM y NUCLEON.

De NUCLEON es del que menos datos se conoce, aunque se sabe cuál es su propósito: guardar el contenido de las llamadas.

De PRISM[66], desde que Snowden filtrara los documentos sobre su existencia el 6 de Junio de 2013[67] se conocen más detalles. A modo general, comenzó su funcionamiento en el 2007 y su propósito es la recopilación de información accediendo directamente a servidores de algunas de las grandes empresas tecnológicas, las de mayor beneficio desde el punto de vista de la inteligencia. Las empresas que forman parte del programa son (por orden cronológico de su unión): Microsoft, Yahoo, Google, Facebook, PalTalk, YouTube, Skype, AOL y Apple.

5.3 - Derecho al olvido en Internet

El Derecho al olvido en Internet[68], relacionado con el Habeas Data[69], hace referencia al derecho que todo usuario tiene de que sus datos personales sean eliminados de los buscadores, de modo que datos pasados desactualizados y que afecten a su privacidad y reputación no puedan ser encontrados por medio de los mismos.

Desde Mayo de 2014, la Unión Europea reconoce este derecho, mientras que Bing y Yahoo comienzan a verse afectados por él desde Diciembre[70].

El caso que detonó este reconocimiento por parte de la Unión Europea fue el de Mario Costeja[71]. Mario es un español que vió como su pasado le perseguía cuando el periódico La Vanguardia digitalizó su hemeroteca, y ésta fue indexada por Google. En su momento, en 1998, dos viviendas que pertenecían al señor Costeja salieron a subasta por embargo, y aunque ese tema estaba zanjado, la información seguía publicada en la red, por lo que cualquiera que buscara su nombre lo encontraría como un moroso. Tras contactar tanto con el periódico como con Google, pasar por la Agencia Española de Protección de Datos, la Audiencia Española y el Tribunal Europeo de Justicia, el tema se solucionó con el reconocimiento del Derecho al olvido en Europa. El proceso comenzó desde 2009.

El reconocimiento por parte de la UE ha tenido una rápida aceptación entre los ciudadanos. En el caso de Google, tres semanas después de que publicara el formulario por el que se solicita que se ejerza el Derecho al olvido, ya se habían registrado 41000 solicitudes[72].

Este tema presenta una controversia y es que hay enfrentamiento de dos derechos. Por un lado el derecho al olvido y por otro el derecho a la libertad de expresión. Por este motivo Google ha anunciado que se estudiarán los casos individualmente, para garantizar un equilibrio entre ambos derechos. También han confirmado que información de interés sobre personajes públicos no será eliminada.

6 - Cibercrimen y privacidad

El mundo del cibercrimen ha crecido rápidamente durante los últimos años. El auge de los servicios online y de la era de la información, ha supuesto un hecho ideal para que los delincuentes puedan realizar su actividades delictivas de una manera más sencilla, pudiendo realizarlas a mayor escala mientras que exponen su identidad mucho menos comparado con el crimen ordinario. La siguiente imagen muestra como ha aumentado la frecuencia del término en estos últimos cuatro años, siendo hoy en día más del triple comparado con 2011 y 2012. Los resultados corresponden tanto a “cyber crime” como a “cybercrime”, ya que se hace referencia de ambos modos.

Figura 37 – Frecuencia del término “cybercrime”

Los resultados en español demuestran una situación desconcertante, y es que primero, previo a Abril del 2013 no había prácticamente ningún texto en el que se hablara de “cibercrimen” o “ciber crimen”, y segundo, comparada la frencuencia con el término en inglés se puede comprobar como es mucho menor.

Figura 38 – Frecuencia del término “cibercrimen”

El cibercrimen es un campo bastante amplio con diferentes vectores para llevarlo a cabo. Ataques de denegación de servicio, pornografía infantil, fraude (scams), blackhat SEO[73]... son algunos de los tipos de cibercrimen existentes. En este caso nos interesan principalmente dos que afectan a la privacidad de los usuarios: el cyberstalking y el robo de identidad.

6.1 - Cyberstalking y cyberbullying

Anteriormente, en la sección de los problemas de las redes sociales, ya se introdujo el concepto del cyberstalking. Recordando, el cyberstalking consiste en utilizar los medios digitales para acosar a una persona. Un término relacionado es el cyberbullying, término que se utiliza sobretodo cuando se refiere a acoso infantil (entre menores). En las gráficas siguientes vemos como ambos temas despiertan un gran interés, en especial este último, ya que es un tema de bastante importancia hoy en día con el fácil acceso que disponen los menores a servicios digitales.

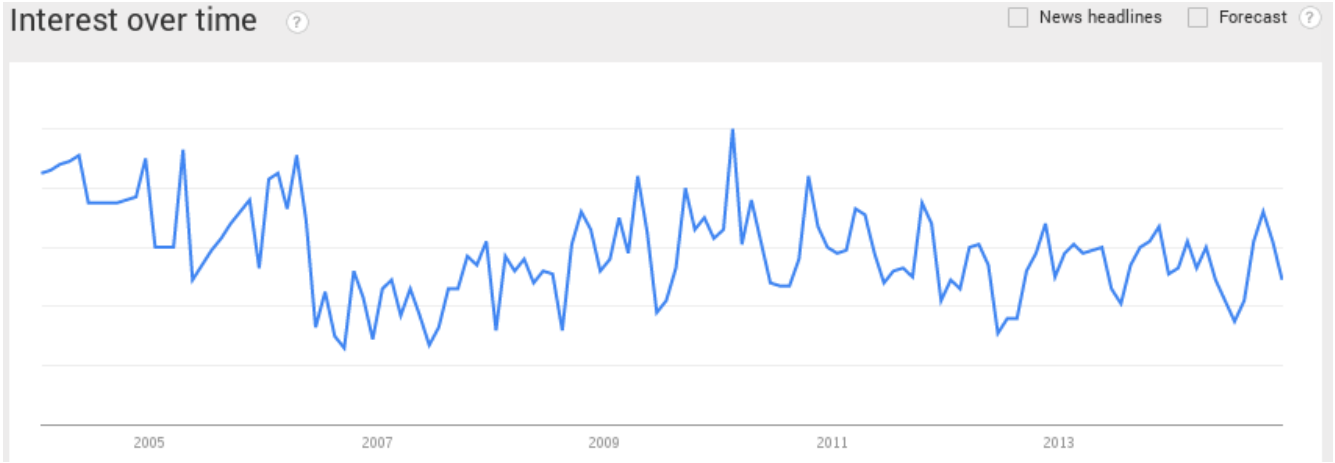

Figura 39 – Interés en “cyberstalking”

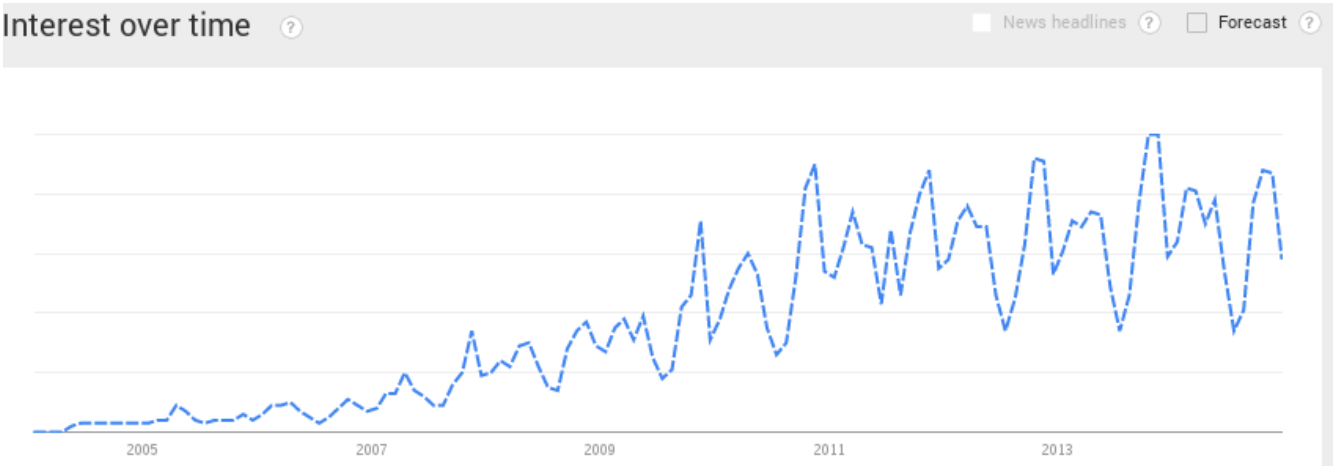

Figura 40 – Interés en “cyberbullying”

En el caso de España vemos como el interés en el ciberacoso es relativamente nuevo comparado con el resto del mundo.

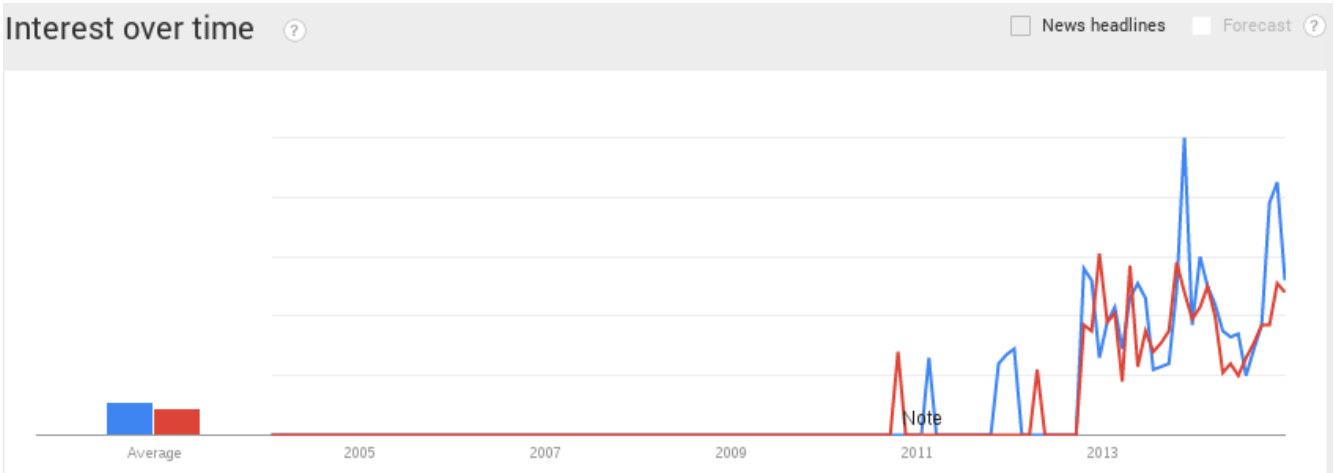

Figura 41 – Búsquedas por ciberacoso y cyberbullying en España

Estos ataques son sencillos de llevar a cabo, debido a la gran cantidad de sitios en los que se comparte información, y sobretodo, al mal uso que se da de los mismos. Por ejemplo, volviendo al caso de la red social, posteando datos personales como fotos de carnets o tarjeta de débito / crédito[74].

El objetivo de este ataque viene determinado por la propia mentalidad del atacante. Un posible objetivo, sobretodo en cyberstalking, es la extorsión, extrayendo la mayor información posible de la víctima buscando situaciones comprometidas con las que extorsionarle. Otro posible objetivo, y es el que más se suele dar en el cyberbullying, es denigrar a la víctima. Cuando se es adolescente, y también adulto, la reputación es uno de los valores que más afectan a una persona, y por medio de estos ataques se pretende mermar dicha reputación. Este último caso es muy importante, ya que se han dado casos de suicido de adolescentes por culpa de este tipo de ataques[75].

6.2 - Robo de identidad

El tráfico de datos personales es uno de los negocios que se lleva a cabo en el mundo del cibercrimen. Por datos personales se conoce como todo dato que sirva para identificar a una persona tales como nombre y apellidos, número de Seguridad Social, fecha de nacimiento, dirección, número de teléfono, email... Estos datos se utilizan posteriormente para diferentes funciones, como pueden ser fraude financiero, creación de identidades falsas o tráfico de información.

En MediaCloud se observa como es un tema del que poco a poco se va hablando más.

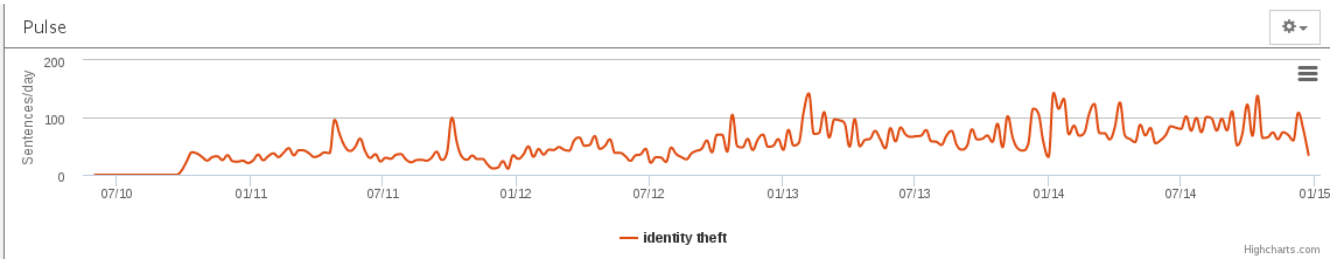

Figura 42 – Frencuencia del término “identity theft”

A continuación se citan algunas de las técnicas utilizadas para la obtención de este tipo de información personal:

- OSINT: OSINT[76] hace referencia a la extracción de datos de fuentes públicas, por ejemplo utilizando buscadores para ver que información hay disponible sobre una persona para que cualquiera pueda leerla.
- Dumpster diving (buscar en la basura): Gran cantidad de papeles con información personal son arrojados a la papelera a diario, por ello una buena fuente para extraer este tipo de información es buscar en la basura documentos para posteriormente analizar si tienen datos personales.
- Fugas de información: cada vez son más las noticias sobre que X empresa ha sufrido una intrusión en la que los datos personales de sus usuarios/personal/clientes se han visto expuestos. Según datos del Identity Theft Center[77], en 2014 se han registrado un total de 761 brechas de

datos, de las cuáles se han obtenido más de 83 millones de registros con datos personales (83176279). Este tema es muy importante, ya que como se pudo ver este mismo año 2014 con las revelaciones de fotos de famosas[78] o de usuarios de Snapchat[79], la privacidad de nuestros datos está íntimamente ligada con la seguridad de las compañías.

- Phising: los ataques de phising son cada vez más comunes. Consisten en engañar al usuario suplantando la identidad de un sitio, para que sea el mismo (el usuario) el que entregue información personal a los atacantes.
- Malware: el malware también tiene su uso para el robo de identidad, en especial troyanos que recopilen todo tipo de documentos y fotos, y que registren las pulsaciones del teclado (keylogger[80]).
- Ofertas de empleo falsas: otro método por el que recopilar gran cantidad de datos personales es el uso de ofertas de empleo falsas, recopilando todos los CVs recibidos para su posterior análisis.
- Ingeniería social: consiste en hacer que la víctima revele información confidencial por medio de técnicas que no requieren una base tecnológica, aunque frecuentemente se usa el phising como parte del ataque.

7 - Conclusión

A lo largo de este análisis sobre el estado actual de la privacidad en Internet se ha visto como es un tema que cada vez tiene más interés, debido a la preocupación que existe sobre los avances de la tecnología y como éstos están afectando a la privacidad.

Se han visto diferentes tecnologías que afectan a la privacidad, tanto para bien como para mal. No existe una solución definitiva para garantizar la privacidad de un usuario, por ello, el mejor modo para hacer frente a problemas de privacidad es tomar consciencia del asunto y aplicar de primera mano medidas para prevenir que los datos personales queden expuestos. Entre estas medidas, la primordial es conocer que información puede ser publicada, y cuál no debe ser nunca compartida.

Hay que agradecer a Edward Snowden por sus reveleaciones, ya que como se ha visto, el interés en este tema comenzó a crecer tras el conocimiento de los programas de espionaje de la NSA. Aunque muchos (con perfil tecnológico) ya pensábamos que este tipo de espionaje existía, las filtraciones de Snowden y la repercusión que tuvieron en todos los medios hicieron que el público creciera mucho más.

Concluyo este estudio reiterando el tema de la consciencia sobre la privacidad, tanto como usuario como proveedor de servicio. Usuarios, controlen la información que comparten. Proveedores, implementan medidas de seguridad y auditen su eficacia constantemente, además de seguir un proceso de privacy-by-design para todos los desarrollos. Como se ha visto, privacidad y cibercrimen son dos temas que van de la mano, y una cosa se puede asegurar con certeza: **el cibercrimen no disminuirá en los próximos años.**

8 - Referencias

1. MediaCloud. http://mediacloud.org/about/
2. Berkman Center for Internet and Society. http://cyber.law.harvard.edu/
3. MIT Center for Civic Media. https://civic.mit.edu/
4. Google Trends. http://www.google.com/trends/
5. EbizMBA. *Top 15 Most Popular Search Engines | December 2014.* http://www.ebizmba.com/articles/search-engines
6. Matthew Kushinka. *The most widely spoken languages in the world.* http://www.redlinels.com/2014/01/10/most-widely-spoken-languages/
7. Google Official Blog. *Updating our privacy policies and terms of service.* http://googleblog.blogspot.com.es/2012/01/updating-our-privacy-policies-and-terms.html
8. Google Public Data. https://www.google.es/publicdata/explore?ds=d5bncppjof8f9_&ctype=l&strail=false&bcs=d&nselm=h&met_y=sp_pop_totl&scale_y=lin&ind_y=false&rdim=region&idim=country:ITA:USA&ifdim=region&tstart=-287020800000&tend=1385510400000&hl=en&dl=en&ind=false&icfg&iconSize=0.5
9. Wikipedia. *Virtual private newtork.* http://en.wikipedia.org/wiki/Virtual_private_network
10. The Tor Project. https://www.torproject.org/
11. Wikipedia. *Friendster.* http://en.wikipedia.org/wiki/Friendster
12. Caroline McCarthy. *Friendster awarded 'compatibility scoring' patent.* http://www.cnet.com/news/friendster-awarded-compatibility-scoring-patent/
13. Caroline McCarthy. *Farewell, 2003: The Friendster apocalypse is nigh.* http://www.cnet.com/news/farewell-2003-the-friendster-apocalypse-is-nigh/
14. Michael Arrington. *Facebook No Longer The Second Largest Social Network.* http://techcrunch.com/2008/06/12/facebook-no-longer-the-second-largest-social-network/
15. Declan McCullagh. *Instagram says it now has the right to sell your photos.* http://www.cnet.com/news/instagram-says-it-now-has-the-right-to-sell-your-photos/
16. Wikipedia. *Ehealth.* http://en.wikipedia.org/wiki/EHealth
17. Chema Alonso. *Róbame que estoy haciendo deporte y estoy así de sano.* http://www.elladodelmal.com/2014/07/robame-que-estoy-haciendo-deporte-y.html
18. Wikipedia. *Peter Chapman.* http://en.wikipedia.org/wiki/Peter_Chapman
19. Wikipedia. *Internet meme.* http://en.wikipedia.org/wiki/Internet_meme
20. Robert Glancy. *Will you read this article about terms and conditions? You really shoud do.* http://www.theguardian.com/commentisfree/2014/apr/24/terms-and-conditions-online-small-print-information
21. Wikipedia. *Tor (anonymity network).* http://en.wikipedia.org/wiki/Tor_%28anonymity_network%29
22. Tor Project Blog – arma. *Tor security advisory: “relay early” traffic confirmation attack.*

https://blog.torproject.org/blog/tor-security-advisory-relay-early-traffic-confirmation-attack
23. Wikipedia. *Internet in Norway.* http://en.wikipedia.org/wiki/Internet_in_Norway#Internet_censorship
24. Wikipedia. *Cloud computing.* http://en.wikipedia.org/wiki/Cloud_computing
25. Arif Mohamed. *A history of cloud computing.* http://www.computerweekly.com/feature/A-history-of-cloud-computing
26. Wikipedia. *Smartphone.* http://en.wikipedia.org/wiki/Smartphone
27. Alasdair Allan. *Got an iPhone or 3G iPad? Apple is recording your moves.* http://radar.oreilly.com/2011/04/apple-location-tracking.html
28. Nielsen. *How smartphones are changing consumers' daily routines around the globe.* http://www.nielsen.com/us/en/insights/news/2014/how-smartphones-are-changing-consumers-daily-routines-around-the-globe.html
29. Privacy Rights Clearinghouse. *Privacy in the Age of the Smartphone.* https://www.privacyrights.org/smartphone-cell%20phone-privacy
30. Jonathan Zdziarski. *Identifying Back Doors, Attack Points, and Surveillance Mechanisms in iOS Devices.* https://pentest.com/ios_backdoors_attack_points_surveillance_mechanisms.pdf?_ga=1.87765276.585154247.1400508931
31. Wikipedia. *Location based service.* http://en.wikipedia.org/wiki/Location-based_service
32. Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel. *Unique in the Crowd: The privacy bounds of human mobility.* http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html
33. Wikipedia. *HTTP Cookie*. http://en.wikipedia.org/wiki/HTTP_cookie
34. Wikipedia. *HTTP referer.* http://en.wikipedia.org/wiki/HTTP_referer
35. Wikipedia. *Samy Kamkar.* http://en.wikipedia.org/wiki/Samy_Kamkar
36. Samy Kamkar homepage. *Evercookie*. http://samy.pl/evercookie/
37. Peter Eckersley. *How unique is your web browser?*. https://panopticlick.eff.org/browser-uniqueness.pdf
38. Electronic Frontier Foundation. https://www.eff.org/
39. Gunes Acar, Christina Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan & Claudia Diaz. *The Web Never Forgets: Persisten Tracking Mechanisms in the Wild.* https://securehomes.esat.kuleuven.be/~gacar/persistent/the_web_never_forgets.pdf
40. Gunes Acar, Christina Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan & Claudia Diaz. *The Web Never Forgets: Persisten Tracking Mechanisms in the Wild.* https://securehomes.esat.kuleuven.be/~gacar/persistent/
41. Geoff Duncan. *Why do not track may not protect anybody's privacy.* http://www.digitaltrends.com/mobile/why-do-not-track-may-not-protect-anybodys-privacy/
42. Twitter Help Center. *Twitter supports Do Not Track.* https://support.twitter.com/articles/20169453-twitter-supports-do-not-track
43. Yahoo! Global Public Policy. *Yahoo's Default = A Personalized Experience.*

http://yahoopolicy.tumblr.com/post/84363620568/yahoos-default-a-personalized-experience
44. Wikipedia. *Ann Cavoukian.* http://en.wikipedia.org/wiki/Ann_Cavoukian
45. Microsoft Security Development Lifecyle. *SDL Process: Design.* http://www.microsoft.com/security/sdl/process/design.aspx
46. Wikipedia. *Google Buzz*. http://en.wikipedia.org/wiki/Google_Buzz#Privacy
47. Estela S. Mazo. *China es ya la primera potencia mundial.* http://www.expansion.com/2014/10/08/economia/1412771929.html?a=VO146b1c8a9d0531f55bb2f50be338cc33e&t=1419625771
48. Wikipedia. *Foreign Intelligence Surveillance Act.* http://en.wikipedia.org/wiki/Foreign_Intelligence_Surveillance_Act
49. Wikipedia. *Patriot Act.* http://en.wikipedia.org/wiki/Patriot_Act
50. USA Department of Justice. *The USA PATRIOT Act: Preserving Life and Liberty.* http://www.justice.gov/archive/ll/highlights.htm
51. Alex Wilhelm. *Google Claims 2015 Will be A "Moment" For Surveillance Reform.* http://techcrunch.com/2014/12/18/google-claims-2015-will-be-a-moment-for-surveillance-reform/
52. Wikipedia. *USA Freedom Act.* http://en.wikipedia.org/wiki/USA_Freedom_Act
53. Spencer Ackerman. *Senate Republicans block USA Freedom Act surveillance reform bill.* http://www.theguardian.com/us-news/2014/nov/18/usa-freedom-act-republicans-block-bill
54. Wikipedia. *Edward Snowden.* http://en.wikipedia.org/wiki/Edward_Snowden
55. GCHQ. http://www.gchq.gov.uk/Pages/homepage.aspx
56. James Risen. *Bush Lets U.S. Spy on Callers Without Courts.* http://www.nytimes.com/2005/12/16/politics/16program.html?pagewanted=all&_r=0
57. Wikipedia. *United States Foreign Intelligence Surveillance Court.* http://en.wikipedia.org/wiki/United_States_Foreign_Intelligence_Surveillance_Court
58. David E. Sanger. *Bush Says He Ordered Domestic Spying.* http://www.nytimes.com/2005/12/18/politics/18bush.html?pagewanted=all
59. Wikipedia. *Terrorist Surveillance Program.* http://en.wikipedia.org/wiki/Terrorist_Surveillance_Program
60. Wikipedia. *Stellar Wind.* http://en.wikipedia.org/wiki/Stellar_Wind
61. Wikipedia. *Foreign Intelligence Surveillance Act of 1978. Amendments Act of 2008.* http://en.wikipedia.org/wiki/Foreign_Intelligence_Surveillance_Act_of_1978_Amendments_Act_of_2008
62. Leslie Cauley. *NSA has massive database of Americans' phone calls.* http://usatoday30.usatoday.com/news/washington/2006-05-10-nsa_x.htm
63. Wikipedia. *MAINWAY.* http://en.wikipedia.org/wiki/MAINWAY
64. Wikipedia. *MARINA.* http://en.wikipedia.org/wiki/MARINA
65. Wikipedia. *Room 641A.* http://en.wikipedia.org/wiki/Room_641A
66. Wikipedia. *PRISM (surveillance program).*

http://en.wikipedia.org/wiki/PRISM_(surveillance_program)
67. Glenn Greenwald. *NSA Prism program taps in to user data of Apple, Google and others.* http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data
68. Wikipedia. *Derecho al olvido.* http://es.wikipedia.org/wiki/Derecho_al_olvido
69. Wikipedia. *Habea Data.* http://es.wikipedia.org/wiki/Habeas_data
70. Stuart Dredge. *Microsoft and Yahoo respond to European "right to be forgotten" requests.* http://www.theguardian.com/technology/2014/dec/01/microsoft-yahoo-right-to-be-forgotten
71. Óscar Corral. *Mario Costeja, el español que venció al todopoderoso Google.* http://www.lavanguardia.com/tecnologia/internet/20140514/54407896513/mario-costeja-google.html
72. Álvaro Ibáñez "Alvy". *Google ofrece un formulario para el "derecho al olvido" y recibe 41000 solicitudes en tres semanas.* http://www.rtve.es/noticias/20140611/google-ofrece-formulario-para-derecho-olvido-recibe-41000-solicitudes-tres-semanas/952526.shtml
73. Wikipedia. *Spamdexing.* http://en.wikipedia.org/wiki/Spamdexing
74. Loucif Kharouni. *The Dangers of Posting Credit Cards, IDs on Instagram and Twitter.* http://blog.trendmicro.com/trendlabs-security-intelligence/the-dangers-of-posting-credit-cards-ids-on-instagram-and-twitter/
75. NoBullying.com. *Bullying Videos: Tragic Victims and Heroes.* http://nobullying.com/cyber-bullying-stories-and-videos-tragic-victims-and-heroes/
76. Wikipedia. *Open Source Intelligence.* http://en.wikipedia.org/wiki/Open-source_intelligence
77. Identity Theft Resource Center. *Data Breach Reports. December 23, 2014.* http://www.idtheftcenter.org/images/breach/DataBreachReports_2014.pdf
78. Wikipedia. *2014 celebrity photo leaks.* http://en.wikipedia.org/wiki/2014_celebrity_photo_leaks
79. James Vincent. *"The Snappening": Has Snapchat been hacked? What is Snapsaved? Your questions answered.* http://www.independent.co.uk/life-style/gadgets-and-tech/the-snappening-has-snapchat-been-hacked-whats-snapsaved-your-questions-answered-9790658.html
80. Wikipedia. *Keylogger.* http://es.wikipedia.org/wiki/Keylogger

9 - Listado de figuras