

USOS Y USUARIOS DE SERVICIOS DE INFORMACIÓN CULTURAL EN INTERNET

Un estudio de analítica web

Javier García-Aráez Martín-Montalvo

Usos y Usuarios de Servicios de Información Cultural en Internet.

El objetivo principal del estudio es obtener datos sobre el funcionamiento habitual de servicios de información cultural en internet: cuántos usuarios hacen uso del mismo y en qué condiciones, y cuál es el uso que se hace de las funcionalidades y herramientas del mismo.

Para ello se van a analizar los datos obtenidos a lo largo de los años 2011, 2012, 2013 y 2014 (hasta el 30 de junio) en portales de Gobierno de Aragón DARA¹ Documentos y Archivos de Aragón y, en menor medida, SIPCA² Servicio de Información del Patrimonio Cultural Aragonés. En ambos casos, el estudio propuesto tratará de resolver estas necesidades de información mediante la configuración y uso de herramientas de analítica web.

Estudios de usuarios en el ámbito online

Herramientas de analítica web

En primer lugar, debe indicarse la especial naturaleza de este análisis, ya que a diferencia de otros estudios de usuarios de servicios culturales, en este caso los datos obtenidos se circunscriben específicamente al ámbito *online*: el usuario del que se van a obtener datos y cuyo comportamiento y uso del servicio va a analizarse nunca ha tenido –como norma general- un contacto físico con el servicio que pretende estudiar.

Su método de interacción con el servicio condiciona, de esta forma, el acceso a los datos sobre los mismos usuarios. Donde en un análisis convencional se disponen de herramientas de comunicación con el usuario y aún un contacto directo (una comunicación, un diálogo), en este caso es imposible este tipo de procedimiento.

Será por tanto necesaria la puesta en marcha de diferentes aproximaciones al problema y la utilización de diferentes herramientas de obtención de datos. Para recabar el tipo de datos necesarios para la elaboración de estudios de usuarios en internet, debe forzosamente acudir a herramientas de *analítica web*, de las que existen diferentes opciones, cada una dotada de sus ventajas e inconvenientes.

En este caso se ha seleccionado *Google Analytics* por su sencillo manejo e instalación y sobre todo por sus grandísimas posibilidades de configuración.

¹ <http://dara.aragon.es/opac/app/simple/>

² <http://www.sipca.es/>

Uso, potencialidades y limitaciones de datos de análisis web y de Google Analytics en particular

El primer problema al que se enfrenta el análisis de uso y de usuarios de un servicio cultural en internet es la separación física y la ausencia de interacción directa entre la gestión del servicio y el usuario que desea satisfacer una necesidad de información. Como se ha indicado, la manera de recabar datos para salvar esta brecha es la utilización de herramientas de analítica web.

Ahora bien, estas herramientas tienen una serie de limitaciones por su misma naturaleza y por el mecanismo de funcionamiento que despliegan, desde las mismas técnicas de obtención de datos hasta la interpretación que debe hacerse de los mismos.

Google Analytics utiliza *cookies*³ para la obtención de datos del usuario y del uso del servicio. Una *cookie* es un fragmento de texto almacenado en el navegador del usuario que interactúa con una web que se lo deposita previamente⁴. De esta forma, el servicio de información cultural sobre el que se desea obtener datos instala en su código una herramienta de analítica web -*Google Analytics*- que “incrustará” en el navegador del usuario un pequeño fragmento de texto. Estas cookies van a almacenar datos sobre el usuario y sobre el funcionamiento de la web, que serán ofrecidos a la herramienta de analítica para su depurado y utilización.

Por su propia naturaleza, la utilización de cookies para la obtención de datos presenta una serie de limitaciones que deben ser tenidas en cuenta a la hora de analizar los resultados almacenados.

En primer lugar, las cookies pueden ser borradas por parte de los usuarios⁵, lo que provoca una serie de distorsiones en los datos obtenidos. Este problema, aunque no incide directamente en los datos relativos a visitas obtenidas o utilización de las funcionalidades del servicio, sí que va a afectar a indicadores clave tales como “usuarios nuevos”, “usuarios recurrentes” (fidelizados) “frecuencia de visita” y otros. Es decir, que los datos obtenidos, por ejemplo, de *nuevos usuarios* no pueden ser utilizados como un valor absoluto, ya que es imposible saber en qué porcentaje estos usuarios detectados como nuevos son en realidad usuarios retornantes que han eliminado las cookies de sus navegadores.

Pero en ningún caso esta circunstancia invalida los datos obtenidos: esta limitación debe sortearse utilizando estos datos de forma comparativa (un segmento poblacional tiene

³ <http://www.w3.org/2001/tag/2010/09/ClientSideStorage.html>

⁴ Protección de usuarios del archivo online:

⁵ Hasta un 37% de los usuarios australianos borran mensualmente sus cookies
<http://www.comscore.com/Insights/Press-Releases/2011/2/comScore-Publishes-White-Paper-on-the-Impact-of-Cookie-Deletion-on-Website-Audience-Measurement-in-Australia>

un porcentaje mayor de usuarios recurrentes que otro) o evolutiva (el porcentaje de nuevos usuarios se incrementa trimestralmente).

En segundo lugar, las cookies se instalan en un navegador, no en un “usuario”. Sin embargo, las herramientas de análisis van a identificar a un navegador de internet como un usuario único. Esta conclusión es también errónea, ya que pueden darse dos circunstancias: un usuario único puede hacer uso del servicio cultural desde dos o más navegadores (por lo que será identificado como dos o más usuarios) o varios usuarios pueden hacerlo desde un mismo navegador (por lo que varios usuarios serían identificados como un solo usuario único).

Estas limitaciones son producto del forzoso desconocimiento directo de los usuarios, por la naturaleza online de este tipo de servicios. Como se ha dicho, en ningún caso invalidarán los datos obtenidos, pero pueden poner en riesgo la veracidad o exactitud de las conclusiones que se obtengan de los análisis de datos: deberán ser tenidos en cuenta para una interpretación correcta del análisis de usuarios de la web cultural.

Tipos de datos obtenidos para su análisis y configuración de la herramienta para obtención de informes personalizados y grupos poblacionales concretos

De la utilización de este tipo de herramientas de analítica web obtendremos dos grandes grupos de datos, según su naturaleza.

En primer lugar obtendremos **datos de los usuarios**, es decir, datos que reflejan información propia de quienes hacen uso del servicio online:

- Datos geográficos: procedencia geográfica de cada una de las conexiones al servicio.
- Datos tecnológicos: Proveedor de Internet (ISP), sistema operativo, acceso desde dispositivos móviles.
- Datos de comportamiento: Tiempo invertido en el servicio cultural (interacción), frecuencia de visita, nuevos usuarios/recurrentes, rebote.
- Datos de captación: propios de la naturaleza telemática del usuario, van a indicar el método mediante el cual los usuarios acceden a la web cultural (tráfico de búsqueda orgánica, referencia por enlace externo, tráfico directo).

En segundo lugar se almacenan **datos de uso del servicio**, referidos a qué funcionalidades, páginas, portales o subportales de los provistos se están utilizando, y cómo:

- Páginas vistas: URL individuales visitadas por los usuarios

- Páginas de destino: “Puerta de entrada” de los usuarios a la web, cuál es la primera página que utilizan los usuarios tras acceder al servicio.
- Uso de herramientas de búsqueda: Uso de las funcionalidades provistas por el servicio para realizar una consulta concreta, términos o cadenas de búsqueda, etc.
- Flujo de comportamiento: Página de entrada, páginas visitadas y página de salida del servicio.

Google Analytics va a almacenar datos tanto de usuarios como de uso del servicio, y va a ofrecer además una serie de informes y herramientas de configuración y de cruzado de datos que permitirán acceder a información depurada y concreta que responde, eventualmente, a las necesidades de información del gestor del servicio.

Para ello la herramienta dispone de una serie de posibilidades de configuración que van a hacer posible el análisis de datos específicos, el cruce y comparativa de indicadores concretos y el aislado para su estudio de grupos poblacionales concretos basados en variables específicas:

- Informes

Google Analytics pone a disposición de sus usuarios determinados informes, tanto de uso del servicio como de los mismos usuarios. Estos informes pueden ser a su vez configurados, permitiendo la visualización de una variable concreta en el marco general de un informe dado.

- Segmentación

La segmentación avanzada de datos es una de las funcionalidades más potentes de Google Analytics, y en el presente estudio se usará a diferentes niveles. Un segmento avanzado es un conjunto de datos, aislado en función de una o más variables a modo de filtro, que permite la visualización específica de conjuntos de datos. Por ejemplo, alguno de los segmentos avanzados que se han utilizado son “usuarios procedentes de Europa” o “usuarios que acceden al servicio vía *Europeana*, *Hispana* y otros de la red de enlaces oficiales del servicio”.

Configurado este segmento, puede ser utilizado para estudiar informes específicos de comportamiento y uso del servicio, para compararse con otros segmentos avanzados, o consigo mismo a lo largo del tiempo

- Filtrado de datos

El tercer nivel de configuración provisto por Google Analytics es el filtrado avanzado de datos. Este filtrado permite, sobre un informe determinado (y al que le podemos haber filtrado los datos mostrados mediante la activación de un segmento), aislar los resultados concretos de dicho informe que se consideren relevantes.

Estos tres niveles de configuración y filtrado de datos pueden y deben combinarse entre sí para seleccionar la información relevante para el estudio de usuarios y para acceder a variables y a datos específicos que están en un primer momento “ocultos” al visualizarse sin procesar ni refinar.

En la Figura 1 puede observarse un ejemplo que ilustra las posibilidades de depurado de información que ofrecen este tipo de herramientas. Sobre un *Informe* concreto se ha aplicado la visualización de un *segmento* específico de usuarios, y se han *filtrado* los resultados para que se muestren solo los más relevantes para obtener la información que se considere relevantes.

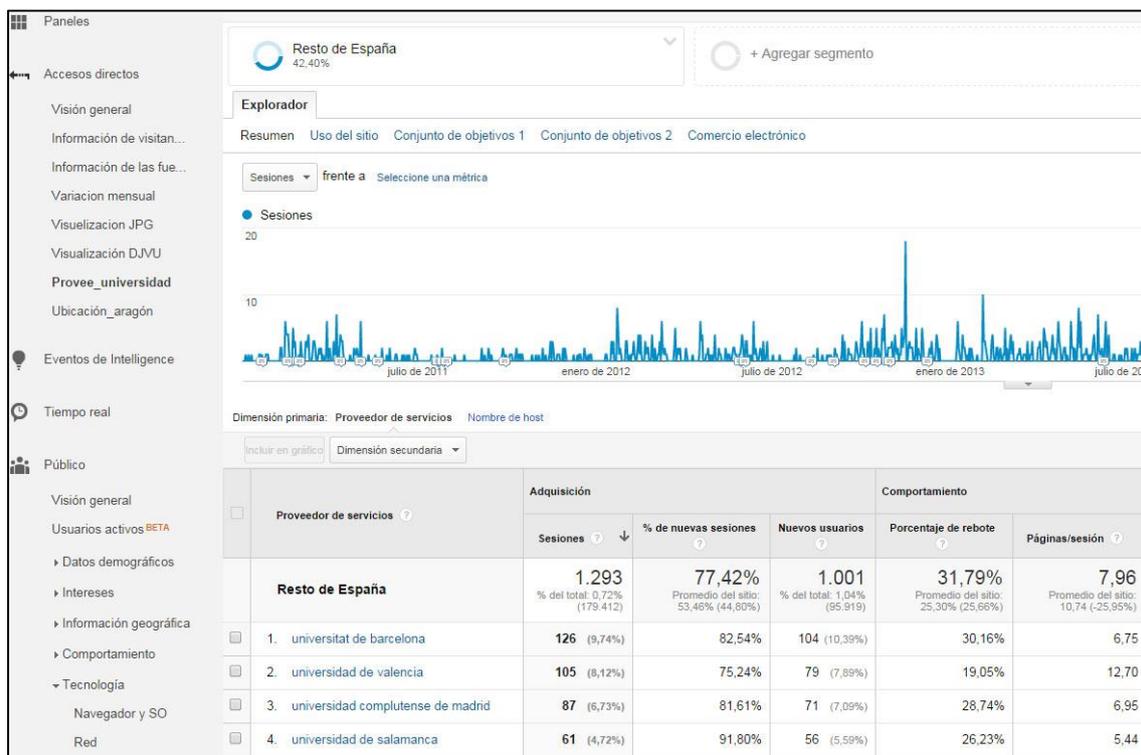


Figura 1: Combinación de informe, segmento y filtro de datos

En el ejemplo propuesto, se ha seleccionado como informe “Proveedor de Servicios de Internet”. Sobre este informe se ha efectuado la visualización exclusiva del segmento “Usuarios de España excluyendo Aragón” y se ha filtrado los resultados para que solo muestre las visitas cuyo proveedor de internet contiene los términos “univ* |escuela|school|institut”.

Como resultado, obtendremos los datos de usuarios que han realizado una visita al servicio cultural desde un dispositivo conectado a una red provista por una universidad, aislando geográficamente los datos de aquellos usuarios que lo hagan desde fuera de Aragón.

Este ejemplo ilustra eficazmente las posibilidades de configuración de la herramienta de análisis y el tipo de datos que pueden filtrarse y depurarse para su análisis.

Selección de indicadores relevantes

Quedan patentes pues las muy elevadas posibles combinaciones de elementos de configuración de Google Analytics a la hora de obtener y mostrar los datos que almacena a lo largo del tiempo. Será necesario por tanto una selección previa de *indicadores relevantes* para el estudio de usos y usuarios de servicios culturales en el ámbito online.

La propia naturaleza del servicio sobre el que se realice un análisis web va a condicionar cuáles son los indicadores más relevantes, cuáles son aquellos que van a mostrar realmente el funcionamiento de la web y van a servir como base de su estudio.

De esta forma, no tiene las mismas necesidades de información una página comercial cuyo fin es la conversión de visitas a su servicio en compras a través de internet, que un Archivo Online o cualquier otro servicio de información cultural de acceso remoto (como tampoco las comparten con portales de difusión de noticias, páginas personales, etc). Una web comercial necesitará conocer el retorno de su inversión en publicidad, los mecanismos de atracción de usuarios a las diferentes funcionalidades de compra de ítems, etc.

En el estudio que a continuación se les ofrece, los indicadores seleccionados combinan elementos propios de los usuarios (procedencia geográfica, identificación de usuarios internos y externos, frecuencia de visita, interacción con la web) y elementos propios de las funcionalidades propias del servicio cultural (como uso de herramientas de búsqueda, impacto de las acciones definidas de difusión cultural, uso de subportales o consulta de URL).

Estudio de usuarios: DARA y SIPCA

Propuesta de estudio

A continuación se procede a presentar los datos de uso y usuarios de los servicios de información cultural online DARA Documentos y Archivos de Aragón y SIPCA Servicio de Información del Patrimonio Cultural en Internet.

El rango cronológico del estudio comprende entre el 1 de enero de 2011 y el 30 de junio de 2014. Si bien se almacenan datos de periodos anteriores, se considera que en las fechas propuestas ambos servicios ya gozan de un funcionamiento “normalizado”.

Se va a proceder de lo general a lo particular. En primer lugar se ofrecerán datos generales de ambos servicios y una visualización cronológica de los mismos, comparando algunos de sus indicadores clave.

En una segunda fase se aislarán segmentos poblacionales diferenciados (usuarios internos/externos y grupos de usuarios por procedencia geográfica) atendiéndose especialmente a sus especificidades de comportamiento y uso del servicio.

DARA: Documentos y Archivos de Aragón

En la figura 2 pueden observarse los datos generales para los 42 meses objeto del estudio:



Figura 2: Datos generales de indicadores relevantes de DARA Documentos y archivos de Aragón

Con un total de casi 180.000 sesiones (visitas) a la página y casi dos millones de páginas individuales visualizadas, obtenemos una media de 10,75 páginas vistas por sesión. Aproximadamente un cuarto de estas visitas se consideran un “rebote” (usuario que entra y sale del servicio rápidamente y solo visualiza una única página).

La duración media de las sesiones es de nueve minutos y medio, y el porcentaje de nuevas sesiones (usuarios que nunca antes habían visitado el servicio) es de poco más del 50%.

En general, estos datos dan testimonio del uso relativamente intenso que se hace del servicio, que se pondrán en perspectiva al compararse con los datos obtenidos por SIPCA.

En la figura 3 puede observarse la evolución trimestral de visitas obtenidas por este portal de archivos online:

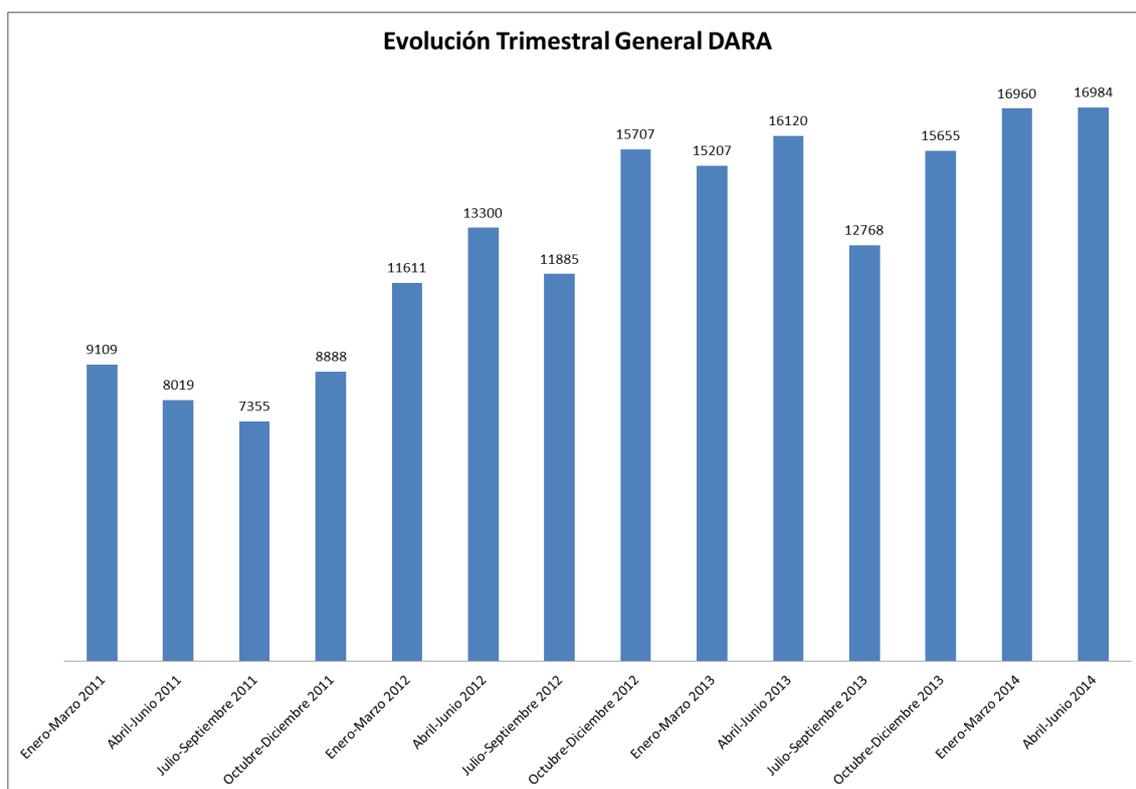


Figura 3: Evolución trimestral de las visitas a DARA

Como puede observarse, existe una tendencia de crecimiento en el uso del servicio, en la que pueden observarse algunas fluctuaciones estacionales, siendo los meses de Julio a Septiembre los que reflejan un menor número de sesiones.

Los dos primeros trimestres del año 2014 muestran los números absolutos más altos respecto a número de visitas, lo que indica el buen estado de salud del servicio. Con casi 17.000 visitas trimestrales, los datos revelan que las sesiones prácticamente se han

duplicado respecto a las obtenidas en los dos primeros trimestres estudiados (que cuentan 9.109 y 8.019 visitas respectivamente).

En la Figura 4 se desglosan los datos, distribuidos semestralmente, de algunos indicadores cualitativos relevantes para el servicio.

Semestre	Sesiones	Usuarios	Páginas	Pág/ses	Duración (min)	Rebote %	Nuevas %
2011/1	17.128	8.467	253.239	14,79	10:31	23,33	47,86
2011/2	16.243	8.108	249.276	15,35	11:55	19,44	46,96
2012/1	24.911	13.220	295.590	11,87	09:42	23,86	50,50
2012/2	27.592	14.891	283.973	10,29	09:49	25,78	51,05
2013/1	31.327	17.857	295.460	9,43	08:40	27,59	53,98
2013/2	28.423	16.863	259.521	9,13	08:55	26,33	56,99
2014/1	33.914	20.934	292.641	8,63	08:27	26,75	58,75

Figura 4: Evolución semestral de indicadores cualitativos

Además del ya previsto crecimiento semestre a semestre relativo al número de sesiones que registra el servicio, puede observarse un incremento parejo en el número de usuarios únicos (navegadores únicos que acceden al servicio, identificados mediante el uso de cookies).

El número de páginas individuales visualizadas alcanza su valor máximo en el primer semestre del año 2012, manteniéndose en números similares los años siguientes, con las variaciones estacionales previstas.

Los indicadores de páginas por sesión y duración de la visita (que indican el grado de interacción de los usuarios con el servicio) son sin embargo decrecientes, desde los mejores resultados obtenidos en el año 2011, con alrededor de 15 páginas por sesión y duraciones superiores a los 10 minutos, los datos de estos indicadores decaen consistentemente en los años 2012 y 2013, hasta las 8,63 páginas por sesión con duraciones de menos de 9 minutos. El porcentaje de rebote, aunque en menor medida, también presenta tendencia al crecimiento.

El porcentaje de nuevas visitas, sin embargo, se incrementa a lo largo de estos años en aproximadamente un 10%, desde el 47,86% de la primera mitad del año 2011 hasta el 58,75% del año 2014.

Puede concluirse que la evolución inversa de estos indicadores está relacionada: el servicio de archivo online, por sus acciones concretas de difusión y por su presencia cada

vez más consolidada como servicio cultural online tiene como consecuencia un mayor porcentaje de nuevos usuarios, es decir, usuarios que nunca antes habían hecho uso del servicio. Esta mayor capacidad de llegar a nuevos usuarios trae consigo un mayor ratio de usuarios que, pese a que visitan el servicio al dárseles a conocer, no hacen uso de él, lo que tiene un impacto negativo en el resto de indicadores de interacción con el mismo. Sin embargo, la tarea de difusión ya ha sido realizada.

SIPCA Servicio de Información del Patrimonio Cultural Aragonés

En la figura 5 se muestran los datos generales de uso del servicio en el rango cronológico propuesto:



Figura 5: Datos generales de indicadores relevantes de SIPCA Servicio de Información del Patrimonio Cultural Aragonés

Con casi 190.000 sesiones el uso de este servicio es mayor que el obtenido en el mismo periodo por DARA. El número de páginas vistas –y por tanto el número de páginas por sesión- es considerablemente menor, con un total de 650.000 y 3,43 páginas por sesión. La duración media de cada visita es en este caso de unos tres minutos y medio

El porcentaje de rebote es también considerablemente mayor, con un 60% de los usuarios entrando y saliendo del servicio desde la misma URL. El porcentaje de nuevas visitas, que usado de forma comparativa puede interpretarse como la capacidad de captación de nuevos usuarios, es sustancialmente mayor en el caso de SIPCA. Hasta casi un 64% de las sesiones que acceden al servicio no lo habían hecho previamente.

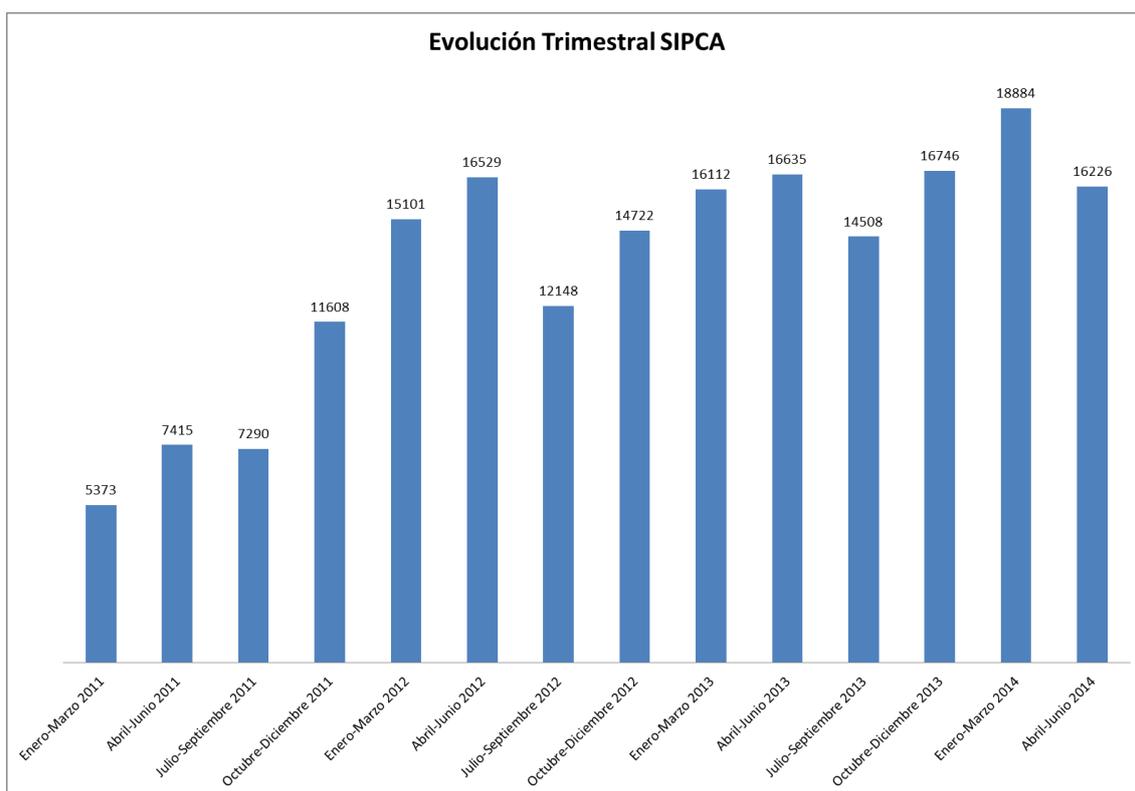


Figura 6: Evolución trimestral de las visitas a SIPCA

En el caso de SIPCA el crecimiento del número de visitas es también muy acusado a lo largo del tiempo, observándose diferencias incluso mayores a las indicadas por DARA entre los primeros datos de la serie (2011) y los últimos (2014). Al igual que en el portal de archivos online, los datos del primer semestre de 2014 son los que muestran mayores valores objetivos respecto al número de visitas, que prácticamente triplican a los obtenidos en el mismo periodo de 2011.

Semestre	Sesiones	Usuarios	Páginas	Pág/ses	Duración (min)	Rebote %	Nuevas %
2011/1	12.788	6.271	82.590	6,46	06:19	46,03	46,59
2011/2	18.898	12.213	68.833	3,64	03:53	56,17	62,41
2012/1	31.630	20.113	110.462	3,49	03:59	53,35	61,14
2012/2	26.870	17.444	95.221	3,54	03:33	58,92	61,89
2013/1	32.747	22.320	94.132	2,87	03:02	64,22	65,67
2013/2	31.254	22.333	87.746	2,81	03:03	65,45	68,63
2014/1	35.110	25.111	111.045	3,16	03:04	66,06	68,57

Figura 7: Evolución semestral de indicadores cualitativos

La evolución de los indicadores propuestos es similar en su tendencia a la observada en el caso de DARA, presentando diferencias aún más acusadas. Frente al drástico incremento en número de sesiones (se triplica entre los datos de 2011 y 2014) y de usuarios únicos (se multiplica por 5), el número de páginas vistas alcanza, prácticamente, su valor máximo en el primer semestre de 2012 para mantenerse en los años sucesivos.

Los indicadores de interacción con el servicio ven sus cifras disminuidas de forma importante. El número de páginas por sesión alcanza su mínimo a finales de 2013 descendiendo desde las 6,46 de la primera mitad de 2011 hasta 2,81 en su valor mínimo. La duración media de la visita se acorta desde los 6,19 minutos de los primeros datos de la serie hasta los 3 de la última medición. El porcentaje de rebote aumenta en 20 puntos desde el 46% de 2011 al 66% de 2014.

El porcentaje de nuevas visitas es un indicador, de nuevo, clave para interpretar los datos de interacción y uso del servicio cultural. Como en el caso de DARA, la tendencia deriva hacia una mayor capacidad de captación de nuevos usuarios. En este caso el crecimiento de este indicador tiene un importante salto adelante en la segunda mitad de 2011 en que alcanza un 62% de las sesiones, para llegar hasta un 68,57% en el año 2014.

Para interpretar estos datos, en primer lugar deben tenerse en cuenta ciertas circunstancias técnicas. A mediados de 2011 se operaron cambios en SIPCA a nivel informático, que se tradujeron en la utilización de “URL amigables” para los resultados de las búsquedas llevadas a cabo por los usuarios en los buscadores provistos por el portal. Este hecho llevó a algo vital para las posibilidades de crecimiento y difusión de un medio online: la indexación de las páginas de resultados de búsqueda interna en buscadores orgánicos externos, mayoritariamente en Google. Una búsqueda en Google podía, a raíz de estos cambios, ofrecer como resultado las páginas desglosadas de SIPCA, lo que provocó el incremento radical en número de usuarios observado en 2011 y consecuentemente del porcentaje de nuevos usuarios.

En segundo lugar, la mayor capacidad de captación de nuevos usuarios lleva asociada, como vimos en el caso de DARA, un decrecimiento paulatino del resto de indicadores cualitativos de uso del servicio.

Además, el caso de SIPCA tiene especificidades propias, ya que el Servicio de Información del Patrimonio Cultural Aragonés sirve como portal que aloja otros servicios y portales culturales, como es el caso de DARA. Es decir, que SIPCA funciona, en muchas ocasiones, como una “lanzadera” o catalizador de visitas y usuarios desde este servicio hacia otros.

Por esta razón los indicadores de interacción y uso del servicio son notablemente menos abultados en el caso de SIPCA respecto al de DARA. Por esta misma razón, la capacidad de captación de nuevos usuarios desplegada por SIPCA se revela vital, ya que muchos de estos nuevos usuarios van a solucionar sus necesidades de información no en el seno

de SIPCA, sino en el de DARA y otros servicios culturales online a los que puede accederse a través de SIPCA.

Frecuencia de visita e interacción con el servicio

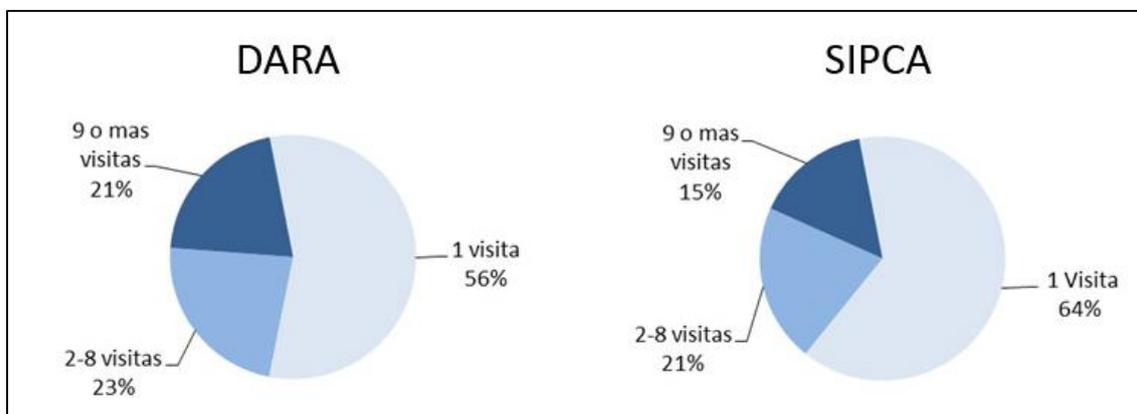


Figura 8: Frecuencia de retorno de los usuarios a DARA y SIPCA

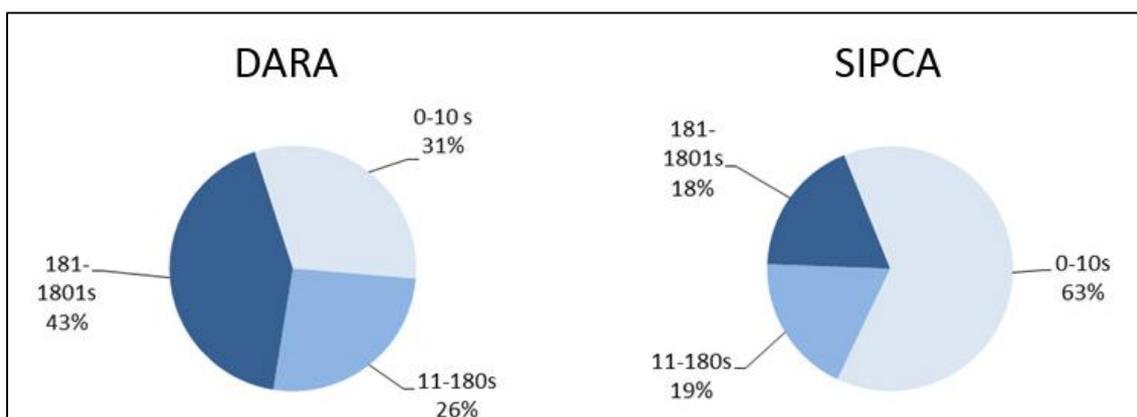


Figura 9: Interacción de los usuarios con el servicio (tiempo de visita)

Las figuras 8 y 9 muestran, enfrentando los resultados de DARA y SIPCA, los datos de frecuencia de retorno (el número de ocasiones en las que un visitante identificado como una cookie individual accede al servicio) y la Interacción de los usuarios con el servicio, es decir el tiempo invertido por cada una de las sesiones que hacen uso del servicio.

Aun teniendo en cuenta la incertidumbre provocada por el borrado de cookies, obtenemos resultados de frecuencia de retorno que alcanzan el 44% en el caso de DARA –con un 21% de las sesiones habiendo accedido al servicio al menos en 9 o más ocasiones- y del 36% en el caso de SIPCA, con un 15% de las visitas indicando una alta tasa de retorno.

Si se observa la interacción con el portal, en primer lugar llama la atención la diferencia notable en cuanto a comportamiento entre los usuarios entre de estos servicios de información. Los datos reflejan que hasta un 43% de los usuarios que visitan DARA

invierten en el sitio de más de 3 minutos (frente al 18% de SIPCA), lo que es clara muestra del buen funcionamiento del portal.

Podemos obtener conclusiones adecuadas si se interpretan estos datos a la luz de la naturaleza y contenidos de los dos servicios comparados. SIPCA es un portal que, además de albergar contenidos culturales (fotografía, descripciones, etc.) en sí mismo, sirve a su vez como puerta de acceso a otros servicios culturales de Gobierno de Aragón. DARA, por su naturaleza de portal de archivos online, alberga una gran cantidad de fondos documentales descritos y también digitalizados, y un mucho mayor número absoluto de ítems.

Los indicadores de uso de SIPCA se ven afectados por esta circunstancia: un gran número de las visitas que recibe se redirige a otros servicios, disminuyendo los valores de los indicadores de interacción y uso, como ya hemos visto. Su capacidad de captación de nuevos usuarios, potenciada por las buenas condiciones de indexación y posicionamiento de que disfruta, van a tener un impacto directo en el número de usuarios que reciben estos otros servicios, como es el caso de DARA. Es posible interpretar, a su vez, que habrá usuarios a los que se dé a conocer así la existencia de estos otros portales, por lo que la tasa de frecuencia de visita se vería disminuida.

DARA Documentos y Archivos de Aragón, como portal de archivos aragoneses, ofrece a los posibles usuarios un número muy elevado de descripciones y documentos digitalizados, fotografías, planos y demás. Los indicadores de interacción y uso del servicio serán por tanto más abultados, de acuerdo con la naturaleza de estos contenidos. El buen funcionamiento del portal conlleva, por tanto, un mayor número de usuarios implicados y que hacen un uso prolongado e intensivo del servicio.

Método de adquisición de usuarios

Tanto DARA como SIPCA son servicios culturales en el medio online. Como tales, es conveniente analizar cuál es el método de captación de los usuarios que los utilizan. Es decir, cómo se distribuye, de entre las opciones disponibles, la procedencia de los usuarios atendiendo al tipo de tráfico que representan.

Sin un depurado posterior de estos datos (que serán segmentados y filtrados para realizar un estudio en mayor profundidad), los métodos de captación de usuarios en internet son los siguientes:

- *Tráfico directo*: Usuarios que acceden a una web escribiendo en su navegador la URL que desean visitar o mediante su inclusión entre los “favoritos” de sus navegadores. En esta categoría también se incluyen los usuarios que acceden a través de un “click” en un enlace existente en un archivo PDF o DOC, entre otros.

- *Tráfico de referencia:* Usuarios que acceden a una web mediante un enlace o hipervínculo existente en otra web, sea cual sea.
- *Tráfico de búsqueda:* Usuarios que acceden a una web desde un buscador orgánico (mayoritariamente, en nuestro caso, Google) tras haber realizado una búsqueda en el mismo.

En la Figura 10 puede observarse la distribución de este tipo de tráfico para los servicios culturales estudiados:

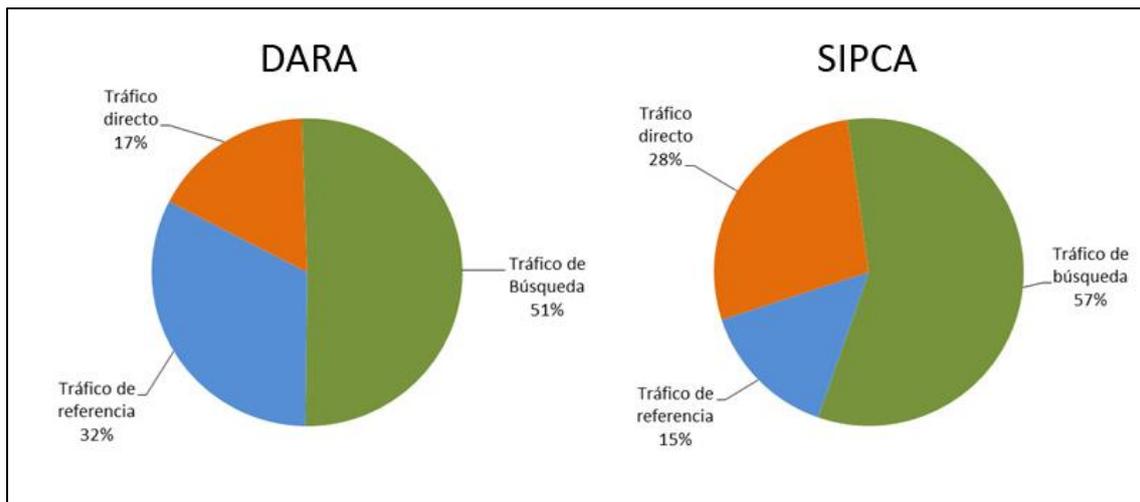


Figura 10: Distribución del tipo de tráfico para DARA y SIPCA

El primer resultado que llama la atención es el muy notable peso específico que, en ambos casos, tiene el tráfico de búsqueda. En ambos servicios más del 50% de los usuarios acceden a través de búsquedas orgánicas, principalmente a través de Google. En el caso de SIPCA un 57% de los usuarios accede al servicio a través de buscadores externos.

DARA obtiene un amplio porcentaje de sus usuarios a través de enlaces procedentes desde otras páginas y servicios afines, con una incidencia mayor de este tipo de tráfico que la observada para el caso de SIPCA. Debe tenerse en cuenta que un porcentaje muy elevado (Como veremos más adelante casi un 20% del total de sus usuarios) proceden de enlaces incluidos en el portal SIPCA.

En fases posteriores de este análisis se procederá a un estudio desglosado de las diferentes fuentes y segmentos poblacionales, diferenciados en función de la fuente de procedencia y tipo de tráfico.

Uso de la herramienta de búsqueda

Tanto DARA como SIPCA son portales que disponen, como parte integral del servicio que ofrecen a los ciudadanos, buscadores internos que propician el acceso de los usuarios a la información que necesitan. Google analytics permite realizar un seguimiento del funcionamiento de estas herramientas. En la figura 11 se muestran los resultados obtenidos:



Figura 11: Uso de la herramienta de búsqueda en DARA y SIPCA

Debe tenerse en cuenta que los datos mostrados para el portal SIPCA comprenden tan solo los meses transcurridos entre marzo y junio de 2014, fechas en las que el seguimiento de este indicador fue activado.

En el caso de DARA obtenemos un total de 93.384 sesiones con búsqueda, es decir, más del 50% de los usuarios realizaron una o más búsquedas en las herramientas provistas por el portal. En total, se realizaron casi 250.000 búsquedas únicas.

Si atendemos a los términos y cadenas de búsqueda realizadas por los usuarios, se ha observado que hasta 30.000 de ellas contienen el término “censo” o derivados, lo que refleja uno de los principales conjuntos de interés y atracción de visitantes. El resto de las búsquedas se distribuye de forma homogénea entre lugares, personas, instituciones o documentos concretos. Un análisis más profundo de este conjunto de datos (disponible y almacenado por Google Analytics) puede ser utilizado para comprender en mayor profundidad las necesidades de información de los usuarios y cómo interactúan con el servicio, lo que puede traducirse en datos fiables para la selección de descriptores y metadatos orientados a las necesidades del usuario.

En el caso de SIPCA, como se ha indicado, los datos obtenidos son muy fragmentarios, ya que sólo se han venido almacenando desde marzo de 2014. En los tres meses estudiados, se realizaron casi dos mil búsquedas únicas.

Los términos utilizados por los usuarios en los buscadores internos provistos por SIPCA son prácticamente en su totalidad toponímicos, concretamente se realizan búsquedas de núcleos poblacionales.

DARA: Uso de subportales

DARA, como Portal de archivos de Aragón, se articula de forma modular, y en su estructura existen diferentes subportales correspondientes a los diferentes Archivos de titularidad tanto pública como privada que lo integran.

Entre otros, los subportales más utilizados por los usuarios son los pertenecientes a los tres Archivos Históricos Provinciales aragoneses, a Archivos Fotográficos y a otros Archivos Municipales, Provinciales y privados.

En la figura 12 se muestra la distribución de los usuarios que acceden a DARA a través de subportales o a través del portal general, atendiendo específicamente a los Archivos Históricos Provinciales:

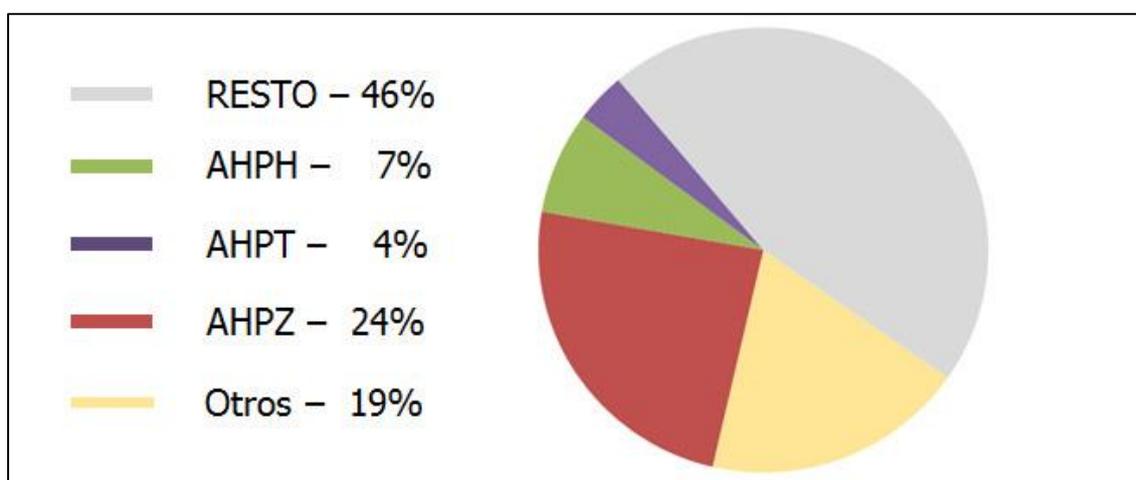


Figura 12: Distribución de uso de subportales según página de inicio de los usuarios de DARA

El uso de los diferentes subportales suma en total un 54% de las visitas que acceden a DARA, lo que es buena muestra de su potencial específico de difusión y crecimiento.

El subportal del Archivo Histórico Provincial de Zaragoza presenta los datos más abultados con un 24% del total de los usuarios del servicio. Le siguen el Archivo Histórico Provincial de Huesca (7%) y el de Teruel (4%). El resto de los subportales, en su conjunto, suman un 19% del total de las visitas a DARA.

Si atendemos a la evolución cronológica de estos segmentos, puede observarse que su peso específico en el uso del servicio se reduce ligeramente a lo largo del tiempo. Por ejemplo, en los primeros seis meses del año 2014 el subportal del Archivo Histórico Provincial de Zaragoza sumó un 23,78% de las visitas totales al portal. Sabemos que las cifras globales de usuarios de DARA crecieron notablemente en el mismo periodo de tiempo, con un crecimiento comparativamente menor del uso de los subportales.

Estudio de segmentos poblacionales específicos: usuarios externos/internos y usuarios según procedencia geográfica

Segmentación

Como se indicó al principio del presente informe, será necesario el uso de las opciones de configuración avanzadas facilitadas por la herramienta de analítica web para la obtención de datos específicos sobre segmentos poblacionales concretos, basados en variables definidas.

Mediante la utilización de este tipo de filtros se han podido captar, de forma diferenciada, los datos e indicadores de comportamiento y uso del servicio de usuarios **internos y externos**, por un lado, y de grupos de población según **procedencia geográfica**, por otro.

Para la identificación de los usuarios externos e internos se ha utilizado la variable “proveedor de servicios de internet”. En el caso de DARA y SIPCA, todos los usuarios internos que acceden a los mismos mostrarán como ISP “Diputación General de Aragón”. De esta forma, es posible separar los usuarios que acceden al servicio desde la red de la institución de aquellos que acceden desde cualquier otro ámbito.

Como otros elementos de la analítica web, este procedimiento está sujeto a varias limitaciones y carencias: es posible que usuarios internos accedan al servicio desde redes externas, y de la misma manera habrá usuarios externos que lo hagan desde dispositivos conectados a la red institucional. Estas limitaciones distorsionarán los datos obtenidos, que sin embargo pueden ser utilizados como indicadores cualitativos del uso del servicio.

Según la procedencia geográfica, se han configurado cuatro segmentos poblacionales, que en este caso si tienen una fiabilidad contrastada:

- Usuarios procedentes de Aragón.
- Usuarios procedentes del resto de España.
- Usuarios procedentes de Europa (excluyendo España).
- Usuarios procedentes de América.

Para el estudio se ha prescindido de los datos de usuarios procedentes de Oceanía, África y Asia, ya que en su conjunto ofrecen cifras limitadas (muy por debajo del 1% del total de los usuarios).

Usuarios internos/externos de DARA

En la figura 13 se muestra la distribución porcentual total de usuarios internos y usuarios externos que utilizaron el portal DARA entre 2011 y 2014, y en la figura 14 la evolución año a año del porcentaje de ambos segmentos poblacionales.

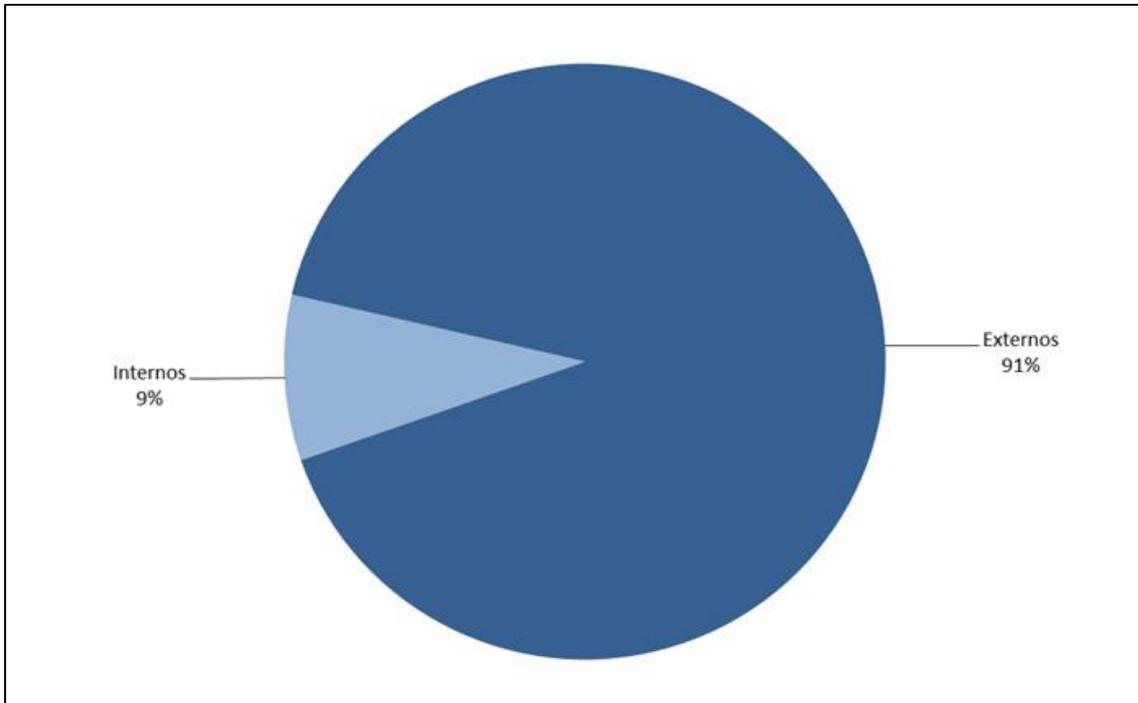


Figura 13: Distribución porcentual de usuarios internos/externos de DARA

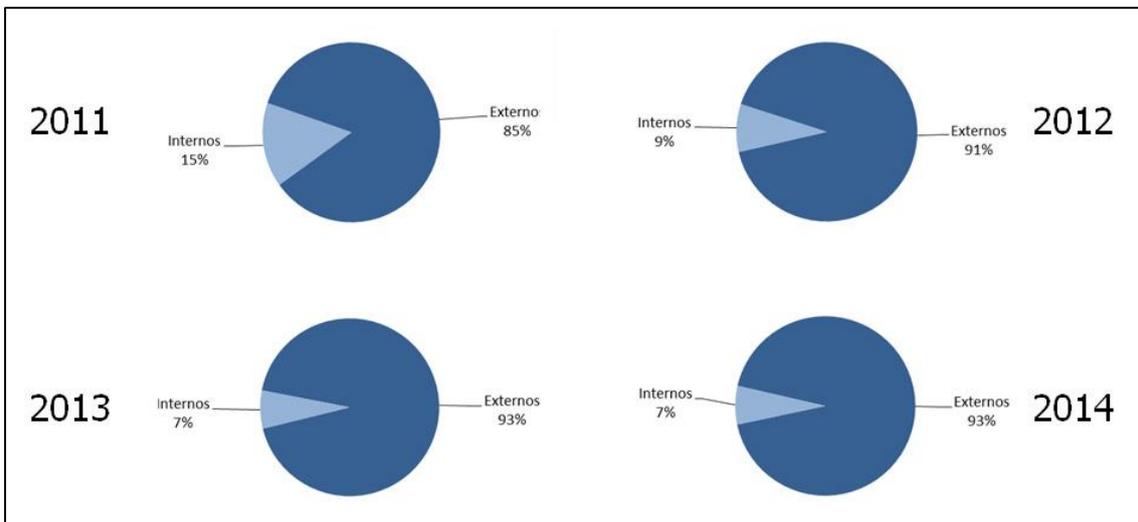


Figura 14: Evolución de la distribución anual porcentual de usuarios internos/externos en DARA

Aproximadamente nueve de cada diez usuarios, según los parámetros utilizados, son usuarios externos. Esta división porcentual, que muestra el elevado peso específico de los usuarios externos en el funcionamiento de DARA, se ha profundizado a lo largo de

los años estudiados. En 2011 el porcentaje de usuarios externos era de un 85%, frente al 93% captado en 2013 y 2014.

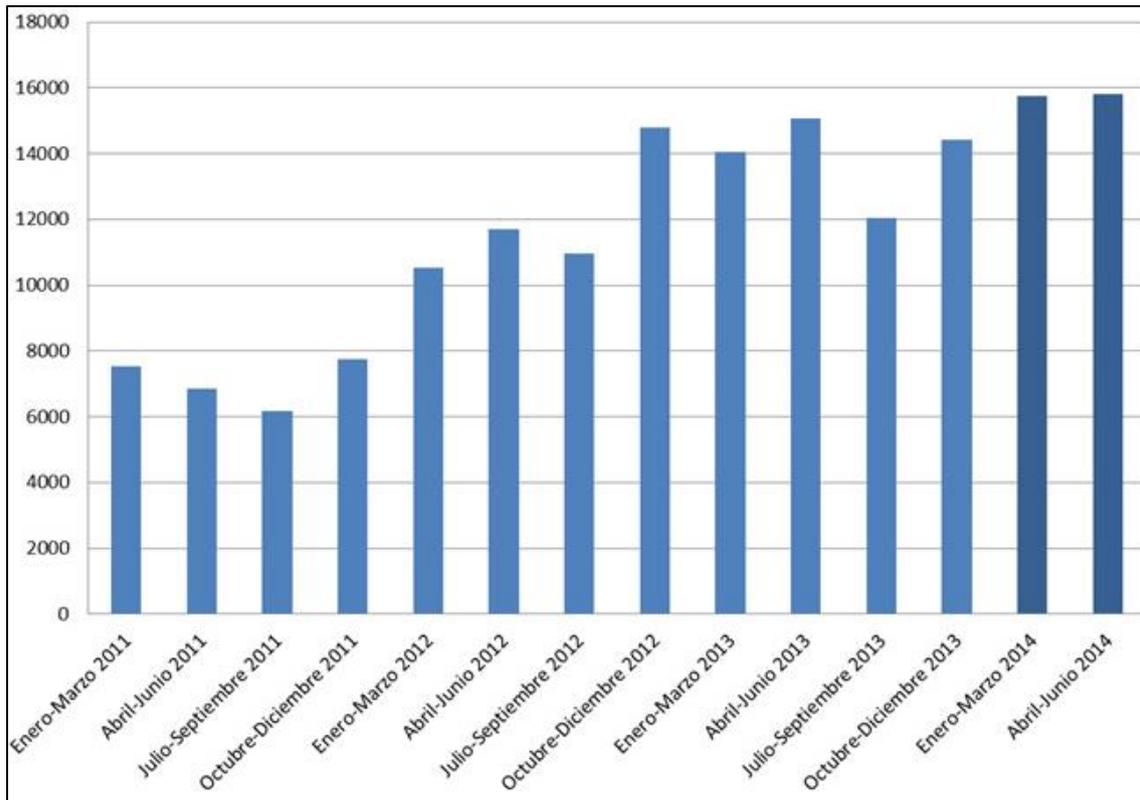


Figura 15: Evolución trimestral de los usuarios externos de DARA

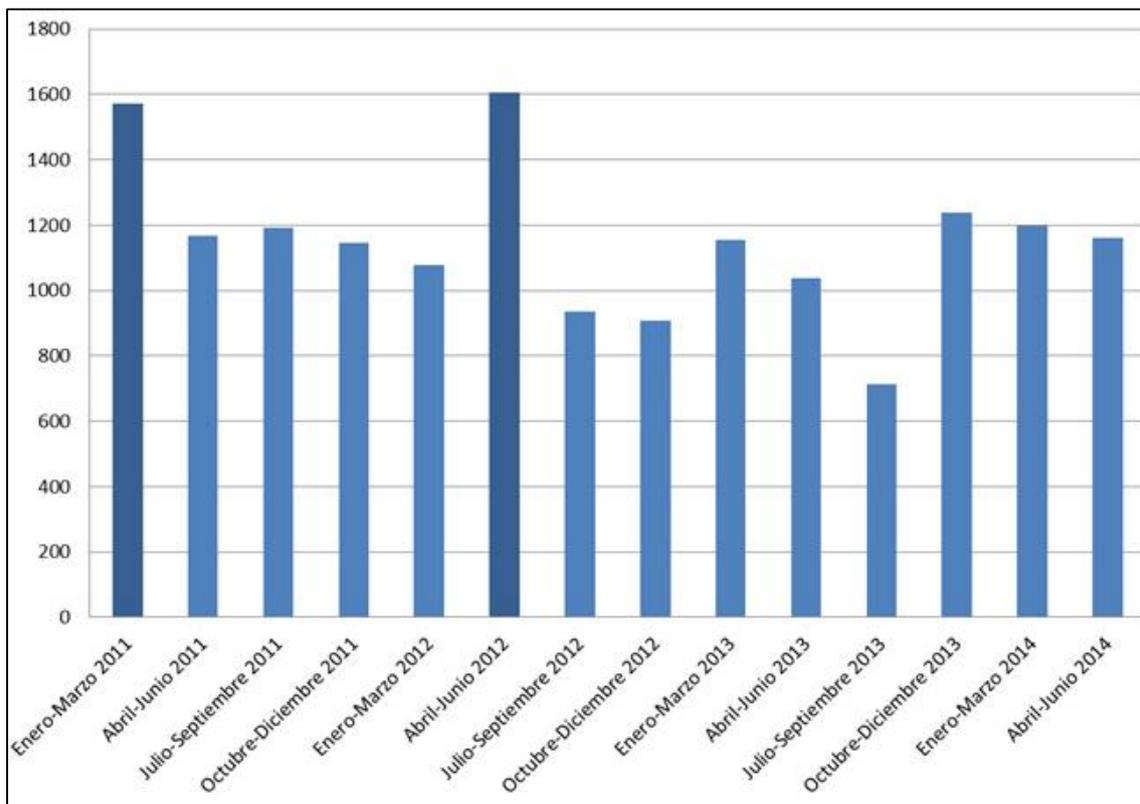


Figura 16: Evolución trimestral de los usuarios internos de DARA

Como puede observarse en las figuras 15 y 16, correspondientes a la evolución trimestral total de los usuarios internos y usuarios externos, ambos segmentos poblacionales tienen tendencias de captación de visitantes muy diferentes.

Cada vez se obtiene un mayor número de visitas de usuarios externos, pudiendo observarse un crecimiento interanual constante a lo largo del periodo. Los datos obtenidos en 2014 (resaltados en tono más oscuro) muestran los mayores valores obtenidos en la serie. Nada indica que la tendencia vaya a cambiar en el segundo semestre de este año.

Respecto a los usuarios internos, no se observa crecimiento, sino ciertas oscilaciones, obteniendo los valores más altos en periodos concretos, provocadas por necesidades del servicio. Existe también una acusada oscilación estacional coincidente con los meses estivales de julio a septiembre.

Respecto al modo de adquisición de cada uno de estos segmentos poblacionales (figura 17), puede verse que la distribución que presenta el segmento de usuarios externos es muy similar al observado en los datos de DARA general –no en vano se trata de más del 90% de los usuarios del servicio- con un porcentaje mayor de usuarios (53%) que acceden a través de buscadores orgánicos. Los usuarios internos, por su parte, lo hacen sobre todo a través de tráfico directo (24%) y tráfico de referencia (55%).

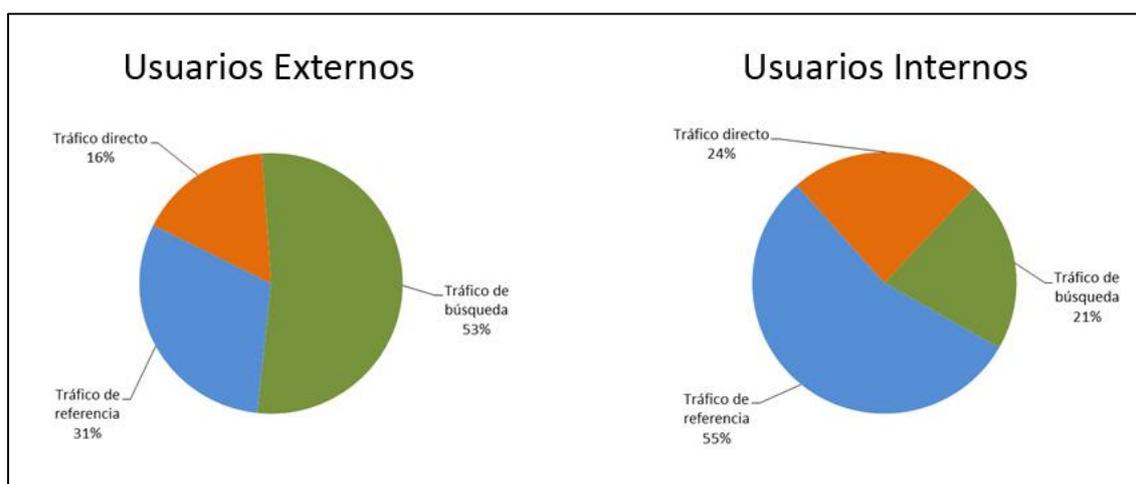


Figura 17 Modo de captación de usuarios Internos/externos de DARA

Usuarios internos/externos de SIPCA

En las figuras 18 y 19 se muestra la distribución porcentual de ambos segmentos poblacionales y su evolución anual:

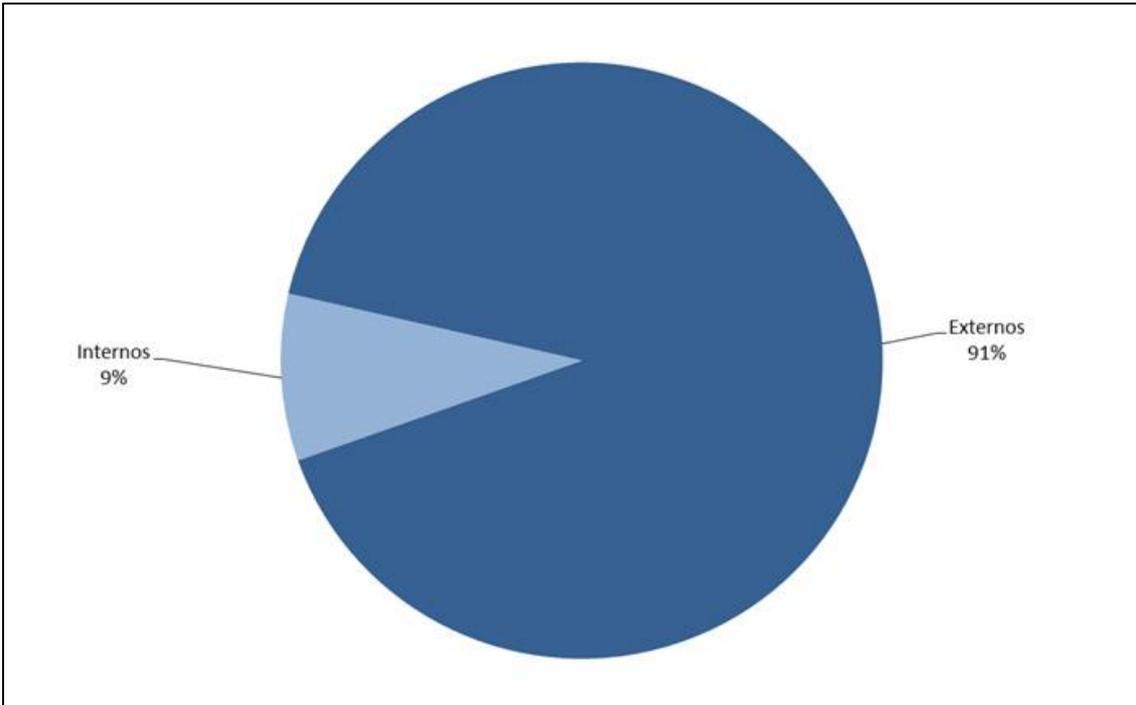


Figura 18: Distribución porcentual de usuarios internos/externos de SIPCA

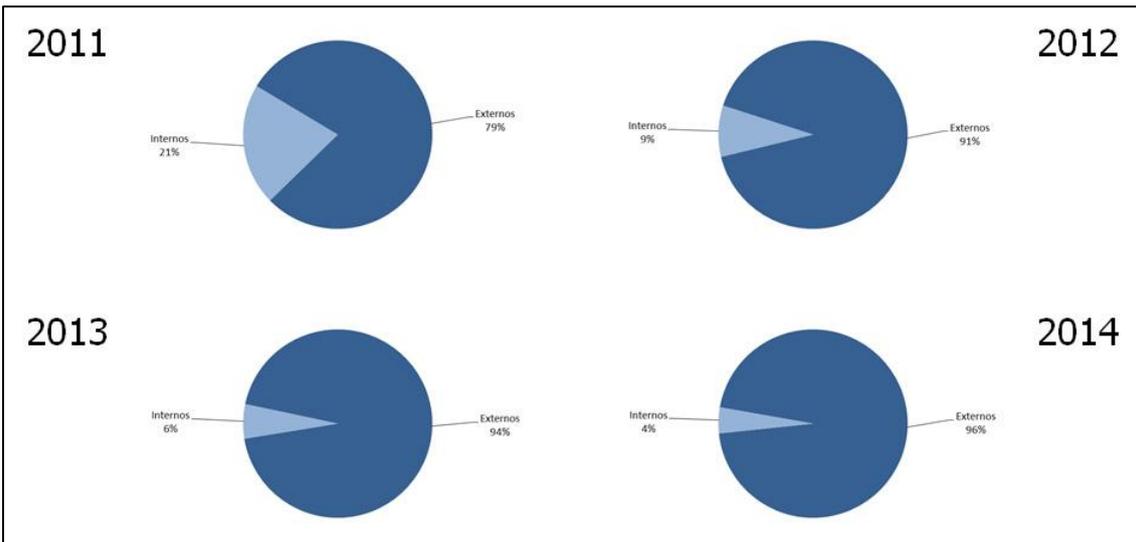


Figura 19: Distribución porcentual anual de usuarios internos/externos SIPCA

La distribución de usuarios internos y externos es idéntica en DARA y en SIPCA: 9% de usuarios internos frente a 91% de usuarios externos. En el caso del Servicio de Información del Patrimonio Cultural Aragonés, esta desigual distribución ha ido aumentando en ratios aún mayores que los observados en DARA. Así, en 2011 un 21% de los usuarios del servicio eran internos, mientras que en 2013 lo era un 6% y en 2014 un 4%.

En la figura 20 se muestra la distribución según método de captación:

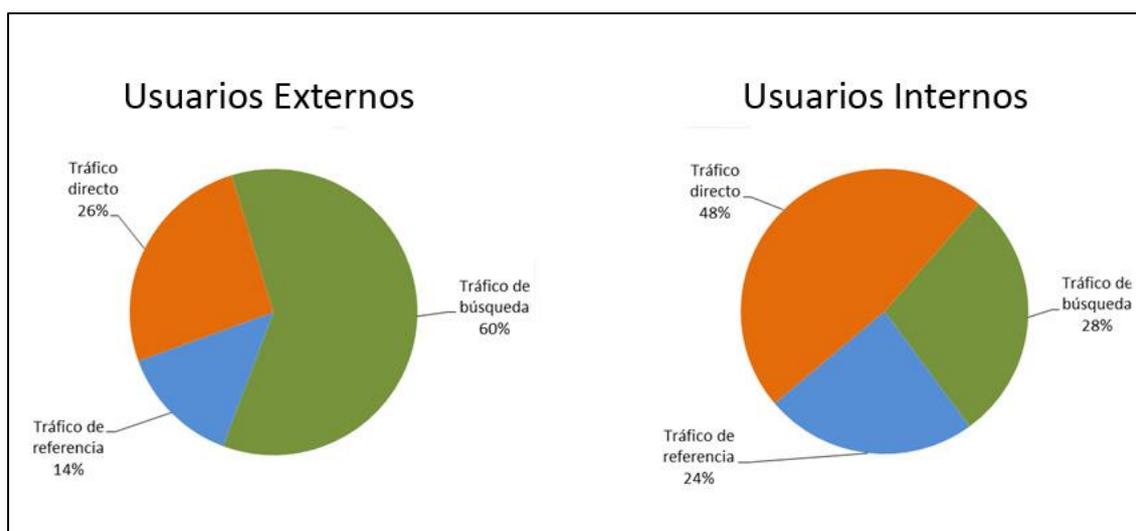


Figura 20: Modo de captación de usuarios externos/internos de SIPCA

La incidencia del uso de buscadores es más intensa en el caso de los usuarios externos de SIPCA, donde alcanza un 60%. El tráfico directo tiene una especial relevancia en ambos segmentos poblacionales, especialmente en el caso de los usuarios internos, en el que casi un 50% utiliza este sistema para acceder al servicio.

Usuarios internos/externos: estado de los servicios

Ambos servicios culturales comparten, para el conjunto general de los datos analizados, los mismos porcentajes de usuarios externos y usuarios internos: 91%-9%. La tendencia de ambos servicios hacia una mayor expansión e incidencia de usuarios y nuevos usuarios –recordemos que en ambos casos el año 2014 registra los mejores resultados de la serie- tienen como necesaria consecuencia el protagonismo, cada vez más acusado, de los usuarios externos. Protagonismo forzoso además dado el medio en el que ambos existen.

Segmentos poblacionales según procedencia geográfica⁶

Distribución general

La figura 21 muestra la distribución porcentual de los usuarios de DARA según procedencia geográfica:

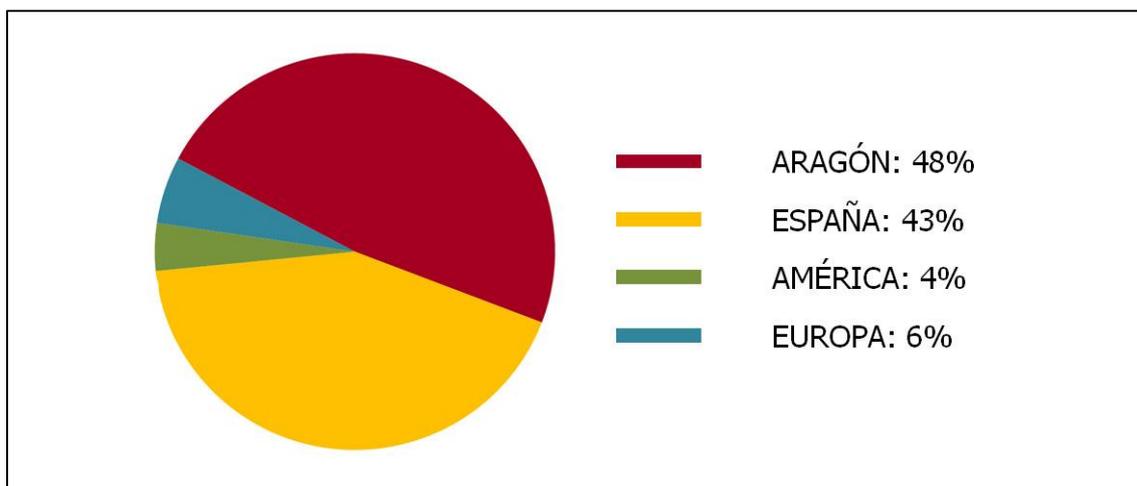


Figura 21: Porcentaje de usuarios según procedencia geográfica

El 90,5% de los usuarios acceden desde España, con una mayoría del 48% procedente de Aragón. Casi un 43% de los usuarios proviene del resto de Comunidades Autónomas del país, con predominio de Madrid y Cataluña.

Los usuarios de fuera de nuestras fronteras casi alcanzan el 10% del total de visitas, con un ligero predominio de los usuarios europeos (casi un 6%) sobre los usuarios americanos (un 4%).

⁶ A partir de este epígrafe el estudio se centra de forma exclusiva en el estudio de los datos de DARA: Documentos y Archivos de Aragón.

Evolución interanual de los segmentos según procedencia geográfica:

ARAGÓN

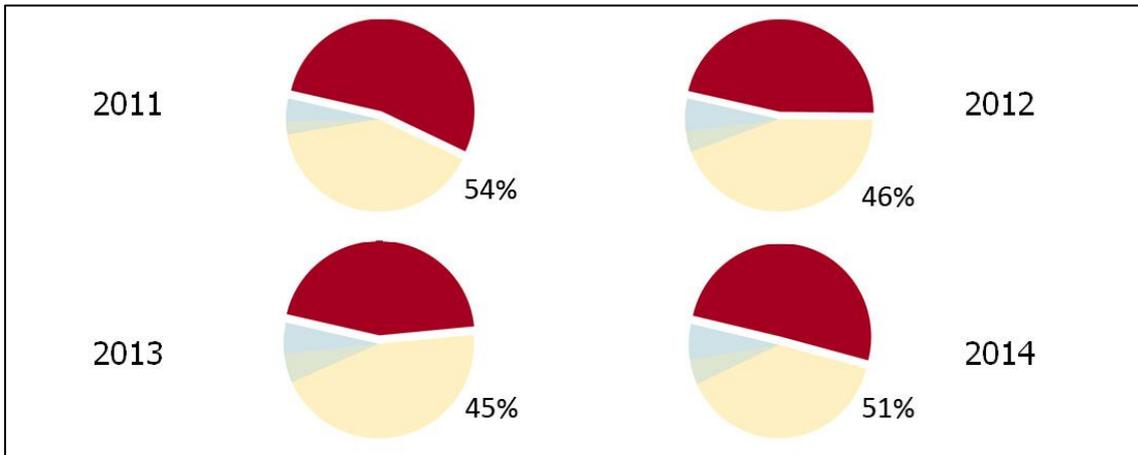


Figura 22: Porcentajes anuales de usuarios aragoneses

RESTO DE ESPAÑA

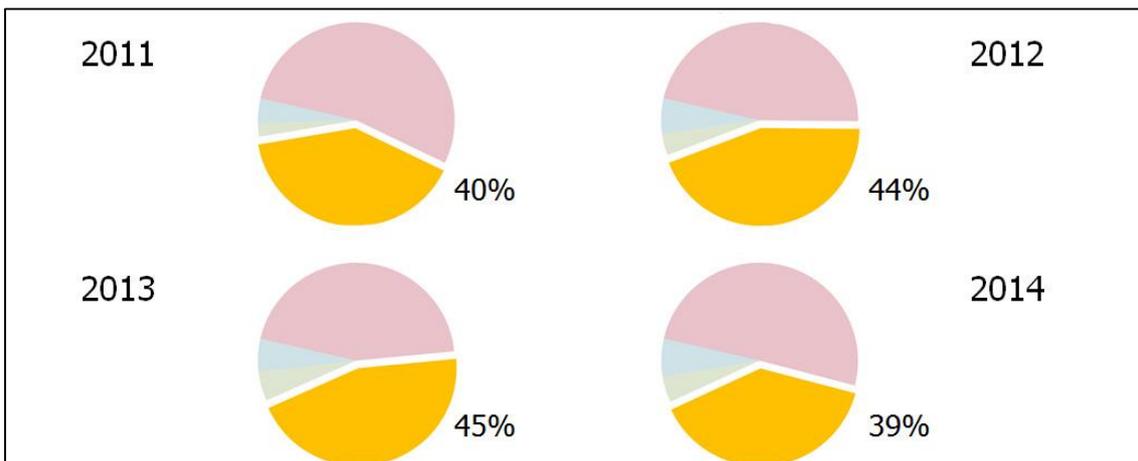


Figura 23: Porcentajes anuales de usuarios del resto de España

EUROPA

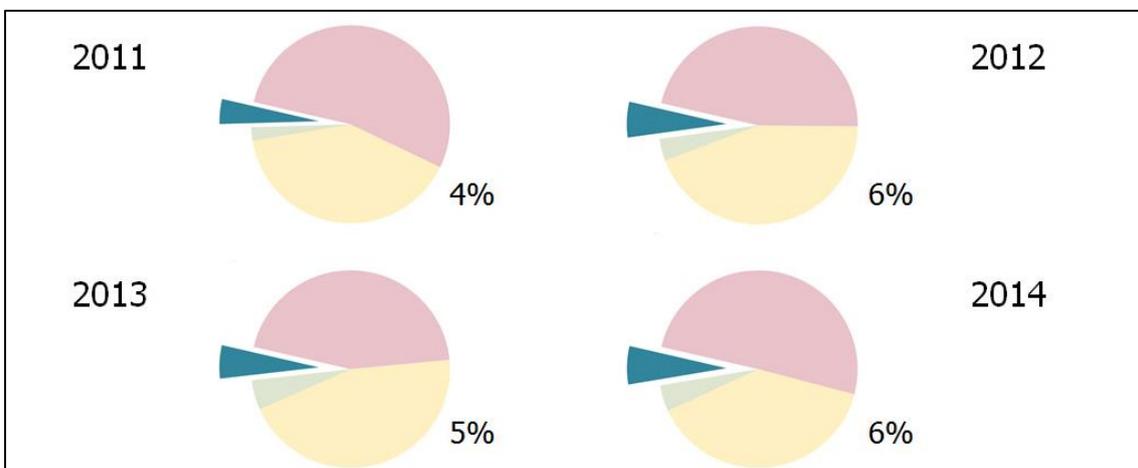


Figura 24: Porcentajes anuales de usuarios europeos

AMÉRICA

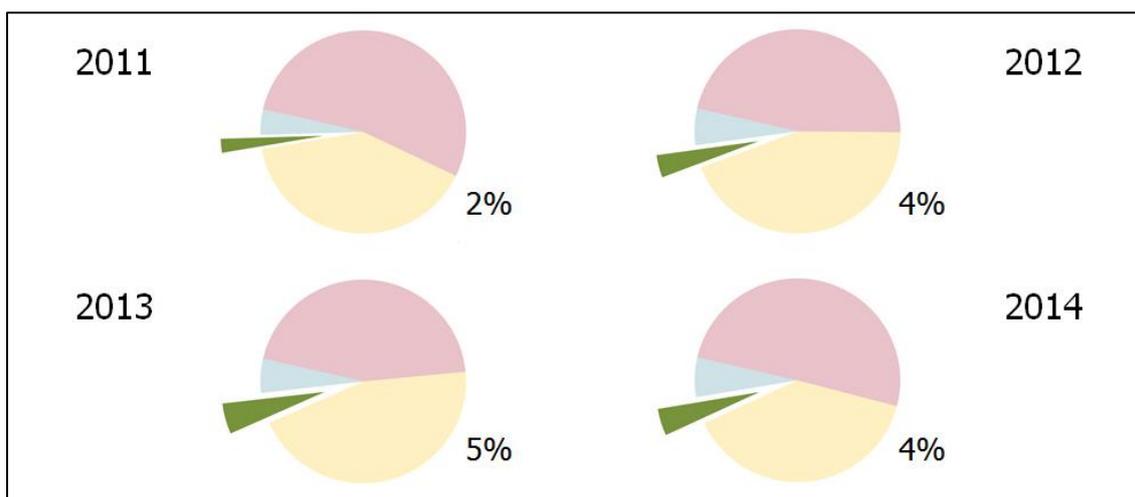


Figura 25: Porcentajes anuales de usuarios americanos

En el año 2011 la mayoría de los usuarios (54%) procedían de Aragón. En esta fase inicial de la evolución de DARA la difusión fuera de las fronteras de la Comunidad Autónoma era menor, como reflejan los datos.

Los años 2012 y 2013 suponen un incremento del porcentaje de usuarios tanto españoles, que alcanza su mayor peso específico en 2013 con un 45% de los usuarios totales, como Europeos (6% en 2012 y 5% en 2013). Los porcentajes de usuarios americanos también se incrementan, desde un exiguo 2% en 2011 al 4% en 2012 y 5% en 2013.

Hasta el año 2013 hay un crecimiento mayor, en porcentaje, de usuarios no aragoneses. El servicio se publicita y difunde también fuera, lo que conlleva un mayor flujo de usuarios de procedencias más alejadas.

En 2014, sin embargo, la tendencia se invierte. En este último año, ha habido un incremento notable del porcentaje de usuarios locales, que alcanza el 51%. Como se muestra en los siguientes gráficos, este crecimiento no conlleva un retroceso en el número de usuarios de otros segmentos, sino en un mayor incremento de usuarios aragoneses, provocado por acciones específicas de difusión, aparición en medios, etc.

Evolución trimestral de usuarios según procedencia geográfica⁷

ARAGÓN

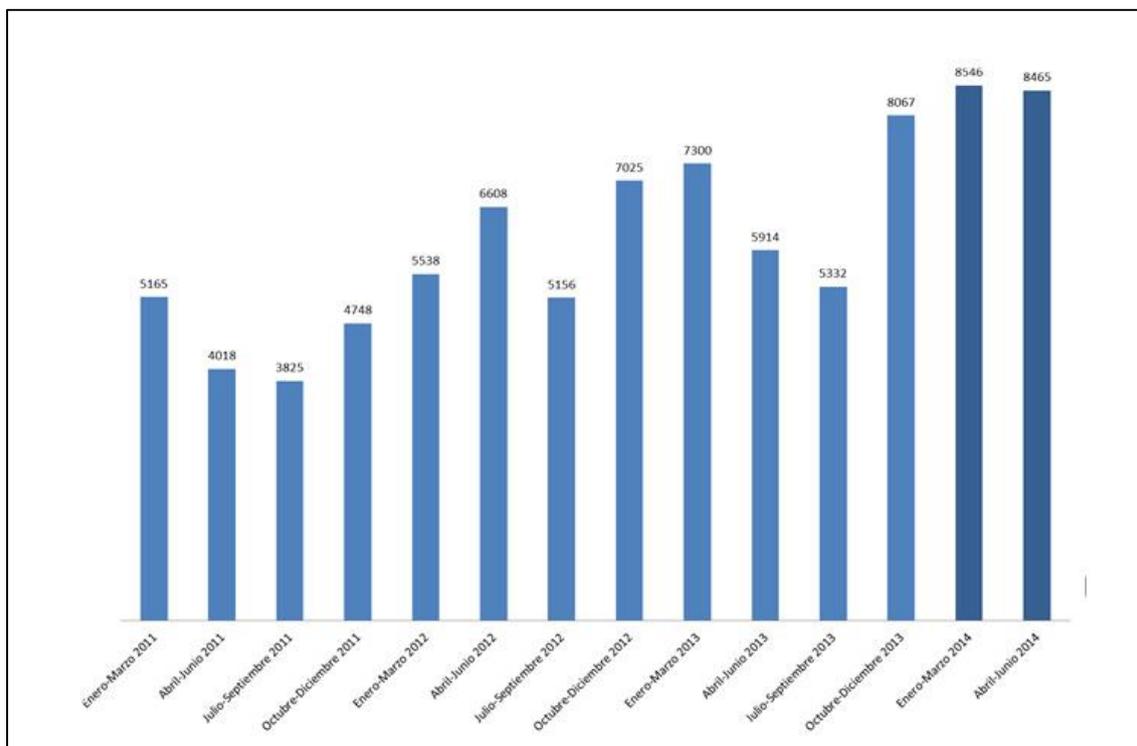


Figura 26: Evolución trimestral de usuarios aragoneses

RESTO DE ESPAÑA

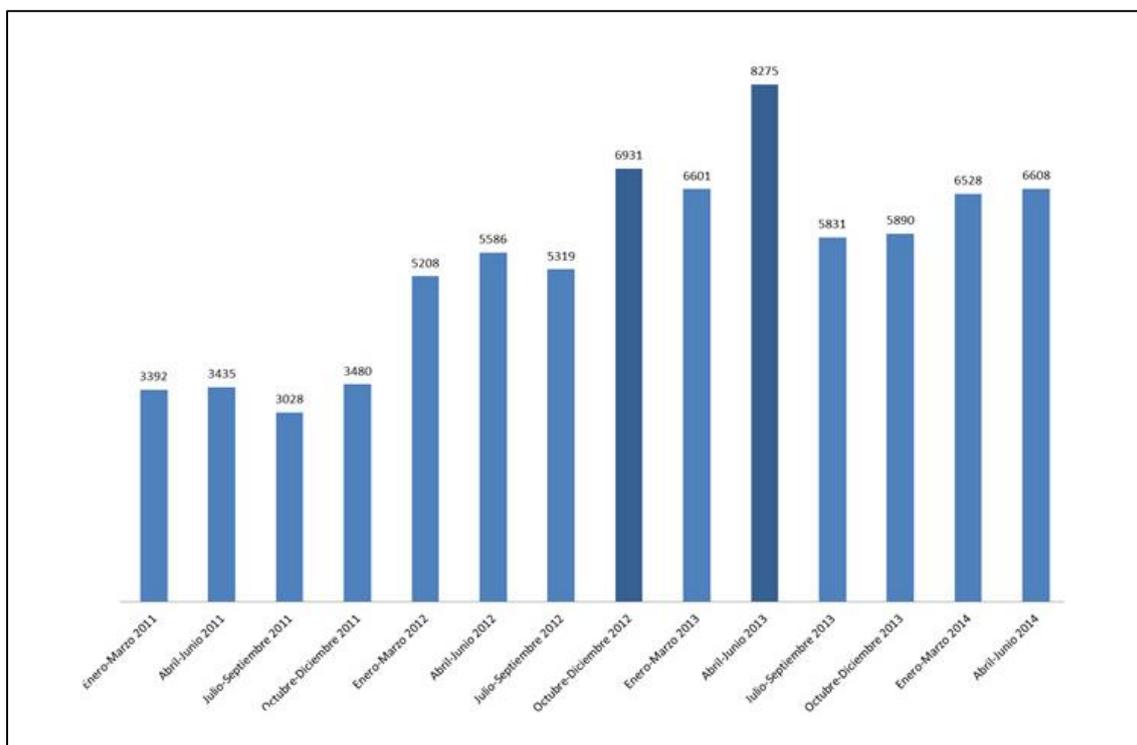


Figura 27: Evolución trimestral de usuarios del resto de España

⁷ Resaltados en tono más oscuro los valores trimestrales máximos

EUROPA

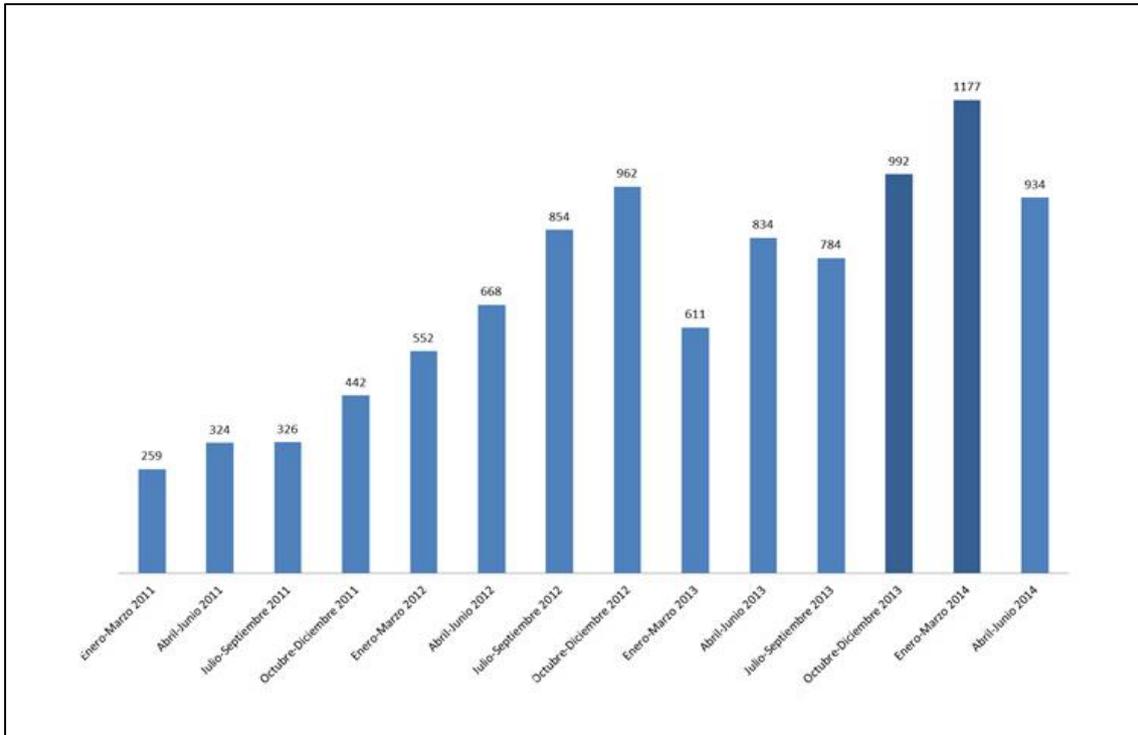


Figura 28: Evolución trimestral de usuarios europeos

AMÉRICA

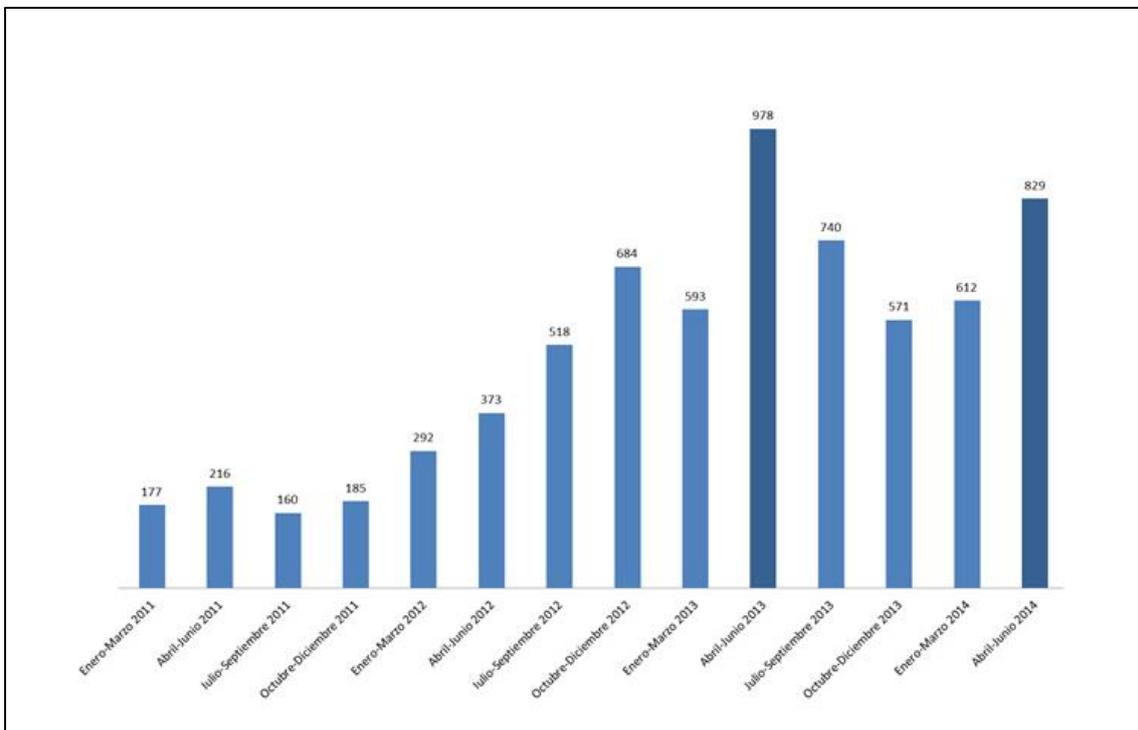


Figura 29: Evolución trimestral de usuarios americanos

De forma similar a la tendencia al crecimiento que pudimos observar en la evolución trimestral general de DARA, los cuatro segmentos poblacionales diferenciados también presentan, por lo general, un cierto crecimiento trimestre a trimestre. En el caso de los usuarios aragoneses, encontramos una constante evolución positiva y los valores máximos de la serie en los dos primeros trimestres de 2014.

En el caso de los usuarios españoles, la tendencia hasta 2013 es similar, aunque encontramos un descenso en el número de visitas de este segmento después de junio de 2013. Si comparamos los datos del trimestre enero-marzo de los años 2013 y 2014, puede observarse que los registros obtenidos el último año superan a los de 2013. No se trata, por tanto, de una disminución en la llegada de usuarios del resto de España, sino una combinación de crecimiento más relajado y un pico de usuarios muy llamativo entre abril y junio de 2013.

Similares circunstancias se observan en los segmentos europeo y americano: Un crecimiento trimestre a trimestre, muy buenos resultados a finales de 2013 y principios de 2014, y picos de afluencia pronunciados en trimestres señalados.

La combinación de las variaciones porcentuales anuales y del número de usuarios trimestrales para cada uno de estos segmentos nos permite resaltar una serie de ideas y conclusiones sobre la procedencia geográfica de los usuarios:

- DARA tiene una presencia eminentemente local. La difusión de proximidad y la naturaleza de los contenidos que pone a disposición de los usuarios fomentan una participación mayoritaria de los usuarios aragoneses.
- Trascendencia en el resto de España: pese la circunstancia anteriormente señalada, más del 40% de los usuarios del servicio, en los tres años y medio estudiados, proceden de fuera de las fronteras de la Comunidad.
- Hasta mediados del 2013 el porcentaje de usuarios procedentes de Aragón disminuía año a año. Los muy notables resultados del último año, sin embargo, invierten esta tendencia.

Usuarios europeos

En la figura 30 se muestran los países de Europa de los que proceden principalmente los usuarios de este segmento:

PAÍS	Nº de Sesiones	% de sesiones
FRANCIA	6.437	66,13%%
ITALIA	614	6,31%%
REINO UNIDO	609	6,26%
ALEMANIA	577	5,93%
BÉLGICA	325	3,34%
HOLANDA	199	2,04%
PORTUGAL	155	1,59%
LUXEMBURGO	145	1,49%
SUIZA	136	1,40%
ANDORRA	76	0,78%

Figura 30: Países de procedencia de usuarios europeos

Destaca sobre el resto de países Francia con el 66,13% del total de usuarios europeos, frente a las cifras más homogéneas de Italia (6,31%), Reino Unido (6,26%) o Alemania (5,93%). El número de usuarios franceses es diez veces más alto que el correspondiente a los tres siguientes países con mayor incidencia en DARA.

Si atendemos a los métodos de captación de los usuarios franceses, pueden observarse dos fuentes principales: Europeana⁸ y páginas de contenido genealógico. Son estas fuentes, principalmente las páginas privadas de asociaciones dedicadas a la genealogía y afines, quienes proveen el mayor número de usuarios franceses a DARA (que contiene miles de documentos de interés para este perfil de usuarios).

⁸ Biblioteca Digital Europea, <http://www.europeana.eu/>

Usuarios americanos

PAÍS	Nº de Sesiones	% de sesiones
ESTADOS UNIDOS	1.333	19,24%
MÉXICO	1.176	16,97%
COLOMBIA	1.000	14,43%
ARGENTINA	991	14,30%
BRASIL	419	6,05%
PERÚ	370	5,34%
CHILE	331	4,78%
VENEZUELA	286	4,13%
ECUADOR	199	2,87%
CANADÁ	111	1,60%

Figura 31: País de procedencia de los usuarios americanos

En el caso de los usuarios americanos la distribución por país tiene una distribución menos desnivelada que en Europa. Destacan por su número los usuarios procedentes de Estados Unidos (19,24%, con un 21,52% de esos usuarios navegando en idioma inglés), seguidos de México (casi un 17%), Colombia (14,43%) y Argentina (14,30%).

Indicadores cualitativos de usuarios europeos frente a usuarios americanos

Indicador	EUROPA	AMÉRICA
Páginas vistas	103.031	28.575
Páginas/sesión	10,58	4,12
% Rebote	23,57%	50,58%
% Nuevas sesiones	44,35%	83,37%

Figura 32: Comparativa de algunos indicadores de calidad de uso de DARA. Usuarios Europeos/Americanos

En la figura 32 se muestra de forma comparada algunos de los indicadores de uso de DARA para los segmentos poblacionales usuarios europeos y usuarios americanos, en concreto *páginas vistas*, *páginas por sesión*, *porcentaje de rebote* y *porcentaje de nuevas sesiones*. Los datos

obtenidos muestran una diferencia notable respecto al comportamiento de los mencionados grupos.

Los usuarios europeos hacen un uso del servicio de información mucho más intenso que los usuarios americanos. Cada sesión (visita) de usuarios provenientes de Europa visualiza como media 10,58 páginas individuales, frente a las 4,12 de los usuarios americanos, para un total de 103.031 páginas individuales visitadas frente a 28.575. También muestran un rebote mucho menor, con un 23,57% frente al 50,58%. Otros indicadores de implicación muestran la misma tendencia.

La tasa de nuevos usuarios es sin embargo mucho más intensa en el caso de los usuarios americanos, con un sorprendente 83,37% frente al 44,35% de los europeos.

A la vista de los datos cabe preguntarse a qué se debe esta diferencia de comportamiento. Descartada la diferencia idiomática –que debería afectar de forma negativa a los usuarios europeos, con diferentes lenguas nativas- ¿Por qué razón los usuarios de los países americanos hacen un uso menos intenso del servicio?

Analizando de forma comparada diversos factores, se observa que la principal diferencia entre ambos segmentos es la *ruta por la que acceden a DARA*, es decir, el método de captación:

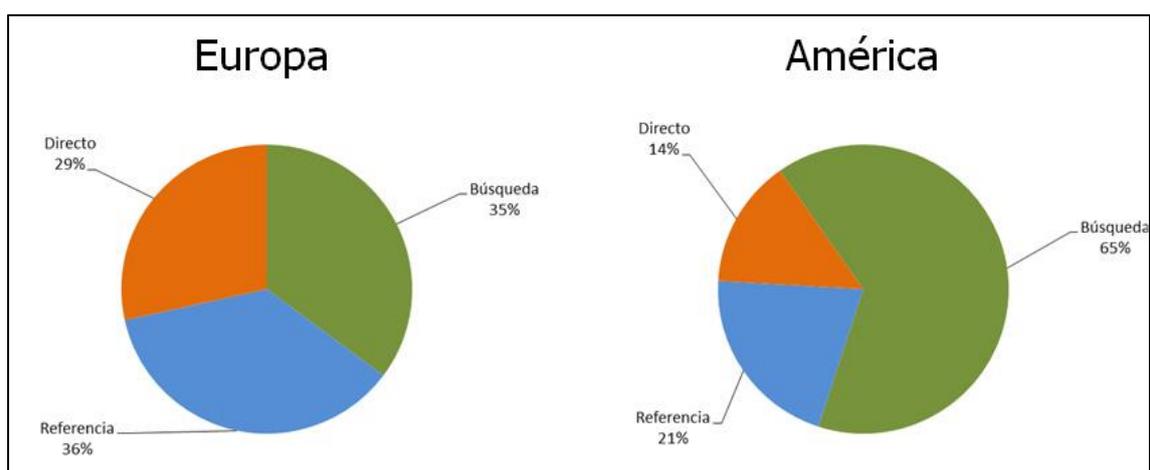


Figura 33: Método de adquisición de usuarios europeos y americanos

Como puede observarse, existe una notable diferencia en la ruta que toman unos y otros usuarios para acceder a DARA. Los usuarios europeos, que recordemos mostraban unos indicadores de uso más intenso del servicio, distribuyen el método de adquisición de forma homogénea entre tráfico de búsqueda (35%), tráfico de referencia (36%) y tráfico directo (29%). Es notable el reducido porcentaje de usuarios que acceden a través de buscadores orgánicos (un 16% menor que la media general de DARA)

Los usuarios americanos, por su parte, acceden de forma mayoritaria a través del uso de buscadores orgánicos, principalmente a través de Google (65%). Sólo un 21% acceden a través de enlaces en otras páginas y un 14% a través de tráfico directo.

A la luz de estos datos ¿puede concluirse que el método de adquisición de usuarios, el sistema que éstos utilizan para encontrar y acceder al servicio de información cultural, va a tener

relevancia en cuanto a los indicadores de intensidad y calidad de uso del servicio? ¿Responden, además, los diferentes métodos de captación de usuarios a diferentes necesidades de información de los mismos?

Para responder a estas preguntas, se han configurado diferentes segmentos poblacionales basados en la captación de usuarios, procediéndose a su estudio individual. Se estudiarán en el siguiente apartado.

Estudio de segmentos poblacionales basados en el método de captación

Selección y configuración de segmentos

Sin una configuración previa y una selección de las variables que es necesario estudiar, Google Analytics solo podría ser usado para estudiar de forma diferenciada tres grandes segmentos basados en el método de adquisición: usuarios de tráfico de búsqueda orgánica, usuarios de tráfico directo y usuarios de tráfico de referencia (a través de hipervínculos externos)

Se ha considerado necesario una configuración avanzada de estos segmentos para responder a las necesidades de información expuestas en el epígrafe anterior. Partiendo de los diferentes segmentos de captación de usuarios, se han añadido nuevas variables que tratarán de relacionar el método utilizado con las necesidades de información de los usuarios y de las diferentes estrategias de difusión emprendidas por el servicio. Son los siguientes:

- Tráfico de búsqueda: búsqueda de servicio de información/archivo/cultural.
- Tráfico de búsqueda: búsqueda de contenidos concretos.
- Tráfico de referencia: red de enlaces institucionales.
- Tráfico de referencia: enlaces externos.
- Tráfico directo.

Segmentos de tráfico de búsqueda

Segmento 1: Tráfico de búsqueda: búsqueda de servicios de información, archivo o culturales.

Este segmento incluye exclusivamente usuarios que acceden a DARA a través de búsquedas en buscadores orgánicos como Google. Además de este primer filtro, se han añadido otros de tal forma que sólo se incluirán aquellos usuarios que estén buscando activamente un servicio de archivo online o directamente DARA Documentos y archivos de Aragón.

El perfil de este segmento responde a aquellos usuarios que

- Están buscando un archivo en línea para responder a sus necesidades de información, conozcan previamente su existencia o no.
- Están buscando activamente DARA, bien porque ya han accedido previamente al servicio o bien por haber tenido noticia de la existencia del servicio.

Las variables usadas para la configuración de este segmento son *método de adquisición: tráfico de búsqueda, términos concretos de búsqueda y página de destino: portada del portal o cualquiera de los subportales que lo componen.*

Segmento 2: Tráfico de búsqueda: búsqueda de contenidos concretos.

Este segmento incluye exclusivamente usuarios que acceden a DARA a través de búsquedas en buscadores orgánicos como Google. Además de este primer filtro, se han añadido otros de tal forma que sólo se incluirán aquellos usuarios que estén buscando un contenido concreto, sea cual sea su naturaleza, y “aterricen” en el servicio de archivo online. Es decir, usuarios que estén buscando en su buscador orgánico de elección, por ejemplo, un documento histórico o de cualquier tipo, y esa búsqueda les dirija al archivo online.

El perfil de este segmento responde a aquellos usuarios que:

- Están buscando un contenido de cualquier naturaleza, un documento concreto, persona o institución.
- Es posible que conozcan de antemano la existencia del servicio de información, pero esta circunstancia no es necesaria para la consecución de su acceso.

Las variables usadas para la configuración de este segmento son *método de adquisición: tráfico de búsqueda, términos concretos de búsqueda y página de destino: URL correspondiente a un contenido concreto, sea éste una digitalización, una descripción u otro.*

Para la captación de estos dos tipos de usuario es condición previa la indexación tanto del servicio como de los contenidos que contiene en buscadores orgánicos. Sin esta circunstancia los usuarios no accederán al servicio cultural.

Segmentos de tráfico de referencia

Segmento 3: Tráfico de referencia: referencias desde la red de enlaces institucionales.

Este segmento incluye exclusivamente usuarios que para acceder al servicio han realizado un “click” en un hipervínculo existente en una página externa al servicio. Además de este primer filtro, se han seleccionado una serie de enlaces de tipo institucional (pertenecientes a Gobierno de Aragón) u otros pertenecientes a otras instituciones y cuya existencia se ha buscado activamente por parte de la gerencia del servicio (como PARES, EUROPEANA, HISPANA y otros).

El perfil de este segmento responde a aquellos usuarios que:

- Buscan el servicio cultural de forma activa en otras páginas institucionales existentes.
- Tienen una necesidad concreta de información, satisfecha a través de otros portales culturales en internet, que redirigen el usuario hasta DARA.

Las variables usadas para la configuración de este segmento son *método de adquisición: tráfico de referencia y referencias desde un corpus de enlaces delimitado basado en una política activa de difusión en internet.*

Para la existencia de este segmento es necesaria la presencia de enlaces en otros portales institucionales o de cualquier tipo, establecida previamente a modo de difusión institucional.

Por su importancia específica en este segmento, en el estudio se ha diferenciado el conjunto de usuarios de este tipo provenientes de SIPCA, considerados como un enlace institucional.

Segmento 4: Tráfico de referencia: referencias desde páginas externas cualesquiera.

Este segmento incluye exclusivamente usuarios que para acceder al servicio lo han hecho a través de un hipervínculo existente en una página externa al servicio. Además, se han excluido las visitas pertenecientes al “segmento 3”, por lo que solo se mostrarán los usuarios que acceden a través de enlaces existentes en páginas privadas de cualquier tipo, sean de empresas, personales, asociaciones, medios de comunicación, portales de corporaciones municipales, etc.

El perfil de este segmento responde a aquellos usuarios que:

- Tienen una necesidad de información que esperan cubrir, en un primer momento, en otras páginas especializadas o de cualquier otra naturaleza.
- Conocen y acceden al servicio a través de páginas de medios de comunicación en línea.

Segmento 5: Tráfico directo

Este segmento incluye mayoritariamente a aquellos usuarios que acceden al servicio bien escribiendo la URL del portal, directamente, en sus navegadores, o bien por haber seleccionado el portal entre sus “favoritos”.

Son usuarios que ya han accedido a servicio con anterioridad, o que conocen su existencia por cualquier medio fuera de internet.

Porcentaje total y evolución anual de cada segmento

En el siguiente gráfico (figura 34) se expone la distribución porcentual de cada uno de los segmentos indicados, para el rango cronológico estudiado al completo:

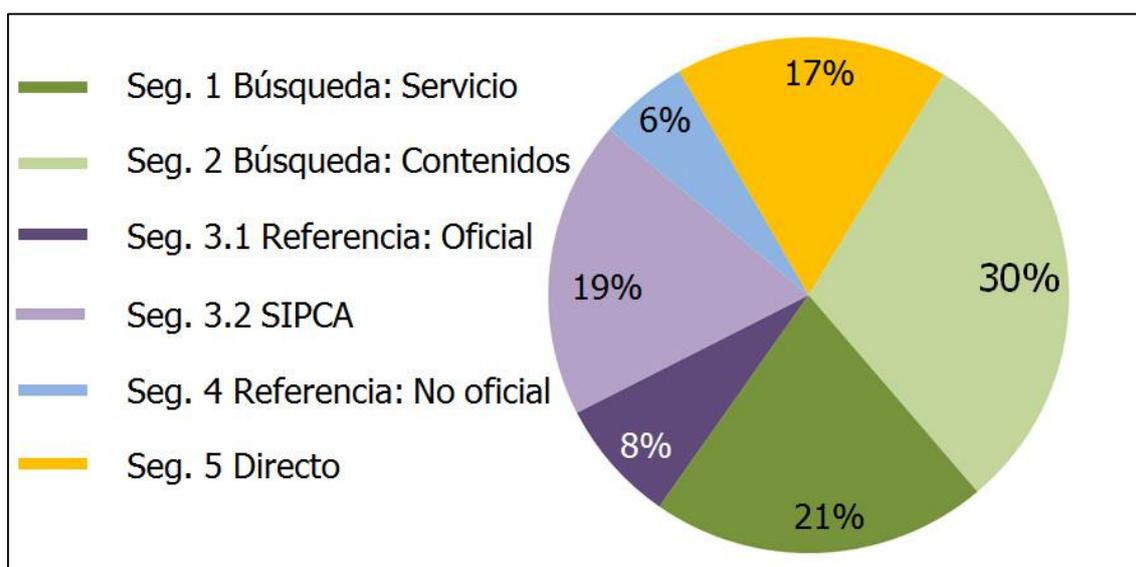


Figura 34: Distribución porcentual según segmentos de adquisición

El tráfico de búsqueda, como se vio en los datos generales, es el principal método de adquisición

de usuarios. La segmentación avanzada muestra, además, que el segmento 2, correspondiente a los usuarios que realizan una búsqueda orgánica con términos correspondientes a un contenido concreto, es el método de captación de usuarios mayoritario, con un 30% del total. Le sigue a continuación el segmento 3, que aísla los datos de usuarios cuya puerta de entrada son los enlaces de tipo institucional. En el gráfico se ha desglosado el 27% total que le corresponde en los datos de SIPCA, que sirve un 19%, y del resto de enlaces, con un 8%. SIPCA se revela como un proveedor de usuarios de primera importancia y su existencia tiene un fuerte impacto en el éxito de DARA. Además, parte de DARA está ubicada en el propio SIPCA, lo que incrementa este suceso.

El tercer procedimiento más usado por los usuarios es el correspondiente al segmento 1, es decir, usuarios que realizan una búsqueda orgánica con los términos “archivo” o “DARA” o “AHPZ” (por poner tres ejemplos), que suponen el 21% del total.

El tráfico directo muestra un 17%, y finalmente el tráfico proveniente de enlaces externos un 6%.

Estos datos muestran en primer lugar la importancia de un buen posicionamiento en buscadores orgánicos, como ya se indicó. Pero además de un buen posicionamiento será de gran relevancia la correcta indexación de los contenidos del servicio cultural, de forma que éstos se relacionen con las búsquedas de ítems concretos por parte de los usuarios en sus buscadores habituales. La mayoría de los usuarios de DARA, como puede observarse, accede al servicio buscando un documento, contenido, institución, persona, etc. concretos en Google: el acceso a DARA es el resultado de esas búsquedas.

Segmentos 1 y 2: Tráfico de búsqueda

Las figura 35 muestra los porcentajes totales, año a año, de visitas obtenidas por estos dos segmentos.

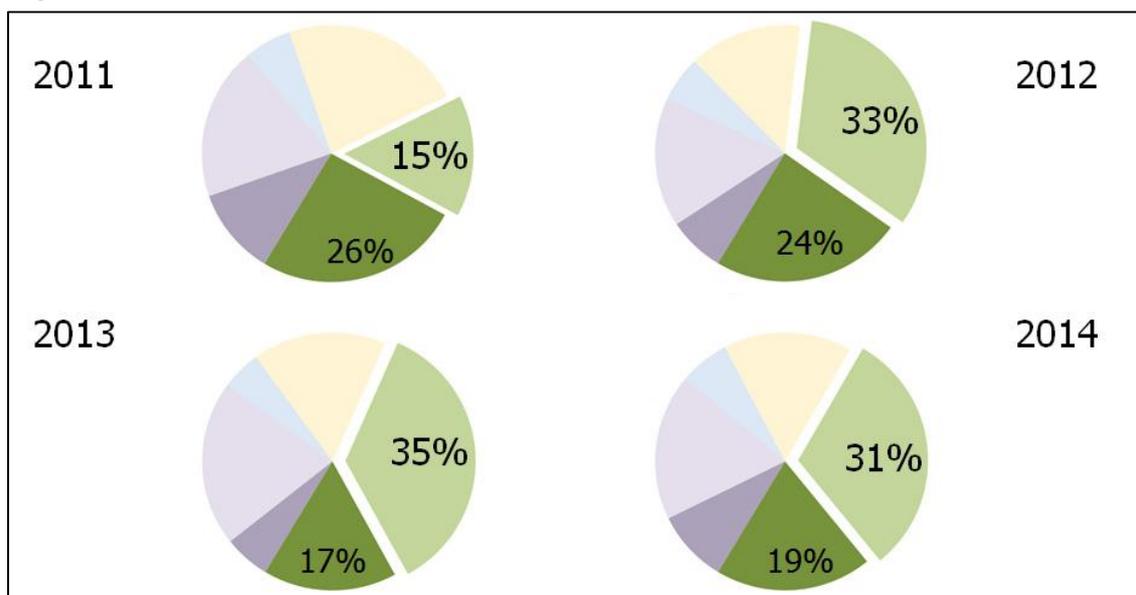


Figura 35: porcentajes anuales de segmentos de búsqueda: Segmento 1 (búsqueda de servicio, más oscuro) y Segmento 2 (búsqueda de contenidos, más claro).

La correcta indexación de contenidos en buscadores orgánicos tuvo lugar mediado el 2011, de ahí el porcentaje mostrado por el segmento 2 (en tono más claro) ese año. En 2012 y 2013 el porcentaje específico de dicho segmento aumenta hasta alcanzar el 35% del total de visitas a DARA, su máximo histórico.

Paralelamente, el porcentaje de usuarios que busca activamente el servicio se reduce desde un 26% hasta un 19%. Finalmente, el año 2014 ve una cierta estabilización, con un recrecimiento del segmento 1 y un decrecimiento del segmento 2.

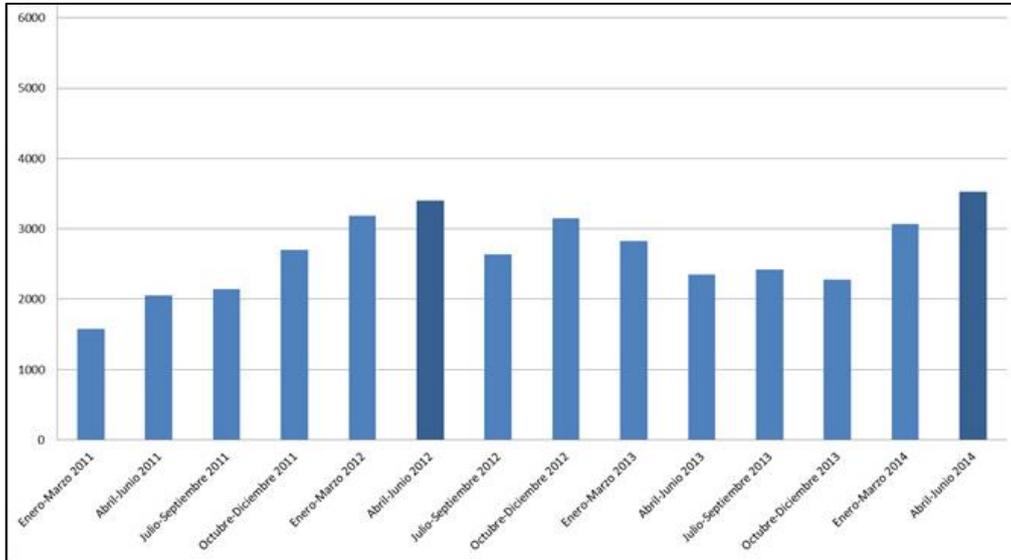


Figura 36: valores trimestrales del segmento1: búsqueda de servicio

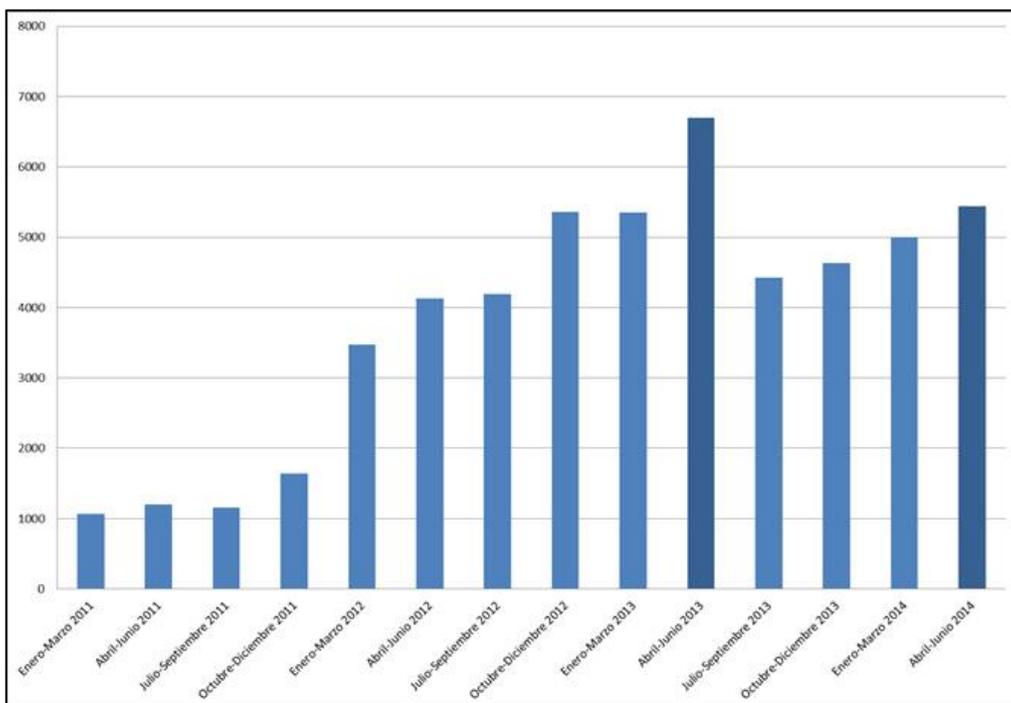


Figura 37: valores trimestrales del segmento1: búsqueda de contenidos

Las figuras 36 y 37 muestran, respectivamente, los valores trimestrales registrados para los segmentos 1 y 2.

En el primero de ellos, se observa que tras una rápida escalada en el año 2011 (correspondiente a las primeras etapas de difusión y posicionamiento en buscadores), el número de visitas correspondiente a este segmento fluctúa y en general se mantiene, con un notable aumento en el año 2014⁹.

El segmento 2 no comienza su crecimiento hasta el trimestre octubre-diciembre de 2011, fechas en las que se lleva a cabo la correcta indexación de contenidos en buscadores orgánicos. A partir de estos meses este segmento tiene un notable crecimiento, alcanzando sus valores máximos en abril-junio de 2013. Tras el descenso estacional de ese año, el crecimiento continúa, si bien a un ritmo más moderado.

Segmentos 3 y 4, tráfico de referencia

Las figuras 37 y 38 muestran la distribución anual de porcentajes de los usuarios procedentes del tráfico de referencia, correspondientes a los segmentos 3 y 4.

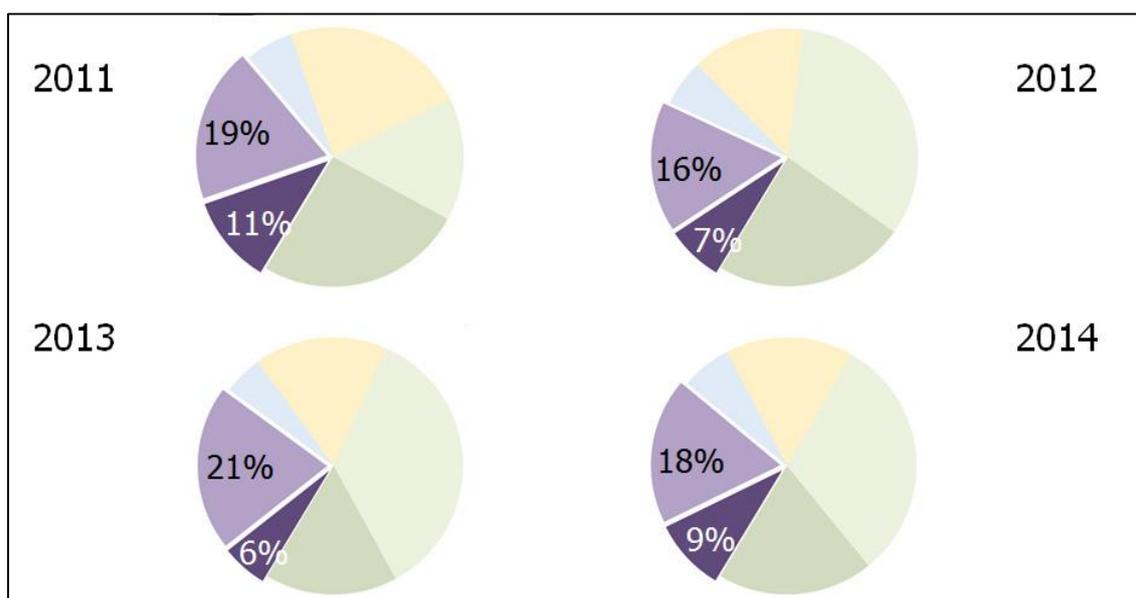


Figura 38: porcentajes anuales del segmento 3, red de enlaces institucionales, con distinción de SIPCA y resto

En la figura 38 se muestran los porcentajes de los usuarios que acceden a través de SIPCA (en color claro) y del resto de enlaces institucionales (en color oscuro). Debe tenerse en cuenta que los valores totales de visitas del año 2011 son muy inferiores a los correspondientes a los años posteriores.

Destaca en primer lugar la notable importancia de SIPCA respecto a la captación de usuarios, con valores que rondan de forma consistente el 19% del total de los usuarios de DARA. El resto de enlaces institucionales presenta un crecimiento sostenido desde el año 2012 hasta el 2014,

⁹ Correspondiente, en gran medida, a la aparición en diarios de la comunidad, edición papel, de noticias referentes a DARA. 23 de junio de 2014.

en número totales, y especialmente el último año, en el que casi un 10% de los usuarios totales del servicio acceden a través de este tipo de hipervínculos.

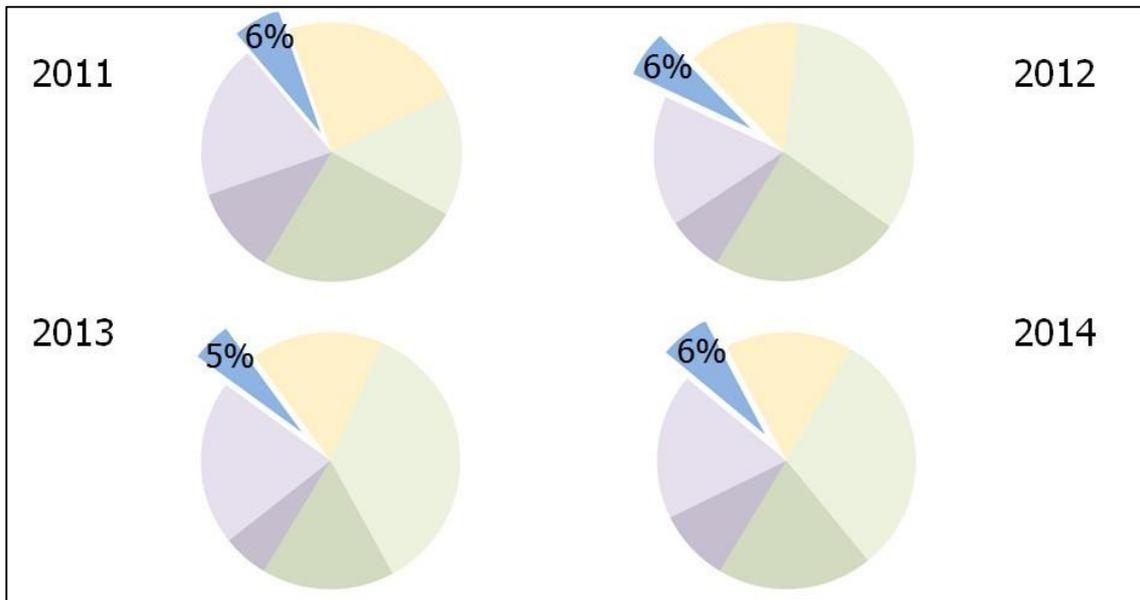


Figura 39: porcentajes anuales del segmento 4, red de enlaces no institucionales.

Los valores porcentuales del segmento 4, correspondiente a usuarios que acceden a través de hipervínculos no asociados directamente con el servicio, se mantienen a lo largo del rango cronológico establecido (alrededor de un 6% de los usuarios totales de DARA).

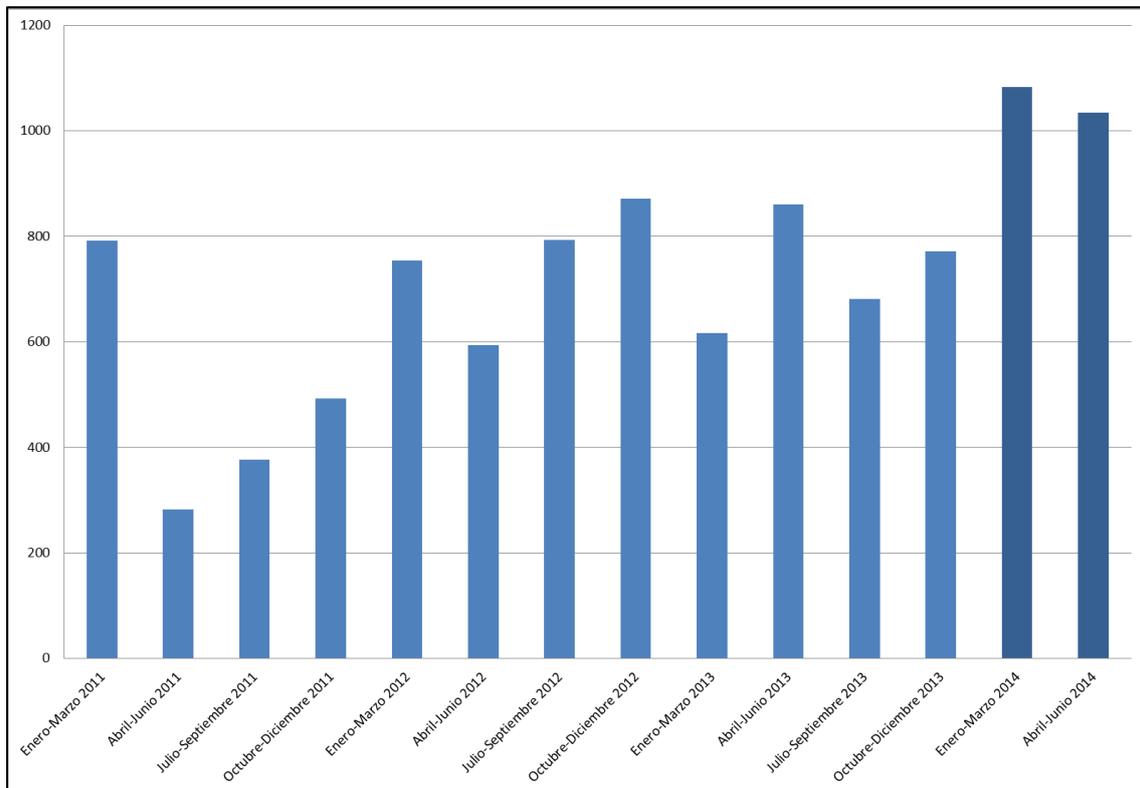


Figura 40: valores trimestrales del segmento 4, red de enlaces no institucionales

La figura 40 muestra los valores trimestrales de este segmento, y en ella puede observarse una tendencia sostenida de crecimiento. Destacan además los datos obtenidos en los dos últimos trimestres de la serie -pertenecientes al año 2014- por sus valores más elevados. En la figura 41 se muestran las fuentes de referencia más representativas de este segmento en los dos primeros trimestres de 2014:

Página de referencia	Total de visitas
Wikipedia.org	588
Aragongen.org	170
Facebook.com	150
Lapuebladealfinden.es	104
m.facebook.es	93
Liberadosdelolvido.org	79
Gensobrarbe.blogspot	77
m.wikipedia.org	57
Genealogie.com	52
Afigen.blogspot	45

Figura 41: Principales páginas no institucionales de referencia

El notable crecimiento experimentado por este conjunto de visitantes en el último año corresponde por lo tanto a Wikipedia (muy especialmente), redes sociales como Facebook y páginas de temática genealógica como Aragongen.org o gensobrarbe.blogspot.

Indicadores de comportamiento

Las diferencias en los indicadores de comportamiento y uso del servicio (páginas por sesión, rebote, porcentaje de nuevas visitas, uso de búsqueda, etc) observadas entre los segmentos poblacionales procedentes de Europa y América ya apuntaba hacia una posible relación entre el método de captación de los usuarios y la manera en que éstos hacen uso del servicio.

En los siguientes apartados se van a mostrar los datos específicos de cada uno de los segmentos basados en el método de captación, con el objetivo de detectar e interpretar posibles diferencias que relacionen estos grupos de visitas con tendencias generales de comportamiento.

Segmentos 1 y 2, tráfico de búsqueda

A continuación se muestran de forma comparada los porcentajes relativos al valor general de DARA para los segmentos 1 y 2, búsqueda de servicio y búsqueda de contenidos.

PORCENTAJE DE NUEVOS USUARIOS

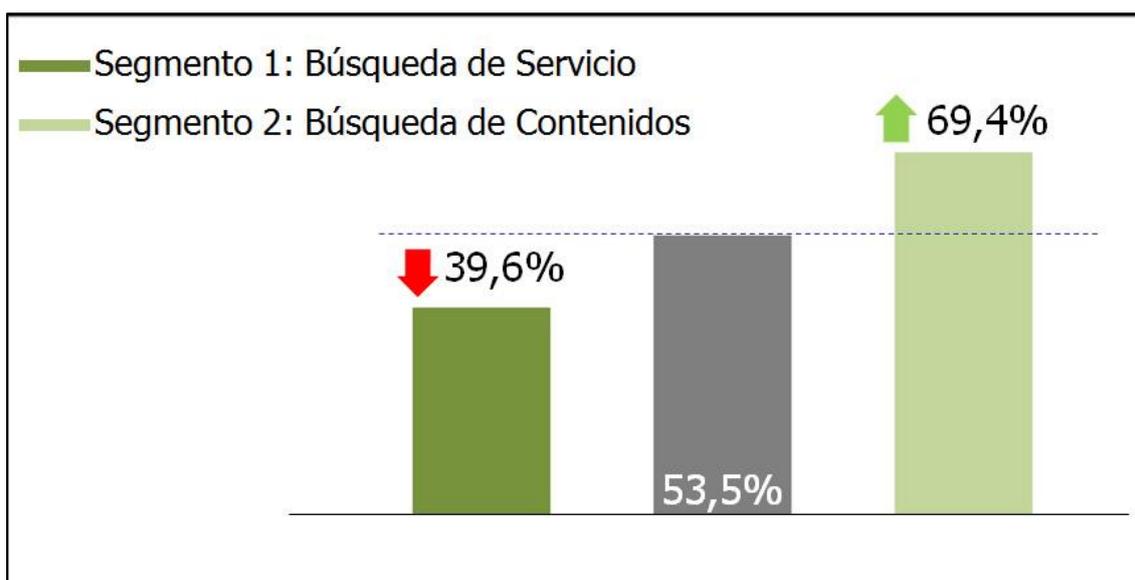


Figura 42: Porcentaje de nuevos usuarios de los segmentos 1 y 2 en relación con el valor de DARA general

PORCENTAJE DE REBOTE

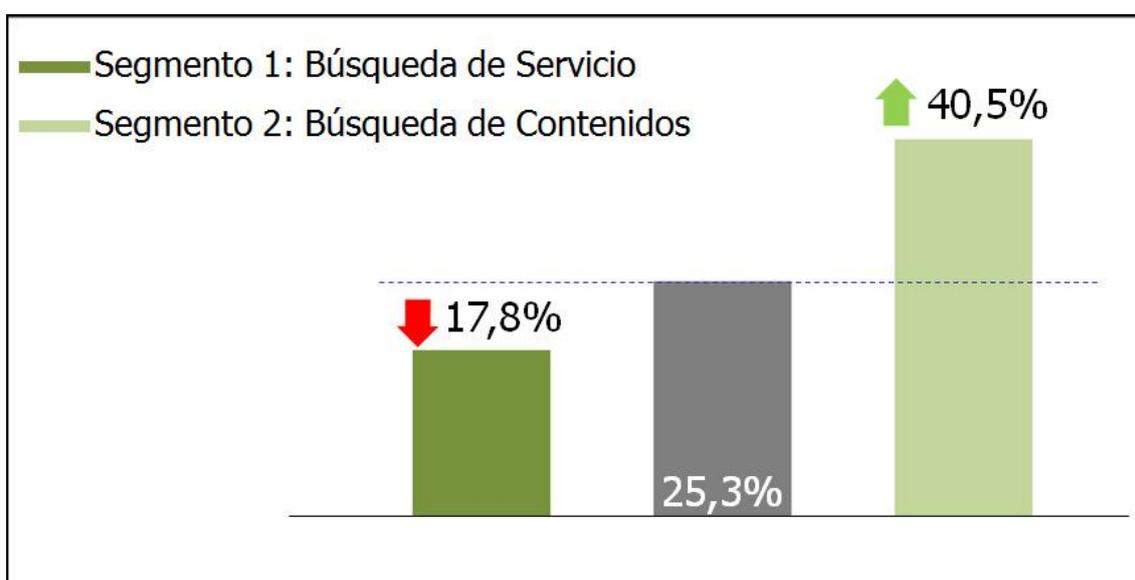


Figura 43: Porcentaje de rebote de los segmentos 1 y 2 en relación con el valor de DARA general

Las figuras 42 y 43 muestran los valores de los indicadores *porcentaje de nuevos usuarios* y *porcentaje de rebote*. El primero de ellos indica el porcentaje de usuarios que acceden por primera vez al servicio –identificando usuario con cookie– y el segundo el porcentaje de usuarios que accede al servicio y sale rápidamente del mismo sin utilizar sus funcionalidades ni consolidar su navegación.

El segmento 2, tráfico de búsqueda de contenidos, muestra un amplio porcentaje de nuevos usuarios, un 69,4%, superior al 53,5% de DARA general. Es una muestra del potencial de

captación de usuarios que tiene el posicionamiento en buscadores orgánicos y concretamente la indexación de contenidos. El segmento 1, búsqueda de servicio (recordemos, usuarios que buscan un archivo online o directamente DARA o alguno de sus subportales) tiene un porcentaje de nuevos usuarios mucho menor, de poco menos de un 40%.

El porcentaje de rebote (visitas de baja intensidad) es también mucho más alto en el segmento 2, con un 40,5%. El segmento 1 muestra un notablemente bajo 17,8%, casi ocho puntos por debajo de la media de DARA general.

A la luz de estos datos pueden extraerse las siguientes interpretaciones:

- El segmento 1 tiene una menor tasa de nuevos usuarios. Esto indica que muchos de los usuarios que acceden de esta manera ya habían visitado con anterioridad el archivo online.
- El segmento 1 tiene un muy bajo porcentaje de rebote: Los usuarios que acceden al servicio buscando activamente un archivo online o DARA permanecen en el portal y hacen uso de él.
- El segmento 2 tiene un gran potencial de captación de nuevos usuarios. Para muchos de los visitantes de DARA, una búsqueda en un buscador orgánico como Google es la primera toma de contacto con el portal.
- El segmento 2 tiene un alto porcentaje de usuarios que salen rápidamente del portal, quizás al no satisfacer éste sus necesidades de información. Sin embargo, la tarea de difusión, la puesta en conocimiento de la existencia del servicio, ha sido realizada.

PÁGINAS VISITADAS POR SESIÓN

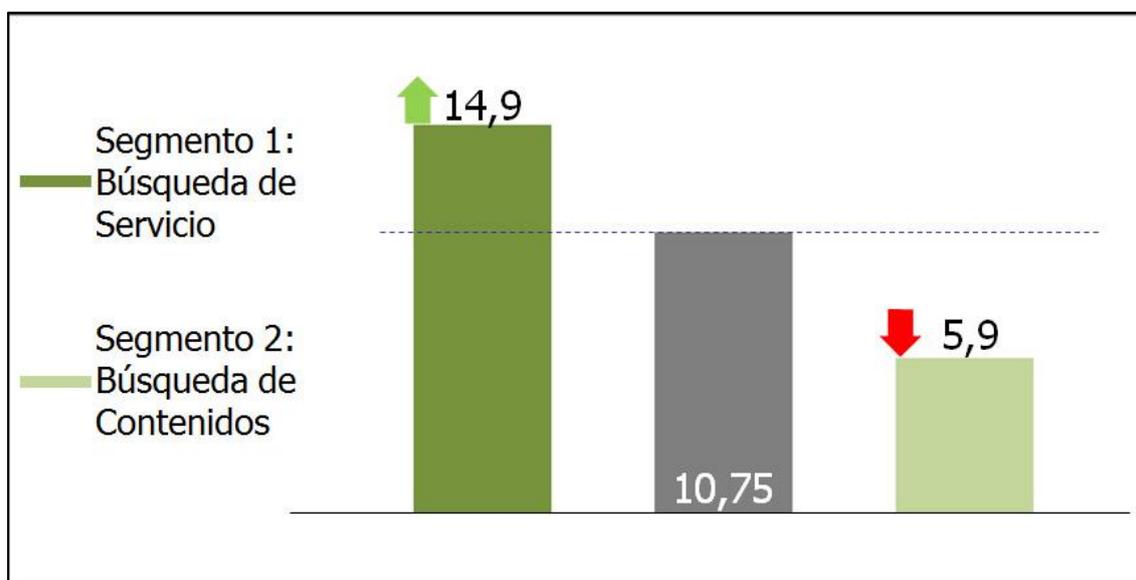


Figura 44: Valores medios de páginas por sesión para los segmentos 1 y 2 en relación con los valores de DARA

PORCENTAJES DE SESIONES CON BÚSQUDA

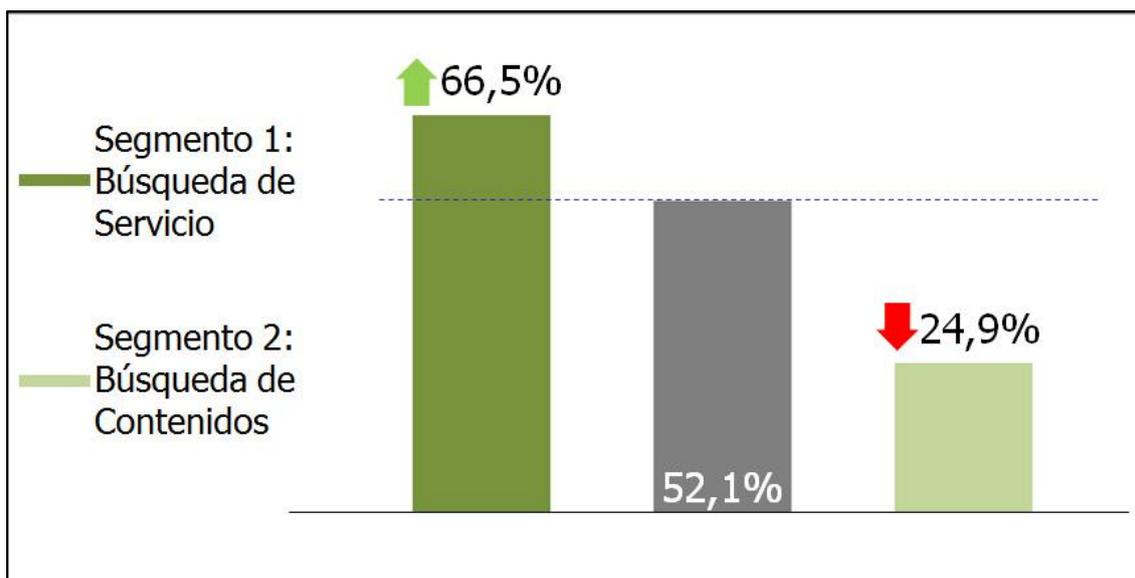


Figura 45: Porcentaje de sesiones con búsqueda de los segmentos 1 y 2 en relación con el porcentaje de DARA

Las figuras 44 y 45 incluyen los valores correspondientes a dos indicadores relevantes para el estudio del uso y las funcionalidades del servicio cultural: *páginas por sesión* y *porcentaje del uso de la herramienta de búsqueda*.

Los valores positivos se invierten en este caso. Los usuarios del segmento 1 muestran unos valores de uso mucho más elevados no ya que el segmento 2, sino en relación con los valores medios de DARA, con una media de casi 15 páginas vistas por sesión y un 66,5% de las visitas haciendo uso de las funcionalidades de búsqueda del portal. Los usuarios del segmento 2 tienen medias y porcentajes mucho más reducidos (5,9 páginas por sesión y 24,9% de porcentaje de visitas con búsqueda¹⁰).

INTERACCIÓN

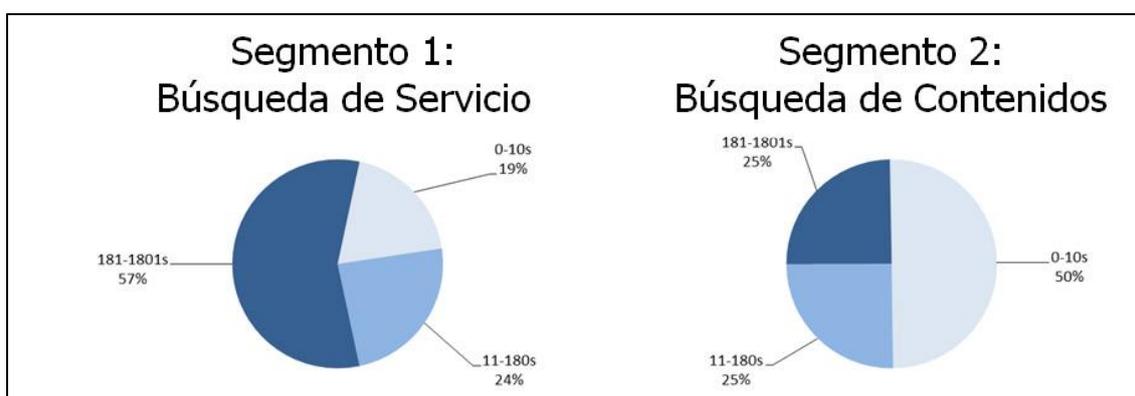


Figura 46: Interacción de los usuarios de los segmentos 1 y 2 con el servicio según tiempo invertido en la página

¹⁰ En este caso, google o el buscador orgánico elegido por el usuario ha funcionado *de facto* como herramienta interna de búsqueda.

Los porcentajes de interacción de los usuarios con el portal, mostrados en la figura 46, replican este modelo de comportamiento. En el caso del segmento 1, hasta un 57% de las visitas invierten 3 o más minutos en la navegación y consumo de contenidos en el portal. El segmento 2 muestra valores menores, con un 50% de los usuarios realizando visitas efímeras (todas ellas rebotes), y un 25% de los usuarios realizando consultas más prolongadas de 3 o más minutos.

Segmentos 3 y 4: red de enlaces

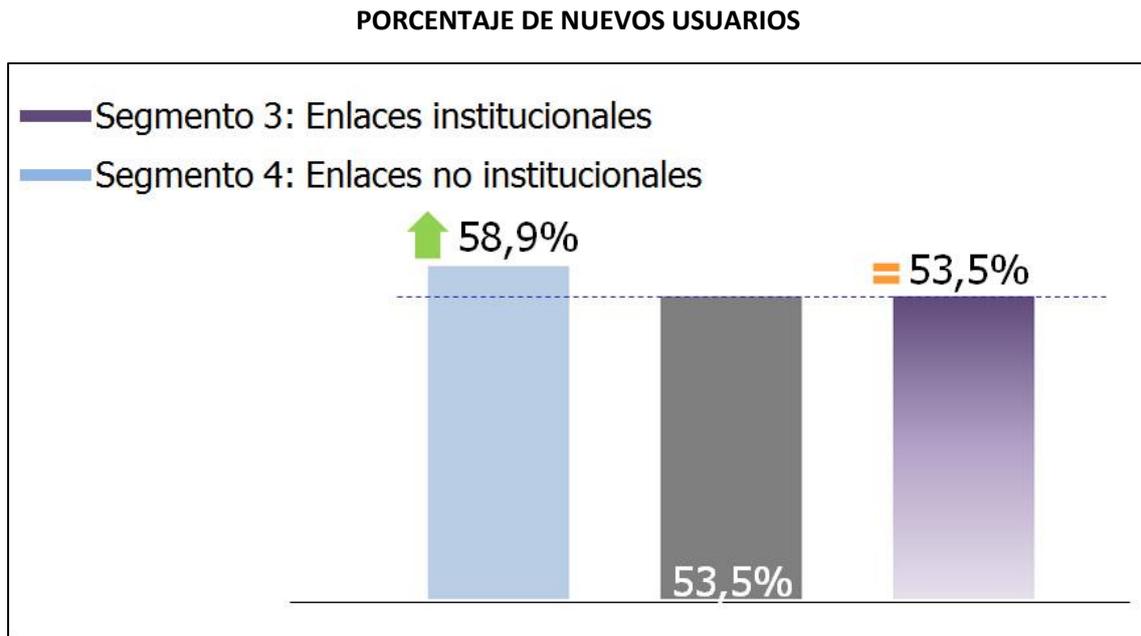


Figura 47: Porcentaje de nuevos usuarios de los segmentos 3 y 4 en relación con el valor de DARA general

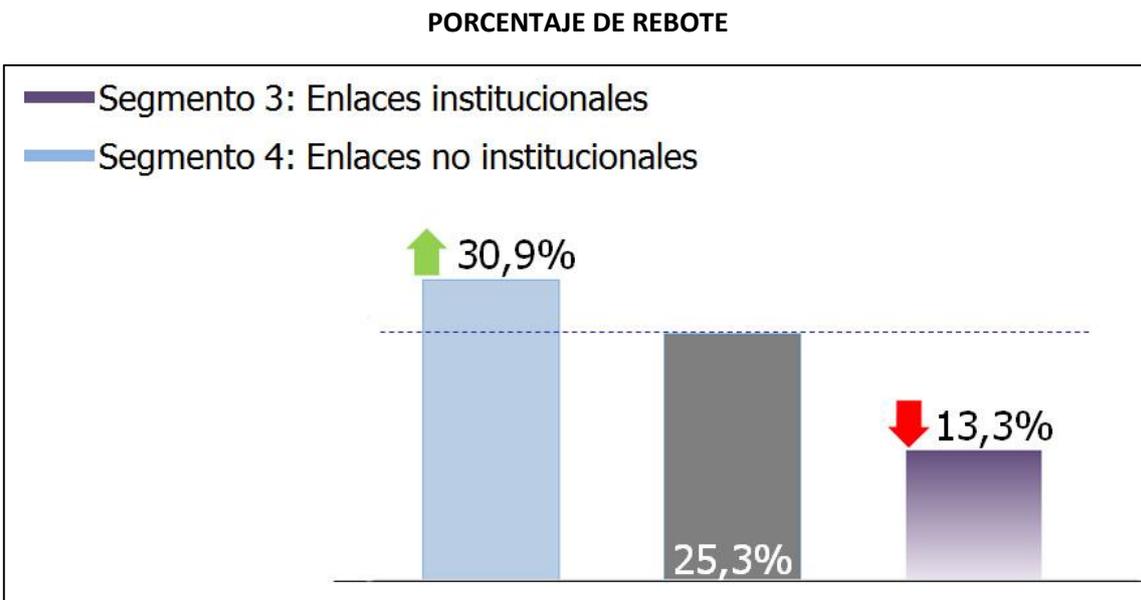


Figura 48: Porcentaje de rebote de los segmentos 3 y 4 en relación con el valor de DARA general

Los valores de los indicadores nuevos usuarios y porcentaje de rebote también muestran comportamientos diferenciados entre los segmentos 3 y 4, correspondientes respectivamente a los usuarios que acceden a través de enlaces institucionales (propios de Gobierno de Aragón o de otras instituciones con las que se ha acordado la inclusión de enlaces y contenidos) y a los usuarios que acceden a través de enlaces existentes en cualquier otra página.

El segmento 3 muestra un potencial de captación de nuevos usuarios idéntico al ofrecido por el conjunto de datos DARA, y ligeramente menor que el segmento 4. Un 58,9% de los usuarios que acceden a través de enlaces ofrecidos por páginas de forma “espontánea” no había accedido con anterioridad al archivo online.

Los datos del indicador *porcentaje de rebote* señalan que el segmento 3, de entre todos los estudiados, es el que ofrece un valor más bajo (13,3%). Los usuarios que acceden a DARA a través de las páginas de Gobierno de Aragón, o en portales especializados como Hispana o Europea son los que tienen una mayor tasa de permanencia en el servicio. El porcentaje de rebote del segmento 4 es superior a la media con un 30,9%.

PÁGINAS POR SESIÓN

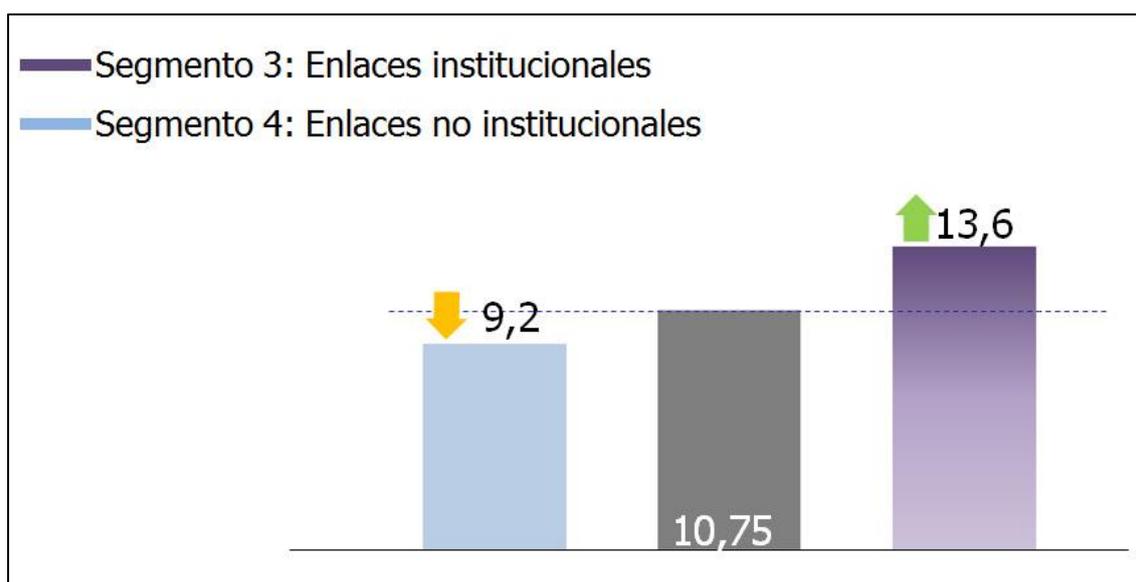


Figura 49: Valores medios de páginas por sesión para los segmentos 3 y 4 en relación con los valores de DARA

PORCENTAJE DE SESIONES CON BÚSQUEDA

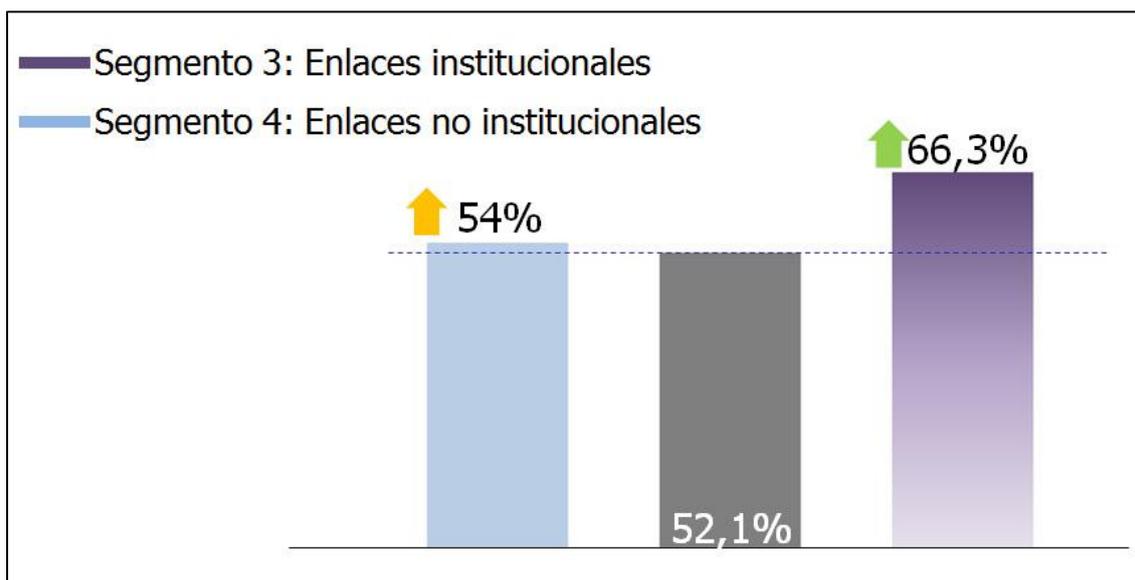


Figura 50: Porcentaje de sesiones con búsqueda de los segmentos 3 y 4 en relación con el porcentaje de DARA

Las figuras 49 y 50 revelan los datos almacenados para los indicadores de calidad de uso del servicio *media de páginas por sesión* y *porcentaje de sesiones con búsqueda*. De nuevo pueden observarse diferencias entre ambos segmentos. Los usuarios que acceden a través de enlaces institucionales consumen un mayor número de páginas por sesión (13,6), valor superior a la media de DARA. Además, este tipo de usuarios hace uso de las herramientas de búsqueda ofrecidas por el portal en un 66,3% de las sesiones registradas.

El segmento 4 muestra valores relativamente próximos a los ofrecidos por la media de DARA en general. La media de páginas por sesión es moderadamente inferior (9,2) y el porcentaje de visitas con búsqueda es ligeramente superior, con un 54% del total.

INTERACCIÓN

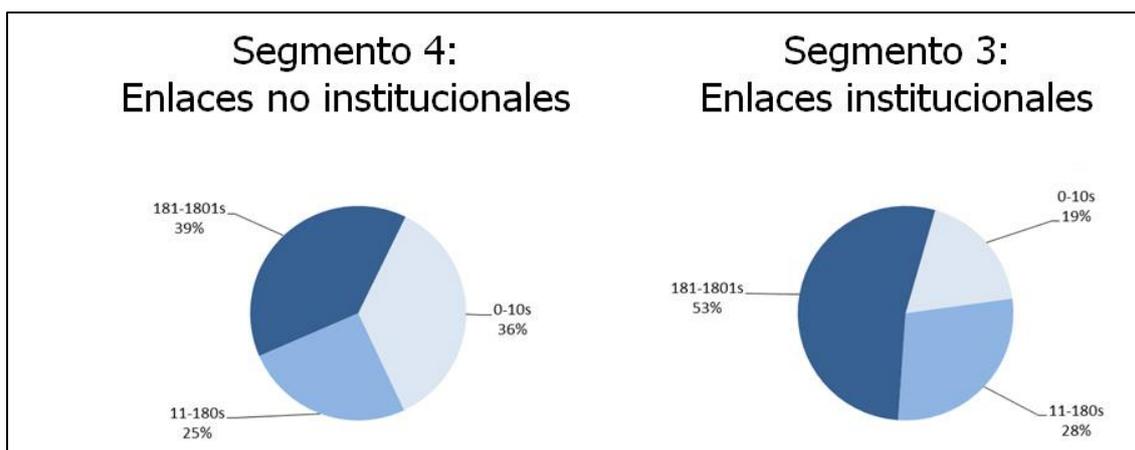


Figura 51: Interacción de los usuarios de los segmentos 3 y 4 con el servicio según tiempo invertido en la página

Respecto a los porcentajes de interacción (figura 51), de nuevo el segmento 3 muestra datos de uso más intenso del servicio con un 53% de usuarios permaneciendo más de 180 segundos en el portal (10 puntos más que la media de DARA). Los valores del segmento 4 son muy similares a la media del conjunto de los usuarios, con un 39% de usuarios de alta interacción.

Interpretación y tendencias de indicadores cualitativos

- Porcentaje de nuevos usuarios:

Este indicador muestra el porcentaje de usuarios que accede por primera vez al archivo online. Usado de forma comparada puede revelar cuáles son los métodos de adquisición de usuarios con mayor potencial de captación de visitas que nunca antes habían estado en el servicio de información cultural.

Los segmentos 2 (tráfico de búsqueda: búsqueda de contenidos) y 4 (tráfico de referencia: enlaces no institucionales) son los segmentos que revelan los valores más altos de este indicador, sobresaliendo especialmente el primero con un 69,4%.

- Porcentaje de rebote:

Un bajo porcentaje de rebote es un valor positivo, ya que indica que un número más elevado de las visitas que acceden al servicio permanecen en él y hacen uso de sus funcionalidades y contenidos.

Debe mencionarse que existe una relación directa entre el porcentaje de nuevos usuarios y el porcentaje de rebote: a mayor número de usuarios nuevos en el servicio, mayor es también el porcentaje de usuarios que rebotan, que no acceden de forma real al mismo.

Según el método de captación, los segmentos con un menor porcentaje de rebote son el número 3 (tráfico de referencia: enlaces institucionales) con un 13,3% seguido del segmento 1 (tráfico de búsqueda: búsqueda de servicio) con un 17,8%. Los usuarios con mayor tendencia al rebote son los adquiridos mediante búsquedas de contenidos (segmento 2) con un 40,5%.

- Media de páginas por sesión

La media de páginas por sesión es el valor medio de páginas o URL (identificables con contenidos) que visitan los usuarios cada vez que acceden al servicio. Los segmentos con valores más positivos de este indicador son el segmento 1 (14,9) y el segmento 3 (13,6)

- Uso de funcionalidades de búsqueda

El porcentaje de sesiones con búsqueda muestra el porcentaje de visitas de un segmento dado que realizan búsquedas en los buscadores internos provistos por el portal de archivos. En este caso, de nuevo son el segmento 1 (66,5%) y el segmento 3 (66,3%) los que revelan porcentajes más elevados.

- Interacción

La interacción es medida a través tiempo que invierten los usuarios en el servicio en cada una de las visitas que realizan, teniendo especial relevancia aquellas que permanecen más de 180 segundos en el servicio. Los segmentos 1 (57% de usuarios de alta interacción) y 3 (53% de usuarios de alta interacción) muestran los valores más altos de este indicador.

Uso de los datos: potencialidades a la difusión y difusión en internet

A la hora de aplicar de forma práctica las conclusiones que puedan obtenerse de este conjunto de análisis debe tenerse en cuenta que todos los segmentos estudiados tienen un potencial relativo a cada uno de los indicadores cualitativos estudiados. De esta forma, las sesiones pertenecientes al segmento 2, búsqueda de contenidos, pese a ofrecer altos valores de rebote y bajos índices de páginas por sesión, será capaz de atraer usuarios que hagan uso efectivo e intenso del servicio y de sus funcionalidades. Del mismo modo, los usuarios pertenecientes al segmento 3, referencias de enlaces institucionales, pese a ofrecer bajos porcentajes de nuevos usuarios también es capaz de atraer nuevos usuarios.

Ahora bien, las características más indicativas de cada segmento pueden utilizarse para trazar estrategias de difusión concretas de las que puedan esperarse resultados similares a los estudiados.

Sabemos ahora del elevado potencial de captación de nuevos usuarios de las búsquedas de contenidos en google y de la existencia de enlaces en páginas no institucionales, tales como las ofrecidas por redes sociales, medios de comunicación, webs personales o de asociaciones.

Una indexación correcta de los contenidos del portal cultural va a tener como resultado seguro la obtención de un gran número de visitas de usuarios que, de otro modo, no tendrían por qué conocer la existencia del servicio o acceder a él. Favorecer y alentar la existencia de enlaces en webs no institucionales también se muestra como una buena estrategia a la hora de captar usuarios que potencialmente tienen pocas posibilidades de conocer y visitar el servicio. Llevando a cabo dicha indexación y optimizando el servicio cultural para ser accesible por Google y otros buscadores orgánicos, es posible convertir estos servicios, de uso mayoritario, en herramientas propias del portal cultural.

Pero no se trata exclusivamente de recibir visitas, sino de que los usuarios detrás de las estadísticas satisfagan sus necesidades de información, que hagan uso del servicio, de sus funcionalidades y subportales. Para **conectar** este perfil de usuarios de alta interacción y uso de funcionalidades con la existencia del servicio cultural deberán potenciarse los mecanismos de adquisición relativos a los segmentos 1 y 3, la búsqueda activa del servicio en buscadores orgánicos y la existencia de enlaces de tipo institucional. Es decir, favorecer la llegada de usuarios de “alta intensidad” o interacción en el uso del servicio.

En el primero de los casos, es esencial una buena optimización en motores de búsqueda (SEO) o posicionamiento. Además, siguen siendo vitales los mecanismos “clásicos” de difusión tales como publicaciones, sesiones y jornadas, aparición en los medios, etc. Dar a conocer de esta forma la existencia del servicio cultural online provocará un incremento en las visitas pertenecientes a este segmento, cuyos altos valores de interacción y uso del servicio hemos podido comprobar.

En el segundo caso, el establecimiento de una red de enlaces institucional tanto en los servicios online propios como de otras instituciones va a favorecer el acceso de usuarios de un perfil definido, de uso intenso del servicio. La colaboración entre servicios culturales afines o de naturaleza similar provocará no solo un incremento en el número de visitas, sino también en los valores del resto de indicadores cualitativos del uso del servicio.

Se ha podido comprobar en definitiva que es posible definir ciertos grupos de usuarios en función de características concretas de captación o método de adquisición.

Estas características, además, están relacionadas con acciones concretas e identificadas de difusión cultural en el ámbito online.

A lo largo del estudio propuesto se ha podido acceder a una serie de interpretaciones y conclusiones mencionadas en cada uno de los epígrafes estudiados. Entre ellos cabe destacar la incidencia de la naturaleza y contenidos del servicio cultural en la afluencia y comportamiento de los usuarios, el éxito en el uso de funcionalidades propuestas, y el impacto de las acciones de difusión emprendidas tanto dentro de nuestras fronteras como fuera de ellas y el grado de impacto que éstas tienen sobre el comportamiento de los usuarios. Sirva el análisis de los datos obtenidos como elemento que cimente y mejore las estrategias emprendidas hasta el momento y, con ellas, el uso, planificación y funcionamiento de los servicios culturales en internet.