

VIII Jornadas sobre Tecnologías de la Información para la Modernización de las
Administraciones Públicas (TECNIMAP 2004)
Murcia, 28 de septiembre-1 de octubre

Grupo 6. Perspectivas de futuro para el desarrollo de la Administración Electrónica. 6.3
Arquitecturas tecnológicas para una nueva Administración.

Integrando la Información de las Administraciones Públicas en la Gestión del Conocimiento. Una Solución desde la Web Semántica y los Servicios Web

David F. Barrero

Investigador
Departamento de Automática
Universidad de Alcalá de Henares
Correo-e: dfbarrero@wanadoo.es

J. Ignacio Criado

Investigador
Departamento de Ciencias Políticas y de la Administración II
Universidad Complutense de Madrid
Campus de Somosaguas s/n, 28223, Madrid, España-Spain
Correo-e: jicriado@cps.ucm.es
Web: <http://www.jicriado.com>

Resumen

Las Administraciones Públicas han adoptado diversas herramientas tecnológicas que han incrementado su potencial para generar, almacenar y distribuir información. Por diversas cuestiones tanto político-administrativas como técnicas se utilizan formatos normalmente incompatibles y, por lo tanto, difíciles de integrar, de manera que la gestión de la información es una tarea compleja que requiere en muchas ocasiones desarrollos de *middlewares* complejos y de elevado coste. En la presente ponencia se analiza la importancia de tener un acceso integrado a la información de una Administración Pública y se ofrece una solución a dicho problema denominada *Searchy*. Éste es un motor de búsqueda de metainformación distribuido de concepto avanzado que se encuentra dentro del estado de arte actual, desarrollado entre RedIRIS y la Universidad de Alcalá de Henares, capaz de integrar la localización de fuentes arbitrarias de información traduciendo semánticamente sus contenidos, lo cual es directamente aplicable en organizaciones intensivas en conocimiento, como las Administraciones Públicas.

Palabras clave

Searchy, recuperación de información, world wide web, , middleware

Integrando la Información de las Administraciones Públicas en la Gestión del Conocimiento.

Una Solución desde la Web Semántica y los Servicios Web¹

David F. Barrero² y J. Ignacio Criado³

1. Introducción

Las nuevas oportunidades que las tecnologías ofrecen a las Administraciones Públicas se orientan a diferentes ámbitos. En el caso concreto de la recuperación de la abundante información disponible en las Administraciones Públicas, nuevas herramientas tecnológicas como la Web Semántica (WS) y los Servicios Web (SW) permiten llevar a efecto algunas de las potencialidades de la Gestión del Conocimiento (GdC) en el ámbito público.

Desde un plano teórico, la información se suele almacenar sin ninguna referencia a su contexto y éste es el origen de los problemas en la búsqueda posterior. La existencia de nociones sobre el contexto relacionado con la información de una Administración Pública implica limitar el espacio de búsqueda a un conjunto más reducido de recursos, es decir, una GdC eficiente requiere considerar la información como un todo y tener capacidad para el manejo de su contexto, en definitiva, resulta necesario trabajar con la semiótica para conseguir traspasar la línea que separa la gestión de la información de la GdC.

Desde un plano técnico, en la práctica diaria la mayoría de las Administraciones Públicas no hace el mejor almacenamiento posible de su información de cara a la integración de sistemas porque se deben satisfacer diversos intereses dentro de un entorno político-administrativo complejo. La información suele estar dispersa en diferentes unidades administrativas y consideraciones técnicas pueden aconsejar utilizar diferentes soportes para la generación y almacenamiento de la información. Los posibles soportes de información pueden llegar a ser sumamente dispares,

¹ Los autores quisieran mostrar su agradecimiento a Diego R. López, responsable del área de *Middleware* y Aplicaciones de RedIRIS, así como a Óscar García Población, profesor del Departamento de Automática de la Universidad de Alcalá de Henares, personas que han hecho posible que *Searchy* sea una realidad. Los agradecimientos no estarían completos sin mencionar a Javier Masa Marín, así como a José Manuel Macías, ambos investigadores de RedIRIS.

² Investigador. Departamento de Automática, Universidad de Alcalá de Henares, dfbarrero@wanadoo.es.

³ Investigador. Departamento de Ciencias Políticas y de la Administración II, Universidad Complutense de Madrid, <http://www.jicriado.com>. Dispone de una beca FPI de la Comunidad de Madrid, co-financiada por el Fondo Social Europeo.

desde complejos gestores de bases de datos, servicios de directorio y otros sistemas de *back end* hasta ficheros singulares creados por el personal de la Administración con un procesador de textos. Esta disparidad de soportes hace difícil la gestión de toda esa información, obligando en ocasiones a desarrollar complejos, a la par que caros, sistemas *middleware* personalizados que los aglutinen.

El objetivo de esta ponencia consiste, teniendo en cuenta la realidad concreta del entorno político-administrativo, en presentar una solución de *software* llamada *Searchy* desarrollada por *RedIRIS*, junto al *Departamento de Automática* de la *Universidad de Alcalá de Henares*, para hacer frente a los problemas señalados. El propósito es doble teniendo en cuenta las arquitecturas tecnológicas para una nueva Administración Pública: por un lado, facilitar la integración de la localización de recursos de información y, por otro, expresarlos en términos semánticos por medio de estándares compatibles con las últimas tecnologías web.

La ponencia opera planteando un marco teórico sobre las Administraciones Públicas en la Sociedad de la Información y del Conocimiento y la GdC en el sector público teniendo en cuenta la realidad de las páginas web. Luego se formula el marco tecnológico de la aplicación en el que se resumen las ventajas de uso de la WS y de los SW. En tercer lugar, se plantea una descripción general del sistema desarrollado. Finalmente, se cierra la ponencia con unas breves conclusiones que resumen el contenido y abren nuevas posibilidades de desarrollo futuro de la aplicación.

2. Marco Teórico

Las Administraciones Públicas cuentan con datos e información que no se explota de la manera más adecuada posible. La información en las organizaciones, pero especialmente en las Administraciones Públicas como consecuencia de su especial naturaleza político-administrativa, suele estar aislada en silos organizativos, vinculada a una diversidad de tecnologías a veces incompatibles entre sí y sin clara referencia a un entorno que permita su transformación en conocimiento directamente utilizable interna y externamente. A pesar de ese panorama, la llamada Sociedad de la Información y del Conocimiento está abriendo nuevas oportunidades a las Administraciones Públicas para mejorar la explotación de su información, especialmente, como consecuencia de la difusión de nuevas tecnologías y de redes de comunicaciones. Junto a ello, el enfoque de la GdC aplicado en el sector público ha planteado la necesidad de que las Administraciones Públicas tengan en cuenta la importancia creciente de los activos intangibles, es decir, que se establezcan sistemas de gestión capaces de transformar en conocimiento el caudal de información organizativa, especialmente, teniendo en cuenta el nuevo entorno Web.

2.1. Las Administraciones Públicas en la Sociedad de la Información y del Conocimiento

Durante el siglo XX las Administraciones Públicas se convirtieron en usuarias intensivas de tecnologías, adoptando el equipamiento más moderno que se encontraba disponible en cada momento. El sistema de organización burocrático se fundamentaba en unas tecnologías que hacían que los documentos en papel pasaran por toda la jerarquía para aprobar o informar de las decisiones, de manera que las estructuras organizativas, en parte, descansaban en esos desarrollos tecnológicos. La Sociedad de la Información y del Conocimiento ha planteado nuevos retos para las Administraciones Públicas.

En la Sociedad de la Información y del Conocimiento el acceso a las Tecnologías de la Información y la Comunicación (TIC) se expande y distribuye gracias a las nuevas redes de comunicaciones y sus servicios asociados. Entre las décadas de 1970 y 1990 se ha producido una incorporación creciente de sistemas de información más flexibles dentro de las organizaciones, que permiten una generación, almacenamiento y distribución de la información como nunca antes se había producido (Seneviratne, 1999). De forma más reciente la generalización de las nuevas capacidades de comunicación externas e internas se basan en el acceso a la red mundial Internet y sus diferentes servicios, sobre todo, la Web (Bellamy y Taylor, 1998).

En un contexto en que las tecnologías de información y las redes de comunicaciones se sitúan en el centro de la actividad de las organizaciones, las Administraciones Públicas están inmersas en profundos dilemas. Las estructuras organizativas que se diseñaron para un sistema tecnológico previo pierden acomodo dentro de un contexto de redes, ya que a nivel interno los datos y la información se pueden transmitir de manera rápida y barata, generar conocimiento potencialmente utilizable por todos los empleados de la organización o permitir su acceso ilimitado desde cualquier terminal conectado a la Red (Criado, 2004). En la esfera externa, las Administraciones Públicas disponen de un nuevo canal de comunicación de información y conocimiento potencialmente accesible desde cualquier lugar y en cualquier momento, además, existe un nuevo medio para la prestación de servicios públicos digitales, que no requieren la presencia física de los ciudadanos en las organizaciones (Silcock, 2001).

2.2. La Gestión del Conocimiento en el Sector Público. El Caso de la Web

Dentro del marco de la Sociedad de la Información y del Conocimiento, las organizaciones perciben que sus activos intangibles son los que aportan verdadero valor. De manera sucinta, los activos intangibles son una serie de recursos que pertenecen a la organización, a pesar de que no están valorados

desde un punto de vista contable, es decir, un activo intangible es todo aquello que una organización utiliza para crear valor, pero que no contabiliza económicamente (Nonaka y Takeuchi, 1995). La GdC supone la gestión de los activos intangibles que generan valor para la organización, ya que la mayoría de estos intangibles tienen que ver con procesos relacionados de una u otra forma con la captación, estructuración y transmisión de conocimiento (Davenport y Prusak, 2000). En el momento actual, la gestión de sus activos intangibles, pero especialmente el conocimiento, es uno de los retos más destacados a los que tienen que hacer frente las Administraciones Públicas. Lo anterior es así, sobre todo, como consecuencia de la diversidad de fuentes de datos e información que existen dentro de las organizaciones y, de manera concreta, como resultado de la preeminencia de la información vinculada a las páginas web y los servicios disponibles a través de medios telemáticos (Harman y Brelade, 2001).

Las Administraciones Públicas destacan por la gran cantidad de datos e información que disponen. Si a los datos e información se añade el conocimiento, se puede decir que los datos están localizados en el mundo y el conocimiento está localizado en agentes (personas, organizaciones, etc.), mientras que la información adopta un papel mediador entre ambos conceptos (Milner, 2000). En realidad, lo que fluye entre agentes distintos nunca es conocimiento como tal, sino datos o información, por lo que es posible aproximar el conocimiento de dos agentes que comparten los mismos datos, pero debido a sus experiencias anteriores y a las diferencias en el modo de procesar los datos (modelos mentales, modelos organizativos, etc.), nunca tendrán las mismas tendencias para la acción, ni estados idénticos de conocimiento. En definitiva, los datos, una vez asociados a un objeto y estructurados se convierten en información, mientras que la información asociada a un contexto y a una experiencia se convierte en conocimiento. La GdC se ha convertido en una de las preocupaciones más recientes de las Administraciones Públicas (Lenk y Traunmüller, 2001).

Como ejemplo concreto, la Web fue concebida a finales de los ochenta con el objetivo de compartir información entre distintos grupos de investigación, y su sorprendente crecimiento a lo largo de los noventa es una prueba irrefutable de su éxito. El considerable tamaño que la Web ha alcanzado genera todo un conjunto de nuevos problemas que ponen de manifiesto sus limitaciones para transformar un enorme caudal de información en conocimiento directamente utilizable por las Administraciones Públicas y los ciudadanos vía telemática (Scavo y Yuhang, 1999).

Por un lado, el volumen de la información almacenada en la Web es demasiado grande y está deficientemente estructurada como para permitir a las organizaciones públicas una gestión eficiente de la misma con las técnicas tradicionales. En ese sentido, los sistemas de búsqueda han evolucionado desde sistemas de fuerza bruta a otros modelos más inteligentes que se basan en el

conocimiento (Zaïane, 1998), pero hay una frontera en la eficacia de la recuperación de conocimiento de difícil superación, dada la naturaleza de las tecnologías tradicionales utilizadas en la Web. En su estado actual de desarrollo, la Web mantiene un gigantesco volumen de información dentro de documentos HTML, pero esta información carece de contexto y por lo tanto no se la puede considerar como conocimiento. Los enlaces, que son la base del hipertexto, no proporcionan dicho contexto, ya que no modifican la naturaleza de la significación de los datos: en términos prácticos, sólo expanden el tamaño de la página web (Berners-Lee, 1999).

Por otro lado, la Web está orientada a la información, sin embargo, cada vez es más utilizada para acceder a servicios, como en el caso de los portales de las Administraciones Públicas (Aibar, 2004). La interfaz utilizada para proveer servicios en los portales del sector público es técnicamente un formulario web que el ciudadano ha de cumplimentar y, por lo tanto, dos servicios muy similares pueden tener interfaces muy diferentes. En este caso, aparece la cuestión de integrar distintos servicios con distintas interfaces, un problema conceptualmente similar al descrito anteriormente con la integración de la información. Estos son las dos cuestiones teóricas que se abordan a continuación en esta ponencia.

3. Marco tecnológico

El enfoque original de las tecnologías que estaban detrás de la Web era común: su orientación a la utilización por seres humanos. Cuando la Web sólo disponía de un número limitado de usuarios, generalmente pertenecientes a la comunidad científica y académica, esta aproximación para la interacción humano-máquina era factible, de hecho, la expansión de la Web a mediados de la década de 1990 muestra el éxito cosechado por este enfoque en esa primera etapa. Sin embargo, la Web ha sido víctima de su propio éxito, y el enorme tamaño que tiene en la actualidad, con el volumen de información que maneja, plantea toda una nueva serie de problemas organizativos a los cuales los ingenieros han de enfrentarse. Además, cada vez más la Web es utilizada no para acceder a información, sino a servicios, en otras palabras, la e-Administración es un magnífico ejemplo de la evolución en el uso de la Web, ya que los ciudadanos no sólo consultan los horarios para realizar un trámite, sino que desean completar servicios en línea. A pesar de que la Web ha sabido adaptarse más que adecuadamente a este nuevo uso que los usuarios reclaman, no deja de ser una evidencia que está orientada a la información, no a los servicios, y que la naturaleza de ambas orientaciones es radicalmente diferente. Teniendo en cuenta esto, el paso siguiente consiste en facilitar la interacción máquina a máquina tanto para el acceso a la información como a los servicios dentro de la Web. Dos tecnologías han sido creadas para conseguirlo: la WS y los SW.

3.1. La Web Semántica

Un ejemplo claro de cómo las tecnologías existentes actualmente en la Web son insuficientes para satisfacer ciertas demandas lo encontramos en los motores de búsqueda. A pesar de poder utilizar una sofisticada tecnología, la búsqueda de información puede ser un proceso frustrante ya que el buscador no ‘conoce’ el significado de la palabra, sólo ve una sucesión de letras. Esto es consecuencia de que la información en la Web está escasamente estructurada, no proporciona información sobre su significado y sí mucha sobre cómo se representa. El proceso sería mucho más fructífero si los motores de búsqueda supieran qué se está buscando: un concepto filosófico, una persona, una Administración, un servicio, etc. Este enfoque requiere que la información en la Web esté estructurada en función de lo que significa, no de cómo se representa, de manera que permitiría todo un nuevo abanico de posibilidades antes no imaginadas en la Web, es decir, facilitaría una nueva experiencia de los ciudadanos con las Administraciones Públicas e incluso de éstas entre ellas mismas y otras organizaciones.

Este es el objetivo de la WS, que nació de una concepción de lo que debe ser la Web en el futuro, tal y como lo plantea su propio inventor. Según esta concepción, la Web es una red, no de información como en la actualidad, sino de conocimiento (Berners-Lee *et alia*, 2001). Las tecnologías base que posibilitan esta transformación, RDF (*Resource Description Framework*) (Miller y Manola, 2004) y OWL (*Web Ontology Language*) (McGuinness y van Harmelen, 2004), ya son un estándar y todo está preparado para que empiece la evolución de la Web hacia la WS.

Las tecnologías de la WS harán que los individuos no necesiten interactuar directamente con la Web, sino que este proceso se automatice mediante un agente *software* personal que buscará en ella para cumplir un objetivo impuesto por el usuario (Hendler, 2001). Por ejemplo, los ciudadanos podrían encontrar de manera más sencilla la sucesión de procesos que hacen falta para dar de alta un negocio, cuadrar en la agenda cuándo realizar dichos trámites, después podrían examinar los horarios de atención al ciudadano de las distintas organizaciones involucradas y contrastarlo con los horarios de transporte público, todo a través de los sitios web de distintas Administraciones Públicas y mediante un proceso completamente automatizado. El usuario recibiría un listado en el que se incluiría cuándo ir a cada Administración, para realizar qué tipo trámite, cómo ir en transporte público y sabiendo que esa planificación sería la óptima.

3.2. Servicios Web

Los SW son esencialmente una tecnología (o conjunto de tecnologías) que permiten la publicación y consumo de servicios entre máquinas utilizando protocolos basados en XML, el más importante de los cuales es SOAP (*Simple*

Object Access Protocol) (Mitra, 2003). A través de los SW se pueden ofrecer servicios independientemente de las arquitecturas y lenguajes de programación utilizados con un nivel de acoplamiento muy bajo entre clientes y servidores.

La aplicación más inmediata de los SW podría ser la modularización de un sitio en Internet. Desde esta perspectiva, se descomponen una serie de servicios que pueden ser reutilizados, por ejemplo, por el *software* de gestión que utilice el personal de una Administración Pública, con una disminución de costes y una mayor flexibilidad. Esto podría incluir los distintos trámites que un ciudadano puede realizar, como son los censos, el pago de tributos o la solicitud de participación en una oferta de empleo público.

Al margen de la mera invocación de servicios, los SW involucran cada vez más todo un conjunto de tecnologías que ofrecen cada día funciones más avanzadas, proporcionando capacidad de transacciones seguras, introspección de servicios, coreografías, orquestación de servicios y un largo etcétera que abre un amplio abanico de posibles usos. Los SW ofrecen un marco para el desarrollo de aplicaciones para las Administraciones Públicas mucho más flexibles, con una capacidad sin precedentes para evolucionar dentro del cambiante entorno político-administrativo y a unos costes mucho menores. Utilizando todo el potencial que ofrecen los SW, puede crearse un *software* de gestión con una gran flexibilidad: podrían transformarse los procesos internos de una organización afectando de manera sustancial la lógica que maneja el *software* y el sistema podría adaptarse a dichos cambios sin necesidad de cambiar los programas.

4. Descripción del Sistema

Las Administraciones Públicas manejan una importante cantidad de información que se almacena en los medios más dispares, desde gestores de bases de datos o páginas web hasta archivos locales generados por un procesador de textos, con formatos a veces escasamente compatibles. Esta diversidad de formatos genera un importante problema cuando se necesita abstraer los recursos particulares y ver la información generada por una organización pública como un todo. Ante esta situación de incompatibilidad de la información, generalmente se crean *middlewares* específicos de desarrollo caro y poco reutilizables en otras situaciones. Por otra parte, el tamaño de las Administraciones Públicas puede variar de una manera sustancial, desde las más pequeñas Administraciones Locales con estructuras poco cambiantes y escaso volumen de información, hasta grandes Administraciones estatales o autonómicas, con múltiples departamentos, una estructura flexible y una gran cantidad de sistemas de información diferentes. Por lo tanto, la manera de abordar técnicamente esta problemática en una u otra organización varía también sustancialmente, aunque resulta deseable que el sistema se pueda adaptar a cualquiera de los entornos descritos. Además, el

personal en las Administraciones Públicas dedicado a las funciones TIC suele ser escaso y saturado en las tareas que desempeña, por lo que no se puede plantear a priori la existencia de personas dedicadas exclusivamente a tareas de integración de sistemas de información. Ahora se pasa a plantear el sistema *Searchy* que pretende ofrecer determinadas soluciones tecnológicas a los problemas propuestos.

4.1. Requisitos de la Aplicación

Teniendo en cuenta el marco teórico y tecnológico, se plantearon una serie de objetivos a alcanzar con el nuevo sistema durante la fase de diseño:

- Proporcionar un servicio de búsqueda homogéneo, independientemente del formato de la información.
- Abstracter la representación de la información, expresándola semánticamente.
- Extensibilidad, ya que los formatos de información accedidos son, a priori, desconocidos.
- Integración con las nuevas tecnologías que nacen en la Web, en particular con los SW para el acceso al servicio y la WS para la representación de la información.
- Altamente distribuido, de manera que pueda escalar adecuadamente con el tamaño de la organización que lo adopte, desde una pequeña organización hasta una de grandes dimensiones.
- Facilidad de uso, para ello la aplicación debe ser completamente autónoma, lejos de complejos contenedores de aplicaciones, muy comunes en el mundo de los servicios web, repletos de complejas funciones a la par que escasamente útiles.
- Portable, de manera que la aplicación estaría programada en Java.

Con estos objetivos de diseño se creó *Searchy*, un motor de búsqueda altamente distribuido de concepto muy avanzado, que actualmente se encuentra en estado de arte respecto a los motores de búsqueda. *Searchy* nació de la colaboración entre *RedIRIS* y el *Departamento de Automática* de la *Universidad de Alcalá de Henares*.

4.2. La Aplicación desde una Perspectiva General

En esencia, *Searchy* es un conjunto de agentes independientes que trabajan cooperativamente. En función de las necesidades concretas del entorno, *Searchy* puede adaptarse con una gran flexibilidad admitiendo muchas maneras diferentes de utilización, de forma que teniendo en cuenta las particularidades de las Administraciones Públicas que lo implementen, puede utilizarse desde un único agente, hasta un número de agentes, en principio, ilimitado⁴.

En este contexto, un agente es un servidor en el sentido tradicional, que expone sus servicios como SW, y cualquier cliente puede acceder a él por medio de SOAP. La utilización de SOAP garantiza una elevada interoperabilidad y proporciona un grado de libertad importante en cuanto al entorno en el que se utiliza. De esta manera, *Searchy* puede ser empleado en entornos web dentro de una página JSP o PHP, o fuera de ellos desde una aplicación gráfica pesada programada en C++, Java o VB.NET. Existe un amplio abanico de lenguajes de programación que actualmente soportan SOAP. Las potencialidades están abiertas en ese sentido.

Funcionalmente un agente está dividido en tres componentes:

a) *Núcleo*. Es la parte común a todos los agentes. En general, el núcleo desempeña todas las funciones que no son específicas de un proveedor.

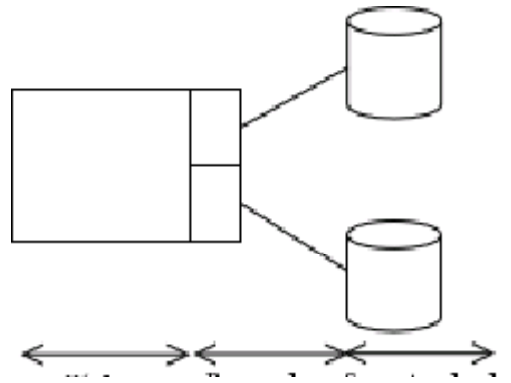
b) *Proveedor*. Maneja el acceso a un soporte de datos particular y genera la metainformación a partir de los datos obtenidos de dicho sistema de información. Añadiendo proveedores, *Searchy* puede manejar nuevos soportes de datos.

c) *Soporte de datos*. Es el elemento de información que es accedido actualmente. Puede ser un gestor de bases de datos, un servidor de directorio, un servidor HTTP, ficheros locales o incluso otro agente *Searchy*.

En la Figura 1 pueden visualizarse estos conceptos en un único agente que dispone de un núcleo y dos proveedores que acceden a dos soportes de datos. Cabe destacar que un agente puede tener cuantos proveedores sean necesarios y que a su vez puede comunicarse con otros agentes mediante un proveedor específico.

Figura 1. Agente Searchy Básico

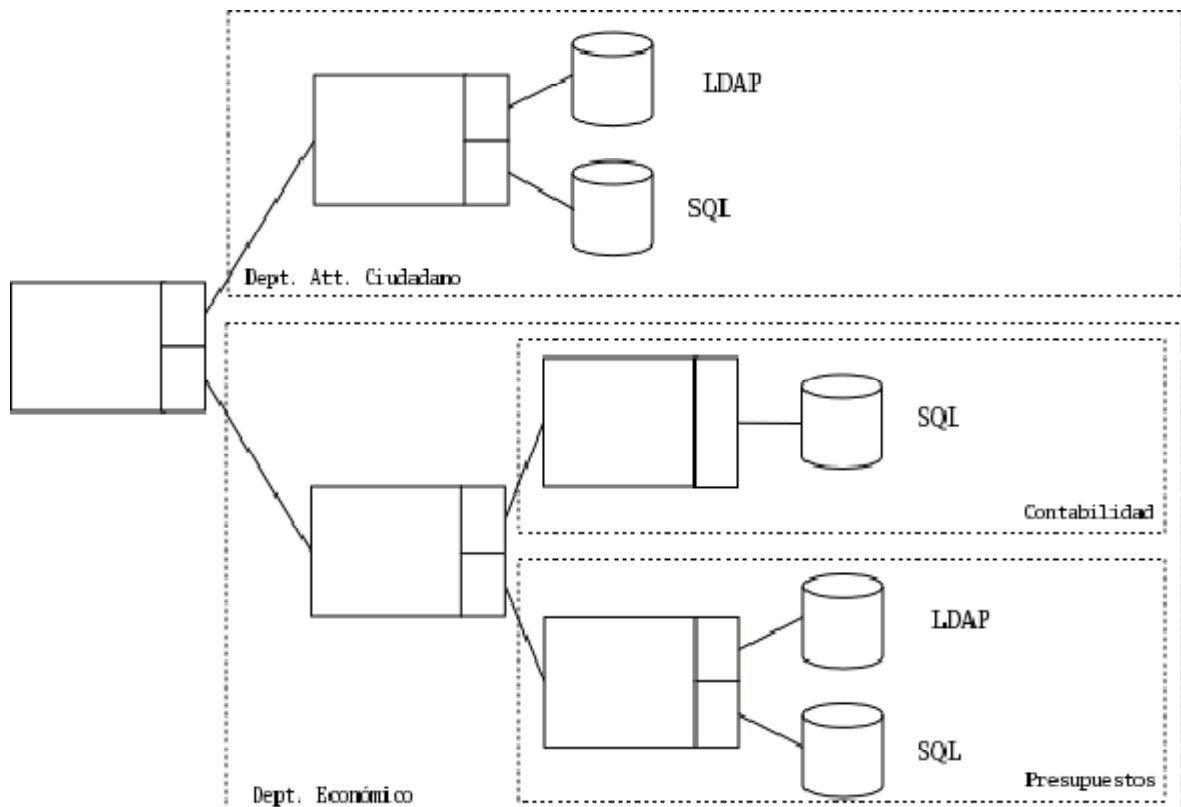
⁴ En realidad este número viene limitado por consideraciones de desempeño. En función de cómo se configure *Searchy*, los agentes pueden trabajar en serie, por lo que el tiempo de respuesta necesariamente será proporcional al número de agentes en serie que se utilicen.



Fuente: elaboración propia

Un caso práctico más realista puede encontrarse en la Figura 2, ya que la filosofía en la que está realizado *Searchy* lo hace sumamente flexible y permite que se adapte perfectamente a cualquier organización pública. La Figura 2 muestra una típica organización, en la que cada departamento o unidad es responsable de administrar un agente que ofrece servicios de búsqueda al resto de la organización. A su vez, un tercer agente realiza la integración de los distintos departamentos, ofreciendo una única interfaz para dar acceso a todos los recursos de información de la organización, que puede realizarse tanto externa como internamente.

Figura 2. Ejemplo de Aplicación de Searchy en una Administración Pública



Fuente: elaboración propia

Puede ser necesario, bien por las dimensiones de la Administración Pública, bien por consideraciones legales, organizativas o de otra índole, subdividir un departamento, por ejemplo, en unidades administrativas de menor ámbito, y que cada subunidad departamental mantenga un agente *Searchy* que proporcione servicios de búsqueda y localización a la información generada.

En la actualidad, *Searchy* tiene implementados cuatro proveedores que permiten el acceso a diversos soportes de datos:

- * **Proveedor SQL.** Proporciona acceso a bases de datos relacionales a través de SQL. Está implementado en JDBC, por lo que soporta cualquier base de datos para el que exista un *driver* JDBC.
- * **Proveedor LDAP.** Da soporte a servicios de directorio LDAP.
- * **Proveedor Google.** *Searchy*, utilizando el API ofrecido por Google, puede realizar búsquedas en el área de todo Internet. Los resultados serían los mismos que los obtenidos en una búsqueda manual en el sitio web de Google.
- * **Proveedor Searchy.** Permite que un agente se comuniquen con un segundo agente, lo que es la base para la distribución de *Searchy*.

El modo de describir los recursos encontrados es fundamental para el correcto desempeño de *Searchy* y es aquí donde las tecnologías de la WS entran en juego. Los agentes, una vez que localizan un recurso, obtienen metainformación y la expresan utilizando la tecnología base de la WS, RDF (Miller, 2004), junto con un vocabulario específicamente creado para la descripción de recursos, *Dublin Core* (DCMI, 2004), que define una serie de atributos como título, autor o materia.

La extracción de metainformación es un proceso que implementa el proveedor y que, evidentemente, es muy dependiente del soporte de la información. Por lo general se utiliza una proyección directa de campos de datos en metainformación, por ejemplo, un determinado campo en una entrada de una base de datos podría servir como título para ese recurso concreto.

En suma, *Searchy* es un sistema de búsqueda altamente distribuido, escalable y modular que facilita la federación de servicios de búsqueda, si bien la generalidad de su diseño lo hace apto para ser utilizado en entornos no previstos inicialmente. La flexibilidad de *Searchy* permite que solucione una amplia variedad de problemas en diversas situaciones a las que hacen frente organizaciones complejas para gestionar su conocimiento. Desde el plano de las Administraciones Públicas, resulta especialmente interesante ya que ofrece la oportunidad de integrar la enorme cantidad de información y datos disponibles en los portales web de una manera más adecuada a las nuevas necesidades de los ciudadanos.

5. Conclusiones

Searchy es un sistema que abre nuevas puertas a las Administraciones Públicas para hacer frente a la problemática de la abundante información de la que disponen desde la perspectiva de la GdC. Este problema se ha agravado como consecuencia de la expansión de la Web, ya que este nuevo medio de comunicación con los ciudadanos supone mayor accesibilidad a los datos e información de la organización, sobre todo, de cara a la prestación de servicios públicos electrónicos. No obstante, todavía pocas Administraciones Públicas son conscientes de la realidad de este problema y de las nuevas soluciones tecnológicas disponibles para hacer frente a la recuperación de información distribuida en toda la organización, sobre todo, las vinculadas a la WS y los SW.

Esta aplicación está disponible gratuitamente y se encuentra en pleno proceso de desarrollo (Searchy, 2004). En un futuro breve *Searchy* tendrá soporte pleno para manejar RDF, con lo que se transformará desde un motor de búsqueda a un motor de búsqueda y traducción semántica de información, proporcionará una interfaz para realizar búsquedas complejas basadas en RDQL (*RDF Data Query Language*) (Seaborne, 2004) y se desarrollarán nuevos proveedores que darán soporte a arquitecturas SOA, servidores HTTP, FTP y archivos locales en distintos

formatos como OpenOffice y MS Word, entre otros muchos. Desde luego, se espera que las Administraciones Públicas puedan explotar en toda su extensión las posibilidades que ofrece esta aplicación para organizaciones intensivas en conocimiento.

Bibliografía

- Aibar, Eduard. 2004. *Govern Electrònic i Serveis Públics: un Estudi de Cas sobre el Portal Interadministratiu CAT365*. Barcelona: Universitat Oberta de Catalunya.
- Bellamy, Christine, y John A. Taylor. 1998. *Governing in the Information Age*. Buckingham: Open University Press.
- Berners-Lee, Tim. 1999. *Weaving the Web. The Original Design and Ultimate Destiny of the World Wide Web, by its Inventor*. New York: HarperCollins.
- Berners-Lee, Tim, James Hendler y Ora Lassila. 2001. The Semantic Web, *Scientific American* 284 (5):28-31.
- Criado, J. Ignacio. 2004. Entre Sueños Utópicos y Visiones Pesimistas. Un Análisis de la Administración Electrónica Local en España. *Gestión y Política Pública* 13 (2).
- Davenport, Thomas H. y Prusak, Laurence. 2000. *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press.
- DCMI Usage Board. 2004. DCMI Metadata Terms.
Disponible: <http://www.dublincore.org/documents/dcmi-terms>. Fecha de acceso: 27 de junio de 2004.
- Harman, Chris y Sue Brelade. 2001. Knowledge, E-Government and the Citizen. A Report into the Role of Knowledge Management in e-Government. *Knowledge Management Review* 4 (3):18-23.
- Hendler, James. 2001. Agents and the semantic web. *IEEE Intelligent Systems Journal*, 16(2):30-37, Marzo-Abril.
- Lenk, Klaus y Roland Traunmüller. 2001. Broadening the Concept of Electronic Government. En *Designing E-Government. On the Crossroads of Technological Innovation and Institutional Change*, ed. C. Prins. The Hague: Kluwer Law International.
- McGuinness, Deborah L. y Frank van Harmelen. 2004. OWL Web Ontology Language Overview, Recomendación W3C, 2004.
- Miller, Eric y Frank Manola. 2004. (eds.). RDF Primer, Recomendación W3C. 10 de Febrero del 2004.
- Milner, Eileen M. 2000. *Managing Information and Knowledge in the Public Sector*. London: Routledge.
- Mitra, Nilo, 2003. (ed.). SOAP Version 1.2 Part 0: Primer, Recomendación W3C.
- Nonaka, Ikujiro y Hirotaka Takeuchi. 1995. *The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford: Oxford University Press.
- Scavo, Carmine, y Shi Yuhang. 1999. World Wide Páginas web Design and Use in Public Management. En *Information Technology and Computer Applications in Public Administration: Issues and Trends*, ed. G. D. Garson. Hershey: Publishing Idea Group.
- Seaborne, Andy. 2004. RDQL: A Query Language for RDF.
- Searchy. 2004. Página web del proyecto Searchy.
Disponible: <http://sourceforge.net/projects/jsearchy>. Fecha de acceso: 26 de junio de 2004.
- Seneviratne, Sonal J. 1999. Information Technology and Organizational Change in the Public Sector. En *Information Technology and Computer Applications in Public Administration: Issues and Trends*, ed. G. D. Garson. Hershey: Idea Group Publishing.
- Silcock, Rachel. 2001. What Is e-Government? *Parliamentary Affairs* 54 (1):88-101
- Zaiiane, Osmar R. 1998. *From Resource Discovery to Knowledge Discovery on the Internet*. Simon Fraser University, no. TR 1998-13.

