

M.1. Software para el acceso a la información. Situación del mercado en 2009

Por Ricardo Eíto-Brun

7 enero 2010

Eíto-Brun, Ricardo. "Software para el acceso a la información. Situación del mercado en 2009".
Anuario ThinkEPI, 2010, v. 4, pp. 300-303



Resumen: Se resumen las conclusiones del estudio publicado en septiembre de 2009 por la consultora norteamericana de TI Gartner Group, sobre el software para acceso a la información. Con este término se reúnen aplicaciones de distintos proveedores como Microsoft, Zylab, Oracle, Autonomy, Google, etc., que ofertan soluciones técnicas para la recuperación, análisis y explotación de información estructurada y no estructurada. Se describen las características del llamado "cuadrante mágico", representación usada por Gartner para presentar el posicionamiento de los proveedores en el mercado atendiendo a dos criterios: su visión y su capacidad de ejecución.

Palabras clave: Recuperación de información, Software, Proveedores de software.

Title: *Software for information access. 2009 market situation*

Abstract: The findings of the "information access software" analyses released in September 2009 by the IT consultancy company Gartner Group are summarized. This term refers to various software applications from publishers like Microsoft, Zylab, Oracle, Autonomy, Google, etc., used for both structured and unstructured data and for information retrieval and analysis. The contribution also describes the characteristics of the "magic quadrant", a visual representation offered by Gartner to reflect the position in the market of the software publishers considering their vision and their ability to execute.

Keywords: Information retrieval, Software, Software providers.

EN SEPTIEMBRE DE 2009 la consultora norteamericana Gartner Group publicó su "Magic quadrant" sobre las empresas dedicadas a la comercialización de software para el acceso a la información.

Esta consultora, con sede en Stamford (Estados Unidos) y amplia representación y prestigio internacional, publica anualmente estudios para distintos tipos de aplicaciones software en los que establece una valoración de los principales fabricantes. Es interesante revisar estos informes para conocer la situación de los actores, qué nuevas empresas están demostrando su potencial y cuál ha sido la evolución general del mercado.

Dentro de la categoría "tecnologías para el acceso a la información" el estudio incluye "fabricantes de software que ofrecen funciones de recuperación de información, búsqueda federada, clasificación y categorización de contenidos, técnicas de *clustering*, extracción de datos y hechos a partir de textos, creación de taxonomías y visualización de información".

Así, bajo esta designación se han agrupado empresas que ofrecen software capaz de extraer y recuperar información de colecciones de docu-

mentos, bases de datos relacionales, sistemas de gestión de contenidos y de otros sistemas de gestión como los ERP (*enterprise resource planning*) o CRM (*customer relationship management*) utilizados para la gestión de las distintas actividades de las empresas.

"El cuadrante de Gartner no recoge todos los fabricantes, por lo que debe evitarse dar una valoración negativa a las empresas que no están presentes"

Obviamente el cuadrante de Gartner no recoge una categorización y análisis exhaustivo de todos los fabricantes existentes, por lo que debe evitarse dar una valoración negativa a aquellas empresas que no están presentes. Para la inclusión de fabricantes en el estudio se han considerado los siguientes criterios: presencia internacional en al menos dos de estos mercados (Europa, África y Oriente Medio, América y/o Asia/Pacífico), bene-

ficios superiores a los diez millones de dólares durante 2008, y comercialización de su software de recuperación y acceso a la información de forma independiente a otras aplicaciones para la gestión de documentos y contenidos.

Principales conclusiones del informe

El estudio pone de manifiesto la necesidad de ir más allá de la mera recuperación de información y del protagonismo cada vez mayor de las tecnologías orientadas a facilitar el análisis y la interpretación de volúmenes crecientes de datos no estructurados (es decir, documentos). Tal y como se indica en el texto, la capacidad de ofrecer este tipo de funciones para el análisis de contenidos y la extracción de información ha sido un aspecto crítico en la evaluación de la capacidad innovadora de estas empresas.

El informe también destaca cómo algunas empresas tienden a orientar su actividad a la resolución de problemas en el acceso a información en áreas de actividad específicas (concretamente se mencionan los casos de la información legal y editorial).

Las previsiones para este tipo de aplicaciones son positivas: en 2008 la venta de licencias y mantenimiento de software de acceso a la información supuso una cantidad global próxima a 1.100 millones de US\$, y se anticipa un crecimiento del 11,7% anual hasta alcanzar una cifra próxima a los 2.000 millones en 2013.

“La situación de liderazgo también se asocia con la situación financiera de las empresas, en cuanto ésta es una garantía de su continuidad en el mercado”

El estudio también demuestra la consolidación de algunos fabricantes en las posiciones de liderazgo. Esto se presenta como consecuencia de distintos movimientos (fusiones y adquisiciones) completados en los últimos años. Concretamente se señala la compra de *Verity* e *Interwoven* por *Autonomy* (a finales de 2005 y en enero de 2009, respectivamente); y la de *Fast Search & Transfer* por parte de *Microsoft* en abril de 2008, lo que le ha permitido ampliar notablemente su oferta de motores de indexación y recuperación integrados en torno al producto *SharePoint*.

En relación con el código abierto, el informe hace mención al programa *Lucene*, aunque se señala que –en esta área en concreto– las capa-

idades de los programas libres se encuentran aún alejadas de las que ofrecen los sistemas comerciales.

El cuadrante mágico

El informe presenta visualmente un esquema donde los fabricantes se posicionan en un eje cartesiano: el eje vertical corresponde a la capacidad de ejecución (*ability to execute*), y el horizontal a la completitud de la visión del fabricante (*completeness of vision*).

Los criterios para evaluar la posición de las empresas en cada eje tienen en cuenta distintos aspectos.

La capacidad de ejecución incluye elementos como la versatilidad del producto y su capacidad de interactuar con diferentes repositorios de contenido, su capacidad de gestionar volúmenes crecientes de datos y de usuarios (escalabilidad), datos financieros de la empresa, ventas y políticas de precios, presencia geográfica en su actividad comercial y de soporte, opiniones de usuarios, etc.

La completitud de la visión es el resultado de evaluar su comprensión del mercado, sus estrategias de marketing y de ventas, y aspectos más concretos relacionados con la innovación como la capacidad de trabajar con documentos no textuales, la inclusión de funciones para el análisis de contenido, o el uso de elementos adicionales al propio texto de los documentos para evaluar su relevancia (se menciona el análisis de citas y de referencias cruzadas).

La intersección de estas dos dimensiones –capacidad de ejecución y completitud de la visión– permite diferenciar cuatro áreas o grupos:



Magic quadrant

a) Líderes; donde se sitúan aquellas empresas que han unido la amplitud de su visión (comprensión de las tecnologías, de las necesidades de los usuarios, etc.), con su capacidad de diseñar productos software que den respuesta a estas necesidades. La situación de liderazgo también se asocia con la situación financiera de las empresas, en cuanto ésta es una garantía de su continuidad en el mercado.

b) Retadores (*challengers*); donde están las empresas que no tienen una visión tan desarrollada de las necesidades globales del mercado, pero sí una alta capacidad de ejecución y que cuentan con los recursos necesarios.

c) Visionarios; empresas que han identificado áreas y necesidades con una gran nivel de detalle, pero que no han llegado a traducir esta visión en una posición de liderazgo.

d) Jugadores de nicho (*niche players*); donde se situarían las empresas con una visión menos completa y que no han demostrado aún la capacidad de ejecución de otros competidores. Este grupo reúne a los fabricantes que ofrecen productos capaces de satisfacer necesidades muy específicas, pero que carecen de los recursos de los líderes y de los retadores para afrontar una tipología de proyectos más amplia.

Inicialmente, en el modelo del cuadrante cuanto más situada esté una organización hacia arriba y hacia la derecha, mejor posicionada estará. Esto indica que tiene una gran comprensión de la evolución y de las posibilidades de la tecnología, de los requisitos presentes y futuros de la comunidad de usuarios, y una gran capacidad para desarrollar productos acordes con esto.

No obstante, no debemos interpretar que las empresas situadas en otros cuadrantes sean necesariamente peores u ofrezcan productos de menor calidad. Por ejemplo, en el eje correspondiente a la completitud de la visión se puede situar a la izquierda a una empresa especializada en una solución muy concreta (por ejemplo, visualización de datos), ya que únicamente centra su actividad en una de las tecnologías para el acceso a la información que analiza *Gartner*.

¿Quiénes están situados en cada cuadrante?

El cuadrante de los líderes del año 2009 está



Autonomy (fuente: <http://www.autonomy.com/>)

“Las capacidades de los programas de código abierto se encuentran aún alejadas de las que ofrecen los sistemas comerciales”

ocupado por *Microsoft* (con un lugar destacado), *Autonomy* y la empresa holandesa *Zylab* (de la que hemos hablado en números pasados de la revista EPI, y que en 2002 colaboró con nuestra editorial para la publicación del cd-rom con los cien primeros números de EPI). El informe señala que *Autonomy* (que sería el principal fabricante por volumen de facturación) y *Microsoft*, son líderes en el rango de proyectos para instalaciones con presupuestos menores de 75.000 US\$ en los dos primeros años.

Dentro de este cuadrante, en una posición menos destacada encontramos a *Endeca*, *Recom-mind* y *Vivisimo*. Éstas son empresas con una amplia visión, pero con menor capacidad de ejecución que *Microsoft*, *Zylab* o *Autonomy*. De *Endeca* se destaca su capacidad de explotar de forma conjunta datos estructurados (bases de datos relacionales) y no estructurados (documentos).

En el cuadrante de los retadores destacan *Google*, *Oracle* e *IBM*, seguidas por *Exalead*, *Kazeon* (comprada este mismo año por *EMC*) e *Isys Search Software*. *IBM*, que el año pasado estaba situada en el cuadrante de los líderes, se ha visto ligeramente desplazada este año, y el informe destaca aspectos positivos como su compromiso con el sistema de indexación libre *Lucene*, su capacidad de gestionar objetos multimedia y de implementar tecnologías relacionadas con la web 2.0.

Finalmente, el grupo de los jugadores de nicho encontramos a *Sinequa*, *Expert System* y *Fabasoft-Mindbreeze*.



Microsoft (fuente: <http://www.microsoft.com/>)

Cambios respecto 2008

Si comparamos el cuadrante de 2009 con el de 2008, podemos apreciar los siguientes cambios:

- Microsoft ha ampliado su distancia respecto a *Autonomy*, al asignársele una mayor capacidad de ejecución, lo que hace más manifiesto su liderazgo.

- *Zylab* se ha situado en una posición superior, aumentando su capacidad de ejecución y pasando a ocupar un lugar más destacado dentro de los líderes.

- *Recommind* y *Exalead* también han aumentado su capacidad de ejecución, si bien la segunda se ha visto desplazada hacia la izquierda, lo que sería señal de una visión menos completa.

- *OpenText* no ha sido incluida en el cuadrante del 2009, cuando sí lo estuvo en el del 2008 con una posición relevante dentro del recuadro correspondiente a los retadores. El motivo por el que se le ha excluido es que ya no comercializa sus soluciones de recuperación y búsqueda de forma separada, sino que las ofrece integradas como parte de sus aplicaciones para la gestión documental y de contenidos.

- *Google*, *Oracle*, *IBM*, *Expert System*, *Isys Search Software*, *Fabasoftware/Mindbreeze*, *Endeca* y *Vivísimo* mantienen posiciones similares a las del año anterior.

- *Sinequa* (empresa francesa) es la única compañía que se ha incorporado al cuadrante

te de 2009. Se destaca su mayor presencia en el mercado norteamericano.

“El cuadrante de los líderes del año 2009 está ocupado por Microsoft (con un lugar destacado), Autonomy y la empresa holandesa Zylab”

Sin duda estos informes de *Gartner* ofrecen una más que interesante visión del mercado del software para la gestión de información, y nos permiten apreciar los cambios y la evolución global de un mercado en el que cabe destacar:

- la necesidad de contar con soluciones avanzadas para el análisis y explotación de la información, que van mucho más allá de la recuperación textual tradicional, y

- la consolidación de posiciones en los cuadrantes correspondientes a los líderes y a los “retadores”, con el papel predominante de *Microsoft*.

Referencias

- Gartner*. <http://www.gartner.com>
- Gartner*. *Magic quadrant for information access technology*. <http://www.gartner.com/technology/media-products/reprints/endeca/169927.html>



Zylab (fuente: www.zylab.es)