

# Informe APEI sobre vigilancia tecnológica

Informe APEI 4

por Lara Rey Vázquez

2009



*A Héctor y a César por los minutos robados.*

Esta publicación ha contado con el patrocinio de:



# Informe APEI sobre vigilancia tecnológica

*por Lara Rey Vázquez*



<http://www.apeiasturias.org>  
[info@apeiasturias.org](mailto:info@apeiasturias.org)



Edición sostenible. Los informes de APEI están editados como documentos electrónicos de lectura en pantalla. Si no es necesario, no los imprimas. Si los imprimes, hazlo a doble cara.

### **Informe APEI sobre vigilancia tecnológica**

Lara Rey Vázquez

Gijón: Asociación Profesional de Especialistas en Información, 2009

Informe APEI 4 - 2009

D. L.: AS-06633-2008

ISBN: 978-84-692-7999-1

Edición: APEI

Director editorial: Raquel Lavandera Fernández

Diseño y maquetación: Digitales José & Cícero, S.L.



Esta obra está bajo una licencia Reconocimiento-No comercial-Compartir bajo la misma licencia 2.5 España de Creative Commons: <http://creativecommons.org/licenses/by-nc-sa/2.5/es>.

<b>1. Presentación</b>	<b>7</b>
Palabras previas	7
A modo de introducción	8
<b>2. Definición y conceptos</b>	<b>9</b>
¿Qué es la Vigilancia tecnológica?	9
Objetivos de la Vigilancia tecnológica	10
¿Qué se debe vigilar?	12
Clases de Vigilancia	12
Vigilancia Tradicional	12
Vigilancia Avanzada	13
<b>3. Metodología y diseño ¿por dónde empezamos?</b>	<b>15</b>
Metodología y fases	15
Proceso completo de vigilancia	17
<b>4. La información y la empresa</b>	<b>19</b>
¿Para qué sirve la Vigilancia en una empresa?	19
La información falsa en Internet	20
La Vigilancia y su ayuda a la empresa	20
La empresa y su entorno	21
Información externa (modelo de Laudon y Laudon)	22
Información interna	22
Las fuentes de información de una empresa	22
Diseño de un sistema de VT en una empresa	23
Principios básicos de un sistema de VT	24
<b>5. Las fuentes de información</b>	<b>25</b>
Conceptos previos	25
Literatura gris	25
Web Invisible	25
Internet	26
¿Cómo evaluar la fiabilidad de una web?	28
Búsquedas y buscadores: más allá de Google	29
Las búsquedas	29
Los buscadores	31
Las fuentes: Gratuitas vs Pago	33

<b>6. Las patentes</b>	<b>37</b>
Las patentes como proceso/trámite	37
Las patentes como fuentes de información	38
Patentes-fuentes de información tecnológica	41
<b>7. Portales, software, buscadores y otras herramientas</b>	<b>45</b>
Herramientas de búsqueda avanzada	45
Software de patentes	45
Software de visualización de la información	48
Agentes de búsqueda	50
Software de Vigilancia Tecnológica	51
Portales de organismos públicos que ofrece servicios de Vigilancia Tecnológica	54
Webs, blogs y otros portales interesantes	54
<b>8. Análisis de la información recuperada</b>	<b>55</b>
Informes de Vigilancia	55
Ejemplos en Internet	56
<b>9. Vigilancia tecnológica 2.0</b>	<b>57</b>
Definiciones	57
La Web 2.0 con ejemplos	57
Vigilancia 2.0	57
<b>10. Bibliografía</b>	<b>61</b>

# Presentación

## Palabras previas

Llevo casi toda mi vida laboral dedicada a las labores de Vigilancia. No tengo muy claro cuántas fuentes habré manejado en este tiempo, pero seguro que muchas (observando mi favoritos diría que más de 1000). Cuando me incorporé por primera vez al mundo laboral, estaba empezando a fraguarse la ola digital en la que hoy estamos inmersos. Mi primer correo electrónico corporativo lo tuve en el año 1998. En mi libreta de direcciones había cuatro contactos (me da la risa!!). En esa época asistí a un curso en Barcelona para aprender a manejar la base de datos Dialog, cuando todavía no era del grupo Thomson. De aquella yo no usaba Google, no tardé mucho, pero todavía no lo utilizaba. Tenía un correo en Hotmail con 2 Mb de capacidad. La conexión era por una frame-relay de 64 kb., utilizaba disquets, Windows NT Workstation, mi ordenador era un Pentium II y utilizaba como Bases de Datos Dialog, TradStat, Datastar, Ardan, Dicodi, etc..Todas por su puesto super-caras y super-restrictivas.

Así era de dura la vida. Pero a pesar de todo era un placer navegar por internet buscando fuentes nuevas con información. Y todos los días nos sorprendíamos con algo novedoso. Los organismos oficiales, las universidades, las bibliotecas, las oficinas de patentes... Se había iniciado una carrera meteórica para estar en la red de redes... facilitando sin querer la labor del documentalista y de qué manera!!!

Creo que este “tsunami” nos ha venido de perlas y creo también que ha valorizado, actualizado y puesto en el punto de mira una profesión un

poco desconocida en la sociedad (aun escucho de vez en cuando... documentalista? Y qué haces... documentales??).

La Vigilancia en sí no es nueva, la forma de hacer vigilancia es lo que ha cambiado. La forma de ofrecerla también ha cambiado. Internet es hoy por hoy una herramienta fundamental para hacer y para ofrecer Vigilancia. El resto lo pone cada uno... como en todas las profesiones.

Lo que está ocurriendo hoy ya lo conocemos todos. El término *infoxicacion* ha pasado de ser un neófito en nuestras vidas a formar parte de nuestro vocabulario. La ingente cantidad de fuentes que ofrecen contenidos a texto completo en Internet es inabarcable. El debate de ahora es el que plantea Isidro Aguillo [Isidro Aguillo ThinkEPI, enviado a IWETEL el 26-10-09] *A pesar de la creciente importancia que ha adquirido la Web, especialmente en el ámbito académico, es todavía muy frecuente observar entre profesores e investigadores actitudes despreciativas hacia sus contenidos Sin duda la proliferación de noticias negativas sobre la red en los medios de comunicación, señalando supuestos peligros, informaciones tendenciosas o flagrantes delitos ha contribuido considerablemente a esta percepción.*

*En este sentido se está abriendo una “brecha digital” entre los que prácticamente solo utilizan la red como principal fuente y citan sin ningún problema los artículos de la Wikipedia como referencia y los que no añaden la URL en una cita bibliográfica, cuyo texto completo está disponible gratuitamente en una revista electrónica o repositorio ya que al parecer ello menoscaba la “calidad” de la referencia.*

Y eso es muy importante. Estamos ante una nueva forma de difundir/añadir/modificar... la información. Las fuentes formales e informales tienen cada vez menos definidas sus fronteras. Está cambiando la forma de comunicarse, la forma de escribir, la forma de difundir la información, la forma de navegar... Proliferan las Redes Sociales, los programas en “la nube”, la sindicación de contenidos, los blogs, etc..

La tendencia es Scivee, nature network, Open Wetware, Innocentive, Google Scholar, Google Patent, etc..

Más información, más datos, más fuentes...

### A modo de introducción

De acuerdo con datos de la consultora inglesa Netcraft, durante el último año la cantidad de dominios registrados en la Web se incrementó casi en un 50% respecto de 2006. En diciembre de 2007 la cifra superó los 155 millones de sitios Web, mientras que en diciembre del año anterior alcanzaba los 105 millones. En total, los *blogs* son los que más crecieron en el último tiempo.

Pese a que los primeros acercamientos a la World Wide Web se dieron hace veinte años, la consolidación de Internet como herramienta y la explo-

sión de sitios Web en forma masiva a nivel global pueden ubicarse a partir de mediados de los '90.

De acuerdo con los datos relevados por la consultora inglesa Netcraft, 50 millones de sitios Web se sumaron durante 2007 a los que ya existían a finales de 2006 (las cifras exactas registradas en diciembre del año pasado y el anterior fueron 155.230.051 contra 105.244.649). Si se tienen en cuenta únicamente los sitios activos, el crecimiento no deja de ser sorprendente: 38% más que los existentes hasta diciembre de 2006.

Sin embargo, lo más llamativo tal vez sea que el mayor crecimiento estuvo dado esta vez por los blogs: los tres mayores proveedores\* de este tipo de sitios Web –MySpace, Live Spaces y Blogger– registraron la mitad del total de nuevos registros del año: 25 millones. Una tendencia que habla del incremento en la participación y generación de contenidos de los internautas.

Durante 2006, el récord había sido de 30 millones de sitios Web nuevos. Pero resulta interesante observar la progresiva de crecimiento del registro de sitios Web a lo largo de la última década, que de acuerdo con Netcraft fue la siguiente: en abril de 1997 existían 1 millón de sitios, en febrero de 2000 pasaron a ser 10 millones y en septiembre de ese mismo año la cifra se duplicó, con lo cual el año 2000 registró el récord histórico de crecimiento. Posteriormente la cifra fue aumentando 10 millones en cada medición, registradas sucesivamente en julio de 2001, abril de 2003, mayo de 2004, marzo de 2005, agosto de 2005, abril de 2006 y agosto de 2006.

\* Fuente: La Flecha. <http://www.laflecha.net/canales/blackhats/en-2007-el-numero-de-paginas-web-crecio-un-48>

# Definiciones y conceptos

## ¿Qué es la Vigilancia Tecnológica?

Definición según Palop y Vicente (1999, b) sobre la Vigilancia: “sistema organizado de observación y análisis del entorno, tratamiento y circulación interna de los hechos observados y posterior utilización en la empresa”.

**Definición de Vigilancia Tecnológica (VT) según la norma UNE 166006:** Proceso organizado, selectivo y sistemático, para captar **información del exterior y de la propia organización sobre ciencia y tecnología**, seleccionarla, analizarla, difundirla y comunicarla, para convertirla en conocimiento con el fin de tomar decisiones con menor riesgo y poder anticiparse a los cambios.

Otra de las definiciones que más se utiliza es: La **vigilancia consiste en realizar de forma sistemática la captura, el análisis, la difusión y la explotación de las informaciones técnicas útiles para la supervivencia y el crecimiento de la empresa.** La vigilancia debe alertar sobre cualquier innovación científica o técnica susceptible de crear oportunidades o amenazas. A las empresas les gusta saber qué pasa, no tener sorpresas, aprovechar las oportunidades, si pueden, y evidentemente combatir o hacer frente a las amenazas que puedan presentarse.

La **Inteligencia Competitiva es definida por Gibbons y Prescott (1996):** Inteligencia Competitiva es el proceso de obtención, análisis, interpretación y difusión de información de valor estratégico sobre la industria y los competidores,

que se transmite a los responsables de la toma de decisiones en el momento oportuno.

### Por lo tanto, ¿son dos conceptos similares?

*Muchas veces se confunden ambos conceptos al referirse a la misma cosa. Es obvio que los dos están estrechamente relacionados, sobre todo, porque la inteligencia competitiva es una evolución de la Vigilancia. Por ello, no son fácilmente distinguibles. De hecho, la Vigilancia está dividida exactamente en las mismas áreas que integran la inteligencia.*

*La Vigilancia nace ante la necesidad de las empresas de observar su entorno y así poder responder a determinados cambios cuando éstos se producen. La Inteligencia Competitiva, sin embargo, parte del conocimiento del entorno, lo cual implica poder adelantarse a los cambios, entendidos en ambos casos, como las amenazas y las oportunidades.*

*De esta manera, la diferencia fundamental entre las dos disciplinas es la actitud activa de la Inteligencia, para “conocer” el entorno y así anticiparse a los cambios. Ello supone no esperar a ver dónde se producen los cambios para actuar después, sino buscar activamente las oportunidades del entorno, lo que conlleva una revolución en la manera de entender todas las actividades gerenciales, comerciales y de innovación de la empresa.*

Fuente: Modelos de vigilancia tecnológica e inteligencia competitiva, ed. Bai, 2007.

Utilizarlos indistintamente no constituye una falta grave, y es habitual hacerlo así<sup>1</sup>. Pero cada vez más, los profesionales que realizan Inteligencia,

<sup>1</sup> Conferencia inaugural de los Estudios de Información y Documentación de la UOC del segundo semestre del curso 2001-2002 (15 de septiembre de 2001): [http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202\\_imp.html](http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202_imp.html) -[vínculo visitado en agosto de 2009].

marcan las distancias. La Inteligencia para ellos tiene unas implicaciones psicológicas y estratégicas que la Vigilancia, bajo su punto de vista, no tiene. Muchas definiciones de ambos conceptos coinciden casi completamente. No obstante, buscando un mayor rigor, se observa una tendencia a reservar la palabra Vigilancia a las primeras fases del proceso, es decir, a la obtención de la información y, tal vez a su procesamiento inicial, mientras que la Inteligencia se aplica sobre todo a las etapas finales, especialmente a la de análisis, que precede a la difusión y a la toma de decisiones.

La vigilancia tecnológica da paso o se incorpora dentro de la inteligencia económica, lo que supone utilizar la información obtenida relevante para la empresa de cara a la toma de decisiones acertada.

Por esta razón se afirma que la vigilancia genera inteligencia:

*La vigilancia tiene un papel de detección mientras la inteligencia competitiva tiene por misión el posicionamiento estratégico de la empresa en su entorno (Cohen).*

*La inteligencia no es sólo observación sino una práctica ofensiva y defensiva de la información. Es una herramienta que conecta el saber de la empresa con la acción (Baumard, 1991).*

Otros autores aportan nuevos matices. En Francia, por ejemplo, emplean la expresión Inteligencia Económica en lugar de Inteligencia Competitiva, que incluye el análisis de las áreas tecnológica, comercial, jurídica, financiera, etc., con un carácter más global que la Vigilancia, que se limitaría a un solo aspecto (por ejemplo, Vigilancia Tecnológica). Otros señalan que la Inteligencia aporta un producto más “elaborado” o “refinado” que la Vigilancia, que, según ellos, aportaría sólo datos en bruto. Por último, algunos señalan que el argumento de más peso en la utilización de la palabra Inteligencia en lugar de Vigilancia consiste en el creciente uso del vocablo *Intelligence* en el mundo anglosajón, que incluye aspectos relacionados con la Seguridad.

La vigilancia tecnológica y la inteligencia competitiva se orientan a proporcionar información

para la toma de decisiones, pero su objeto de análisis marca la diferencia entre ellas. Mientras la primera se enfoca en el seguimiento de la evolución de la tecnología y de sus implicaciones, la segunda lo hace en otros factores de competitividad, como los competidores actuales y potenciales, clientes, proveedores, entorno normativo, etc., y sus repercusiones en la competitividad de las empresas (regiones, sectores, clusters, cadenas productivas, entre otras).

Dado que es un proceso, se trata de un conjunto de acciones que se llevan a cabo entre varias personas, por tanto no corresponde con la idea del trabajo de una persona, sino de un equipo.

Dado que es un proceso, no hay que identificar a la vigilancia con herramientas de software, bases de datos o publicaciones especializadas. El proceso está por encima de las herramientas que se utilizan, que pueden cambiar y evolucionar con el tiempo. Es el proceso el que indica, además, el mejor uso que hay que dar a cada herramienta o a cada fuente de información.

La Vigilancia Tecnológica está basada en la captación y análisis sistemático de informaciones disponible en fuentes de información gratuitas o comerciales. No tiene nada que ver con la obtención ilícita de información mediante espionaje, copias ilegales etc.

Las tareas de la VT son pura gestión de la información. Se trata de un método claro, riguroso y neutro de alerta temprana para la dirección.

El fin de la Vigilancia Tecnológica no es captar rumores o informaciones sin contrastar. Todo lo contrario, se trata de localizar informaciones fiables de alto valor que creen nuevas oportunidades de negocio y mejoren la situación competitiva de la organización.

Hay que hacer hincapié en que el principal beneficiario de la implantación de esta metodología es la dirección de la empresa, por lo que su implantación conviene que esté impulsada por la propia dirección<sup>2</sup>.

### Objetivos de la vigilancia tecnológica

La Vigilancia Tecnológica tiene como objetivo la obtención continuada y el análisis sistemático de

<sup>2</sup> Fuente: ¿Qué es la Inteligencia Competitiva? Pere Escorsa. Madrid, 29-30 de noviembre de 2007.

información de valor estratégico sobre tecnologías y sus tendencias previsibles, para la toma de decisiones empresariales.

También es destacable el objetivo que se señala desde la Comunidad de Madrid para la vigilancia tecnológica, que es “la obtención continuada y el análisis sistemático de información de valor estratégico” (2000).

Las empresas necesitan conocer las últimas tecnologías en las que se está trabajando, así como los últimos productos que están en el mercado. En definitiva, conocer las líneas de investigación y las patentes en la disciplina correspondiente. También interesa conocer las tecnologías emergentes y aquellas que quedan obsoletas. Por tanto, saber en una etapa temprana, los avances tecnológicos en el área de interés, es un tema prioritario. La incorporación de una tecnología eficiente puede ofrecer una ventaja competitiva a la empresa.

En **segundo** lugar, interesa conocer qué hacen los competidores, en qué investigan, dónde patentan, por dónde se mueven, para intentar seguirles o adelantarlos.

La cooperación entre las diferentes universidades, centros tecnológicos, empresas, consorcios y asocia-

ciones innovadoras resulta fundamental para obtener la mejor tecnología dentro de un área determinada que permita a la empresa destacar del resto.

**Tercero**, se intentará armonizar y organizar los estudios e investigaciones sobre los distintos estándares para, de este modo, impedir el duplicado del trabajo.

Y, **finalmente**, las empresas de una tecnología determinada quieren saber quién es el líder, y en qué instituto, en qué universidad, en qué empresa está, para intentar ponerse en contacto con él para conseguir ventaja.

El objetivo es conseguir que la información adecuada esté en el momento preciso en manos de quien la necesita.

Es decir, si se tiene la información adecuada, pero no se ha establecido quien o quienes son receptores de esta información, cuando por fin lo hayamos localizado, la información ya no tendrá el valor añadido del principio. El factor “novedad” o “sorpresa” en una información es tremendamente efímero, y todo el tiempo que se tarde en transmitirla irá en detrimento de su valor.

Por lo tanto el factor información valiosa + interlocutor válido = vigilancia tecnológica.



## Preguntas a las que la VT intenta contestar

- ¿Cuáles son las principales líneas de investigación?
- ¿Qué tecnologías emergentes están apareciendo?
- ¿Qué hacen los competidores?
- ¿Quiénes son los líderes? (Centros de Investigación, equipos, personas...)

## ¿Qué se debe vigilar?

Como el propio término indica, la vigilancia supone estar informado de lo que está ocurriendo en un determinado entorno. Para el gestor de la información esto significa detectar información de interés procedente de recursos variados y dispersos, pero también seguir y analizar aquellas otras fuentes de obligada consulta que habitualmente aportan información relevante y que deben explorarse sistemáticamente.

Conocer el entorno, identificar tendencias (nuevos temas, nuevos actores, nuevos productos, nuevos procesos, etc.), riesgos, oportunidades, etc., y actuar en consecuencia, son los objetivos básicos de la vigilancia tecnológica y la inteligencia competitiva. El éxito de estos procesos depende de distintos factores pero podemos afirmar que hay dos determinantes. Por una parte, la participación de todos los miembros de una organización; por otra, la existencia de un sistema de gestión del conocimiento eficaz, entendiendo por tal un modelo de gestión de información formal e informal.

Sin embargo, en las últimas décadas esa tarea ha vivido desarrollos sustanciales, y hoy coexisten prácticas de vigilancia con objetivos y niveles heterogéneos de avance. Cuando dicha labor se hace de manera rutinaria en la empresa, sobre un amplio conjunto de fuentes de datos, con la expectativa de encontrar información de interés para el desarrollo de su negocio, se está en presencia de una vigilancia *pasiva o scanning*. Cuando la búsqueda de información sobre actividades relevantes seleccionadas tiene un carácter investigativo, con el fin de proveer un conocimiento continuo sobre los desarrollos y las tendencias emergentes, se le llama vigilancia *activa o monitoring*. Si en ella se incluye la búsqueda puntual de información sobre un determinado tema, se denomina *search* (Escorsa y Maspons, 2001). Cuando la búsqueda incluye tanto al *scanning* como al *monitoring*, es decir, tiene un carácter más general, se denomina *watching*. En este caso el proceso de vigilancia está mucho más organizado y es más sistemático, y se está en presencia de un sistema de organización de la observación, análisis y difusión precisa de la información

para la toma de decisiones. Ese sistema de vigilancia filtra, interpreta y valora la información para permitir a sus usuarios actuar con más eficacia<sup>3</sup>.

## Clases de Vigilancia

Se habla de cuatro tipos de vigilancia:

- tecnológica** o centrada en el seguimiento de los avances del estado de la técnica y en particular de la tecnología y de las oportunidades /amenazas que genera.
- competitiva**, implica un análisis y seguimiento de los competidores actuales, potenciales y de aquellos con producto sustitutivo.
- comercial**, dedica la atención sobre los clientes y proveedores.
- del entorno**, centra la observación sobre el conjunto de aspectos sociales, legales, medioambientales, culturales, que configuran el marco de la competencia.

## Vigilancia Tradicional

**“Todas las acciones económicas están insertadas en redes de relaciones interpersonales”.**

**La implantación de acciones de Vigilancia fiables y eficaces pasa por la creación y la gestión de un cierto número de redes.**

La vigilancia tradicional es la que se ha realizado toda la vida en las empresas. Aunque quizá el término Vigilancia se ha puesto más o menos de moda en los últimos diez años, las empresas llevan realizando Vigilancia desde hace mucho tiempo.

La asistencia a ferias, congresos, la consulta de catálogos, las revistas especializadas que se recibían por correo en formato papel...

La Vigilancia tradicional utiliza fuentes clásicas para atender a estos frentes:

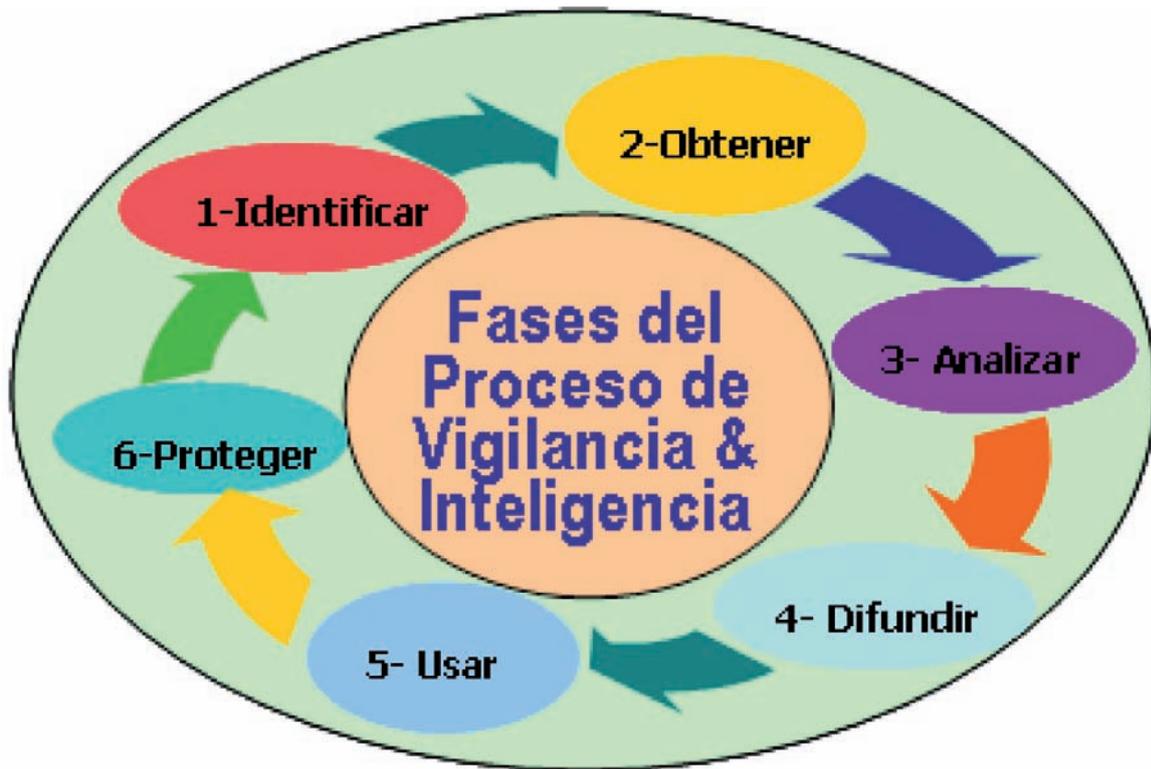
**Vigilancia competitiva** competidores actuales y potenciales.

**Vigilancia comercial** clientes, mercados, proveedores.

**Vigilancia tecnológica** tecnologías disponibles y/o emergentes.

**Vigilancia del entorno** sociología, política, reglamentaciones, medio ambiente...

<sup>3</sup> MALAVER RODRÍGUEZ, Florentino; VARGAS PÉREZ, Marisela (ed.): Vigilancia tecnológica y competitividad sectorial: lecciones y resultados de cinco estudios, ISBN: 978-958-44-1156-3, Bogotá, 2007.



Y de forma tradicional contestaba en la medida de lo posible a estas preguntas, aunque de forma no estructurada:

- ¿Cuáles son las principales líneas de investigación/innovación?
- ¿Qué tecnologías/productos emergentes están apareciendo?
- ¿Qué hacen los competidores/proveedores?
- ¿Quiénes son los líderes? (Centros de investigación, equipos, personas...).

La necesidad de saber qué es lo que está ocurriendo en el entorno (nuevos mercados, nuevos competidores, nuevas amenazas...) ha sido junto con la aparición de Internet la causa del acercamiento de la Vigilancia a la empresa.

#### Vigilancia Avanzada

Los volúmenes de datos, informaciones y conocimientos que se almacenan en bases de datos y en los que se han dado en llamar grandes repositorios de datos permiten una exploración mediante diferentes opciones de búsqueda.

**Minería de datos:** Es un mecanismo de explotación, consistente en la búsqueda de información valiosa en grandes volúmenes de datos. Está muy relacionada con las bodegas de datos que proporcionan la información histórica con la cual los algoritmos de minería tienen la información necesaria para la toma de decisiones.

Según *Fallad* y sus coautores (1996): "La minería de datos es un proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos".

Según *Molina* y sus colaboradores (2001): "Es la integración de un conjunto de áreas que tienen como propósito la identificación de un conocimiento obtenido a partir de las bases de datos que aporten un sesgo hacia la toma de decisión".

**Minería de textos:** Debido a que el 80 por ciento de la información de una compañía está almacenada en forma de documentos, las técnicas como la categorización de texto, el procesamiento de lenguaje natural, la extracción y recuperación de la información o el aprendizaje automático, entre otras,

apoyan al Text Mining (minería de texto). En ocasiones se confunde el Text Mining con la recuperación de la información (Information Retrieval o IR) (Hearst, 1999). Esta última consiste en la recuperación automática de documentos relevantes mediante indizaciones de textos, clasificación, categorización, etcétera. Generalmente, se utilizan palabras clave para encontrar una página relevante. En cambio, el Text Mining se refiere al examen de una colección de documentos y el descubrimiento de información no contenida en ningún documento individual de la colección; en otras palabras, trata de obtener información sin haber partido de algo. (Nasukawa y otros, 2001)<sup>4</sup>.

La **minería de texto** pretende facilitar el tratamiento por parte de las computadoras de la semántica del lenguaje natural. La mayor parte del conocimiento humano está representado en lenguaje natural. Para poder acceder a dicho conocimiento es necesario poder contestar a estas preguntas:

- ¿Cómo buscamos la información?
- ¿Cómo comparar fuentes de información diferentes y sacar conclusiones?
- ¿Cómo manejamos los textos para, por ejemplo, traducirlos o editarlos?

La **Minería de texto** consiste en la búsqueda a partir de técnicas de aprendizaje automático de regularidades o patrones que se encuentran dentro de un texto.

En la literatura sobre el tema, se refiere que las aplicaciones de la minería de textos se utilizan principalmente para:

- Extraer información relevante de un documento.
- Agregar y comparar información automáticamente.
- Clasificar y organizar documentos según su contenido.
- Organizar depósitos para búsqueda y recuperación.
- Clasificar textos e indizarlos en el Web.

Pero la realidad es que son muchos los casos de organizaciones que han aplicado la minería de textos con el afán de alcanzar objetivos más reales y ambiciosos (en el mejor sentido) que estos. Por ejemplo, el uso que le dan algunas empresas para identificar el contenido de los correos electrónicos que les envían los clientes y redirigirlos a los departamentos apropiados, incluso si el sistema es capaz de identificar el contenido de una consulta frecuente, envía una respuesta estándar, sin necesidad de intervención humana. En el campo de la vigilancia tecnológica y en la “*Business Intelligence*”, para bucear en las bases de datos textuales y seguir la evolución de la competencia. Además de que se apunta también la posibilidad de usar esta técnica en la investigación de mercados en el Web, sobre la base de recoger estadísticas sobre la utilización de determinados conceptos y temas en la red con el objetivo de estimar la demografía y las curvas de demanda de productos asociados con ellos. Todas esas aplicaciones son perfectamente transferibles a la gestión de información que ocurre en el interior de las bibliotecas y están llamadas a redimensionar la función de la entidad, así como del bibliotecario, tanto hacia adentro como fuera de esta<sup>5</sup>.

<sup>4</sup> Fuente: Bressán G. 2003. Almacenes de datos y minería de datos. 2003. Disponible en: <http://www/banners/interstitial.html?http://exa.unne.edu.ar/> [Consultado: 29 de julio de 2009].

<sup>5</sup> Fuente: Copiado y pegado de “Minería de textos: una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital” MSc. Eleazar BOTTA-FERRET<sup>1</sup> y Lic. Jania E. CABRERA-GATO<sup>2</sup>, Acimed 2007; 16(4) [http://bvs.sld.cu/revistas/aci/vol16\\_4\\_07/aci051007.html](http://bvs.sld.cu/revistas/aci/vol16_4_07/aci051007.html), [consulta: 3 de agosto de 2009].

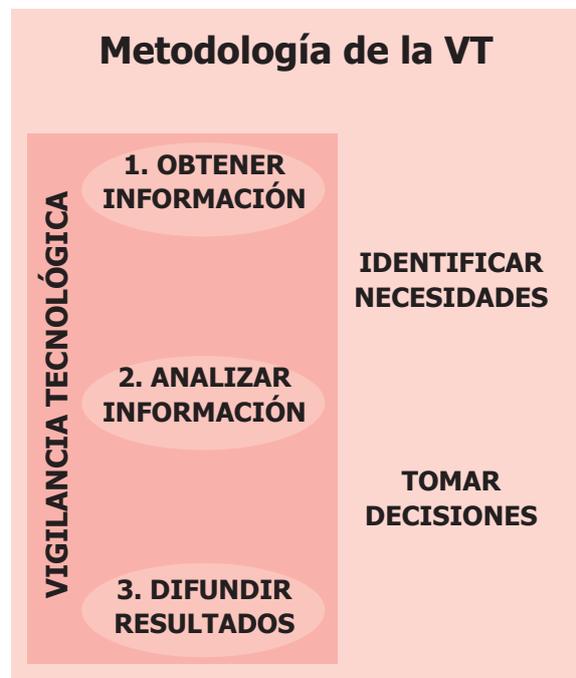
# Metodología y diseño: ¿por dónde empezamos?

La información útil no llega a las empresas en la forma que éstas necesitan, a menudo se halla inmersa junto con un cambalache de información inútil. La tarea del documentalista es evaluar esa información de acuerdo con unos criterios:

- Fiabilidad de las fuentes.
- Riqueza (en datos o referencias).
- Vulnerabilidad (facilidad de manipulación).
- Variación dependiendo del formato y del entorno en donde se recoge esa información (utilizar indicadores [ver presentación de Daniel Torres-Salinas, 2009]).

## Metodología y fases

1. **Implicación de la Dirección:** Como cliente y soporte de la estrategia, la implicación de la Dirección se antoja fundamental en el proceso de implantación de un sistema de gestión de la información.
2. **Nombramiento del animador:** El valor del sistema de gestión de la información se encuentra en su dinamismo, aspecto que debe fomentarse desde la figura de un responsable o animador del sistema.
3. **Definición del equipo de gestión de la información:** El desarrollo de las actividades lleva asociada la creación de roles y el reparto de responsabilidades a un colectivo que debe presentar un determinado perfil de competencias.
4. **Diagnóstico de gestión de la información:** La gestión de la información es un enfoque que nunca parte de cero en las organizaciones, por lo que es necesario identificar patrones actuales y potenciales para el manejo de la información.
5. **Identificación de factores y fuentes de información:** Desde la colaboración y comunicación de la Dirección y el equipo de gestión de la información, se establecen los factores clave a vigilar, identificando y seleccionando las fuentes más relevantes. (FCV).
6. **Creación del Manual de gestión de la información:** En línea con las iniciativas de calidad organizativa es conveniente la documentación de las prácticas que se siguen para el desarrollo del proceso de gestión de la información.
7. **Revisión de procedimientos:** Una vez sentada la base de actuación en materia de gestión de la información habrá de establecerse una revisión periódica para adecuar cada uno de los aspectos a los cambios del contexto interno y externo a la organización.



<i>Fases del ciclo de VT</i>	<i>Ashton y Klevans (1997)</i>	<i>Rodríguez (1999)</i>	<i>Vargas y Castellanos (2005)</i>
<p><b>FASE I</b> Planeación e identificación de necesidades</p> <p><b>FASE II</b> Identificación, búsqueda y captación de información</p> <p><b>FASE III</b> Organización, Depuración y Análisis de la información</p> <p><b>Fase IV</b> Procesos de Comunicación y Toma de decisiones / Uso de resultados</p>	Necesidades Planeación de actividades Fuentes y Métodos	Planeación	Información previa Planeación
	Recolección de fuentes de información	Selección de las fuentes de información y Acopio	Preparación de la Búsqueda Búsqueda en bases de datos
	Análisis de Datos	Análisis	Depuración y convalidación de registros Procesamiento de Registros Análisis e Interpretación de los resultados
	Entrega de Información Evaluación de los resultados Uso de los resultados	Difusión de resultados Procesos de decisión Acciones	Diseño de estrategias Impactos

Fases de procesos de la vigilancia tecnológica<sup>6</sup>.

<sup>6</sup> MAURICIO LEÓN, Andrés; CASTELLANOS, Óscar y VARGAS, Freddy: Valoración, selección y pertinencia de herramientas de software utilizadas en Vigilancia Tecnológica, En: Revista Ingeniería e Investigación, vol. 26, nº 1, abril 2006.

La norma UNE 166.006:2006 EX tiene por objeto facilitar la formalización y estructuración en cualquier organización del proceso de escucha y observación del entorno para apoyar la toma de decisión a todos los niveles de la organización hasta devenir en la implantación de un sistema permanente de Vigilancia (se volverán a tratar estos puntos de la norma en el siguiente capítulo).

Utilizar esta norma como guía para establecer una metodología de lo que sería nuestro sistema de VT es un comienzo.

A parte de la metodología descrita aquí arriba que no forma parte de la norma, la norma UNE recoge estos pasos:

1. Identificar áreas y objetivos prioritarios de actuación del sistema de VT.
2. Asegurarse de la disponibilidad de recursos e información necesarios para permitir la operación y seguimiento del proceso.
3. Realizar el seguimiento, la medición y el análisis del proceso.
4. Llevar a cabo las acciones necesarias para alcanzar los resultados planificados.

Fases de una Vigilancia Ocasional.

1. Definición del problema.
2. Identificación de las fuentes.
3. Búsqueda.
4. Análisis.
5. Validación de la Información.
6. Elaboración de un Informe.

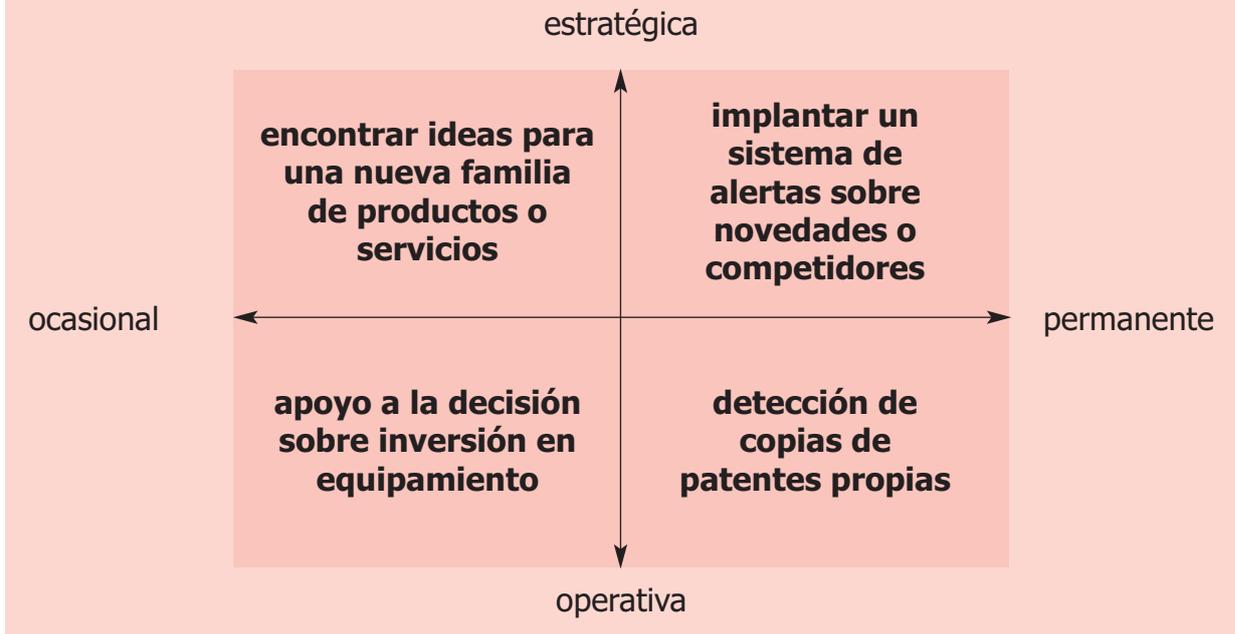
Fases de una Vigilancia Permanente.

1. Definición de los Factores Críticos de Vigilancia.
2. Identificación de las fuentes.
3. Identificación de las personas/grupos receptores de esa vigilancia.
4. Búsqueda periódica y análisis.
5. Envío de la información o elaboración de informes si es necesario.

### Proceso completo de Vigilancia

- Diagnóstico estratégico del sector.
- Identificación de los problemas competitivos relevantes.
- Identificación de los factores críticos competitivos y tecnológicos.
- Identificación de los factores tecnológicos críticos y las tecnologías a vigilar.
- Definición del objetivo de la vigilancia tecnológica.
- Búsqueda de la información.
- Identificación de las palabras clave asociadas al tema a vigilar.
- Identificación y selección de las fuentes de información relevantes.
- Formulación de la ecuación de búsqueda.
- Análisis de la información.
- Análisis de los resultados.
- Distribución de los resultados.
- Inteligencia Competitiva.
- Distribución de esa inteligencia.

## ¿Cuál es la finalidad de la Vigilancia Tecnológica?



# La información y la empresa

«Tu vigilancia necesita extenderse más allá de las fronteras tradicionales de tu sector y sus actuales productos. Tus competidores tradicionales son a menudo el punto de vista equivocado para detectar las señales de cambios profundos.» Eckhard Pfeiffer, Presidente de Compaq Computer.

## ¿Para qué sirve la Vigilancia en una empresa?

Las empresas europeas pierden más de 3,5 billones de Euros en investigar y desarrollar productos ya patentados (Farreras 1994).

Se estima que el 80% de la información Técnica está almacenada en las bases de de patentes. Sin embargo el uso de las Patentes como fuente de in-

formación en Europa está muy lejos del de Japón o Estados Unidos.

*Una empresa que tenga acceso a información valiosa para su competitividad, en una forma oportuna y adecuada, que cuente con una cultura en la que se promueva la discusión y donde los miembros clave de la empresa mantengan contacto con fuentes externas de información, y más importante aún, que pueda transformar esta información en un producto “inteligente”, es decir, que brinde resultados de alto valor estratégico, verá incrementada significativamente sus posibilidades para obtener innovaciones exitosas.*

Fuente: Escorsa y Maspons (2001).

### Las empresas de la UE gastan 31.800 millones en investigación redundante

Son esfuerzos repetidos y, por lo tanto, costosos e inútiles. Las empresas europeas gastan 31.800 millones de euros anuales en realizar investigaciones redundantes. Según la Unión Europea, la cifra es una muestra más del desconocimiento que persiste entre los investigadores comunitarios sobre los beneficios y la utilidad del sistema de patentes. Una carencia que Bruselas quiere suplir lo antes posible.

Con una precaución tan simple como la de solicitar información en sus respectivas oficinas nacionales de propiedad industrial antes de invertir tiempo y dinero en investigación, las empresas europeas podrían ahorrarse no miles, sino millones de euros al año. Según cifras publicadas por la Comisión Europea, las compañías comunitarias gastan 31.800 millones de euros anuales en investigaciones redundantes, es decir, en invenciones en muchos casos ya patentadas.

La propia Comisión atribuye a la falta de información de los investigadores sobre el funcionamiento del sistema de patentes buena parte de esta situación. En el caso de España, por ejemplo, la Oficina Española de Patentes y Marcas (OEPM), organismo dependiente del Ministerio de Ciencia y Tecnología, ofrece un servicio de búsqueda documental que permite rastrear la existencia de investigaciones ya patentadas o en fase de solicitud de patente.

Según Carmen Toledo, jefa del servicio de búsqueda de la oficina, la OEPM realiza dos tipos de búsqueda: una hecha a la medida de cada empresa y otra a través de boletines trimestrales de vigilancia tecnológica. La primera tiene un coste de entre 18 y 82,5 euros, según la búsqueda sea en el ámbito nacional o mundial. La segunda, que hasta el momento cubre sólo ciertos sectores, alcanza los 298 euros.



### La información falsa en Internet

La Nasdaq (National Association of Security Dealers-Automatic Quotation, bolsa electrónica norteamericana) ha introducido un dispositivo de vigilancia (surveillance device) que permite detectar información falsa en internet lanzada con la intención de afectar un determinado valor cotizado.

Desarrollado por la conocida empresa especializada en seguridad de datos RSA (iniciales de los apellidos de sus fundadores Ronald L. Rivest, Adi Shamir y Leonard M. Adleman), utiliza sistemas criptográficos y pretende controlar todo lo que se dice en chatrooms sobre bolsa, descubriendo qué datos erróneos se vierten sobre las 4.000 principales empresas norteamericanas.

Más detalles sobre el sistema en:

<http://www.news.com/News/Item/0,4,14008,00.html> [*Financial Times*, 05/09/97, p. 1].

### La Vigilancia y su ayuda a la empresa

La puesta en marcha de **procesos de vigilancia** permite anticiparse, encontrar respuestas y localizar la información relacionada con un tema previamente definido, reducir riesgos y aumentar

nuestra capacidad de innovación. A partir de ahí, recuperados los documentos que cumplen con los parámetros objetivos, éstos se deberán analizar, valorar, ordenar, clasificar, haciéndolos accesibles y de fácil uso.

Existen opciones de búsqueda de diferentes características, gratuitas o de pago, con más o menos prestaciones a la hora de localizar información. No obstante, obtener resultados depende, en la mayor parte de los casos, de la habilidad del usuario a la hora de formular la consulta, priorizar necesidades, analizar la información disponible y los “gaps” existentes, identificar y validar la fiabilidad de las fuentes, así como las herramientas a utilizar, ser flexibles y, cuando sea necesario, contar con expertos. Teniendo en cuenta estas premisas la cuestión sobre la inversión necesaria puede ser, en la mayoría de los casos, no demasiado relevante.

Existen distintas opciones y herramientas a la hora de implantar un sistema de estas características. Las herramientas de software, son fácilmente abordables y haciendo un uso inteligente de las mismas podremos conseguir una amplia variedad de objetivos.

El tamaño de la empresa no será, por tanto, una variable que condicione o defina la pertinencia o no de la **función de vigilancia**. Por el contrario, cuando en nuestro día a día percibamos que la tecnología se puede comprar sin demasiada dificultad, que el producto es conocido, que no existen diferencias destacables en costes y que, sin embargo, el valor se encuentra en aportar algo más cada día, dando soluciones innovadoras y diferentes al resto, necesitaremos, posiblemente, conocer y aplicar alguna de las herramientas disponibles en el mercado para realizar funciones de vigilancia<sup>7</sup>.

### La empresa y su entorno

Podemos identificar tres tipos de información en función de su valor dentro de la empresa:

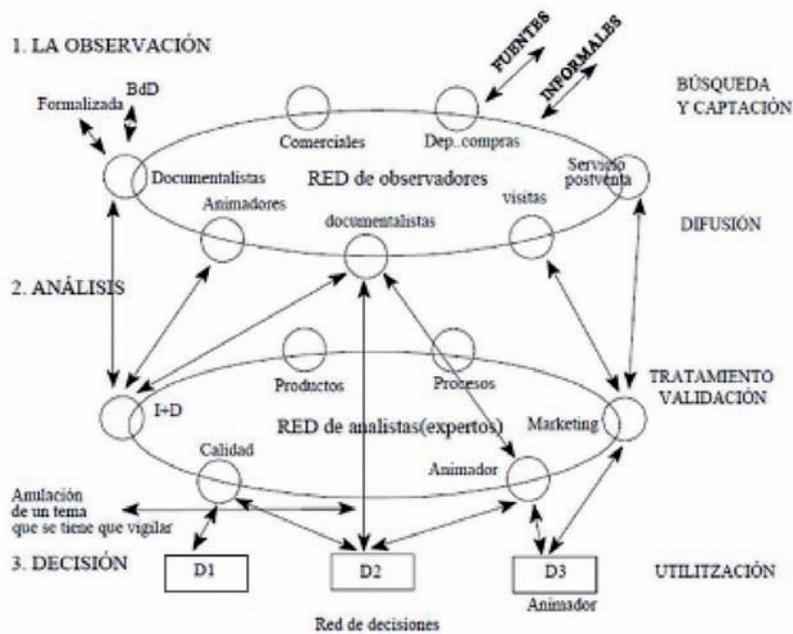
a) **Difusa**. Por supuesto la más difícil de detectar, y al mismo tiempo la que tiene más valor. Puede conseguir modificar el desarrollo de un proyecto. Muchas veces llega a nosotros de forma indirecta y cuesta identificarla ya que una vez detectada no es fácil cotejarla. Normalmente se encuentra dentro de documentación difícil de localizar:

- patentes, tesis, artículos científicos, actas de congresos especializados, “papers”, etc.

b) **Latente**. Se encuentra disgregada y no es de difícil localización, pero sí es complicado unificarla. Suele aparecer en fuentes diferentes y con formatos diversos. Se puede utilizar para documentar el estado de la técnica, sin hacer un análisis profundo de los resultados obtenidos. Para localizar información de este tipo podemos recurrir a los clientes, a bases de datos de proyectos de I+D, a buscadores especializados, foros, etc.

c) **Del entorno**. De carácter generalista, de poco valor para la estrategia de la empresa, pero que marca una pauta o ayuda a ver la tendencia del entorno. Está formada por noticias de prensa, novedades de las ferias, catálogos de proveedores, etc.

Alfons Cornella cita a Jakobiak cuando establece una tipología de la información que denomina crítica, y que está formada por los siguientes tipos: científica, técnica, tecnológica, técnico-económica, jurídica y reglamentaria, de entorno y seguridad, e información cualitativa y general.



Fuente: Jakoviak 1992.

<sup>7</sup> Vigilancia Tecnológica e Inteligencia Competitiva: usos y aplicaciones en la empresa navarra, 2005. En: <http://www.navac-tiva.com/web/es/documentacion?tag=210> [consultado el 3 de septiembre de 2009].

## Información externa (modelo de Laudon y Laudon)

Entre los estudios para clasificar la información del entorno de la empresa destaca, también, la clasificación de Laudon y Laudon que sirve, además, para estructurar la captura y absorción de información.

Así, identifican dos entornos, inmediato y remoto:

**El entorno inmediato**, lo conforman los activos que una empresa trata a diario, como clientes, distribuidores, competidores, proveedores, financiadores y reguladores.

**El entorno remoto**, está formado por aquellos elementos que una empresa debe tener en cuenta para controlar el entorno en el que se encuadra, y que está formado por la información sobre la situación política, la sociedad, los cambios tecnológicos o la evolución económica.

## Información interna

Alfons Cornella (1994, 104) destaca dos aspectos a tener en cuenta a la hora de acometer el análisis de la información interna y sus repercusiones sobre la competencia de la empresa. Por un lado distingue entre la información operacional, aquella que se genera por el propio funcionamiento rutinario de la empresa y los conocimientos o "Know-how", que no son otra cosa sino el resultado de la integración de la información interna y externa y las capacidades creativas de los miembros de la empresa. En segundo lugar, afirma que la información tiene sentido sólo cuando alguien la usa para algo.

Aunque resulta obvio que todos los miembros de una empresa utilizan información en mayor o menor medida, la mayoría de las investigaciones sobre el uso de la información interna de las empresas, se han centrado en el estudio del com-

portamiento de los directivos (McKinnon, 1992). Ello se debe a que una de las características de la gestión de una empresa según Forrester (1968) es la de convertir información en acción en el proceso de toma de decisiones, actividad que desarrollan por excelencia los directivos.

Estas consideraciones permiten enfocar el estudio de la información interna de una empresa como recurso competitivo, sobre la base de cómo es utilizada dicha información por sus directivos en el desempeño de sus tareas directivas, actividades todas ellas que involucran el manejo de información.

## Las fuentes de información de una empresa

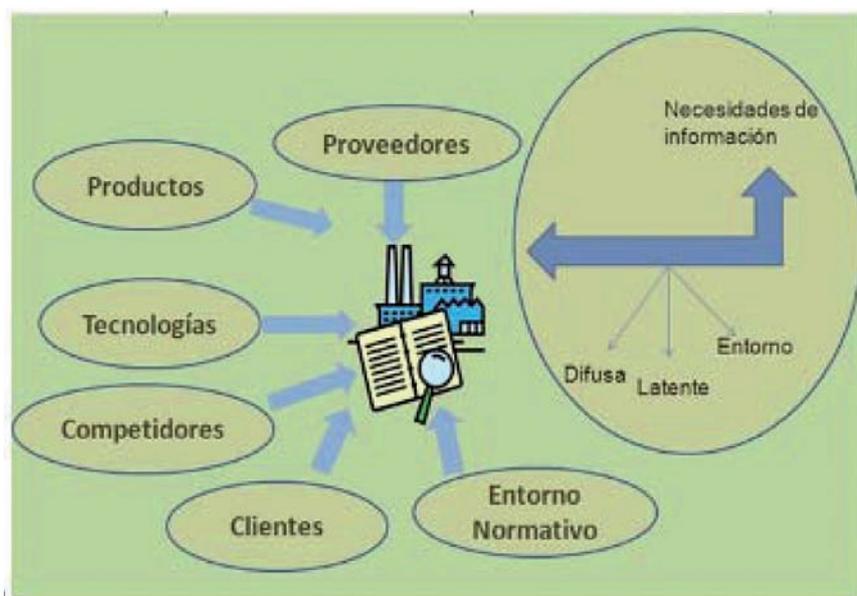
Tipos de fuentes de información.

Las principales fuentes de información que pueden alimentar la función de vigilancia tecnológica son:

- los contactos directos personales con los competidores, proveedores, centros de investigación, universidades, etc...
- la participación en coloquios, congresos y otras manifestaciones científicas,
- las revistas especializadas,
- las patentes como sistemas de información ya que reflejan cuales son las grandes tendencias las bases de datos (P. Morcillo Revista Madri+d Número 17, junio- julio 2003).



Fuente: Clasificación de Juan Carlos Vergara.



### Diseño de un sistema de VT en una empresa<sup>8</sup>

Antes de iniciar cualquier intento de diseño es conveniente plantearse, a modo de reflexión, las siguientes preguntas:

1. ¿Cuál es el objeto de la vigilancia?
2. ¿Qué debemos vigilar? ¿Qué informaciones buscar?
3. ¿Dónde localizarlas?
4. ¿Cómo tratar y organizar la información?
5. ¿A quién comunicar la información en la empresa?
6. ¿Cómo promover la implicación de todo el personal.
7. ¿Qué recursos vamos a destinar?

Las respuestas a estas preguntas deben plasmarse en el manual de vigilancia, completando el contenido de las acciones o etapas descritas en el esquema inicial. En el manual se recogen los objetivos, las fuentes, los recursos y el proceso operativo de la vigilancia, constituyendo en cierta medida «las instrucciones de uso» del sistema, adaptadas a las especificidades de la empresa.

Los Factores Críticos de Vigilancia son aquellos aspectos críticos de los que depende la buena mar-

cha de la empresa”. Son inherentes al sector de actividad, dependen de los objetivos y la estrategia de la empresa, son variables en el tiempo y tienen un carácter arborescente.

Las actividades posteriores a la definición de los FCV deben ser en su orden:

- a) Identificar recursos de información.
- b) Definir el plan y realizar el manual de vigilancia tecnológica.
- c) Constituir el grupo de trabajo, asignar funciones, motivarlos.
- d) entrenarlos.
- e) Evaluar el funcionamiento.

Las fases de la vigilancia tecnológica según la Norma UNE 166.002 son:

1. Identificación de los *factores críticos de vigilancia*, las cuestiones externas a la organización cuya evolución es crucial para su competitividad: tecnologías emergentes, competidores actuales y potenciales, desarrollo de los mercados y del entorno.
2. Identificación de las *fuentes o tipos de información* relevante para seguir los factores críticos de vigilancia: formales (patentes, prensa, bases de datos, informes, publicaciones...) o informales (conversaciones, apuntes, reuniones, congresos, ferias, exposiciones, encuestas...).
3. Elección de los *medios de acceso y seguimiento* de las fuentes de información, en muchos casos mediante software especializado o sistemas y servicios automatizados: control de estudios e informes de mercado, vigilancia de prensa, servicios de bases de datos, monitorización y rastreo de internet, etc.
4. Ejecución regular de la *búsqueda de información*, mediante una estrategia sujeta a revisión constante.
5. Análisis, evaluación y organización, *puesta en valor, de la información* obtenida, para sacar

<sup>8</sup> Para ver otro diseño: León Aliz, T; Diseño e implementación de un Sistema de Vigilancia Tecnológica en una empresa de escasos recursos, Congreso Internacional de Información info 2004 En <http://www.bibliociencias.cu/gsdll/cgi-bin/library?e=d-000-00-0eventos-00-0-0-0prompt-10-4-0-11-1-es-50-20-about-00031-001-1-0utfZz-8-00&a=d&c=eventos&cl=CL1.4&d=HASH63935b5e7ca3232eaa7606> [Consultado el 3 de septiembre de 2009].

conclusiones útiles a la organización, en muchos casos también mediante software de análisis, tratamiento y presentación de datos.

6. *Difusión selectiva de la información* elaborada, por los canales y a las personas adecuadas.

Principios básicos de un sistema de VT (Fuente: <http://www.emotools.com>)

Un sistema organizado e integrado en los “hábitos” de funcionamiento de la organización, de una manera “natural”.

Requiere un enfoque multidisciplinar y no acotar la visión a un departamento.

La vigilancia es tarea de todos: cualquier empleado puede aportar información de valor. Esto lleva consigo que la gente adquiera una “cultura de la alerta. La eficacia (y eficiencia) de un sistema de V.T. se multiplica con una adecuada “cultura de la información” en todo el personal. La participación activa de todo el personal se consigue con “incentivos” claros que premien la cantidad y calidad de información aportada (no basta con “cultura”).

Es fundamental definir las prioridades informativas: ¡Quien no sabe lo que busca, no sabe lo que encuentra! Esto es de sentido común. A veces la se-

rendipia nos puede dar un buen resultado pero es poco probable.

La V.T. debe impulsarse al más alto nivel, desde los máximos directivos de la organización, que deben dar el ejemplo como “proveedores” y “usuarios” de información del sistema, pero debe tener un “dinamizador”. Una persona que se encargue de dinamizar, canalizar y asesorar al resto de personal.

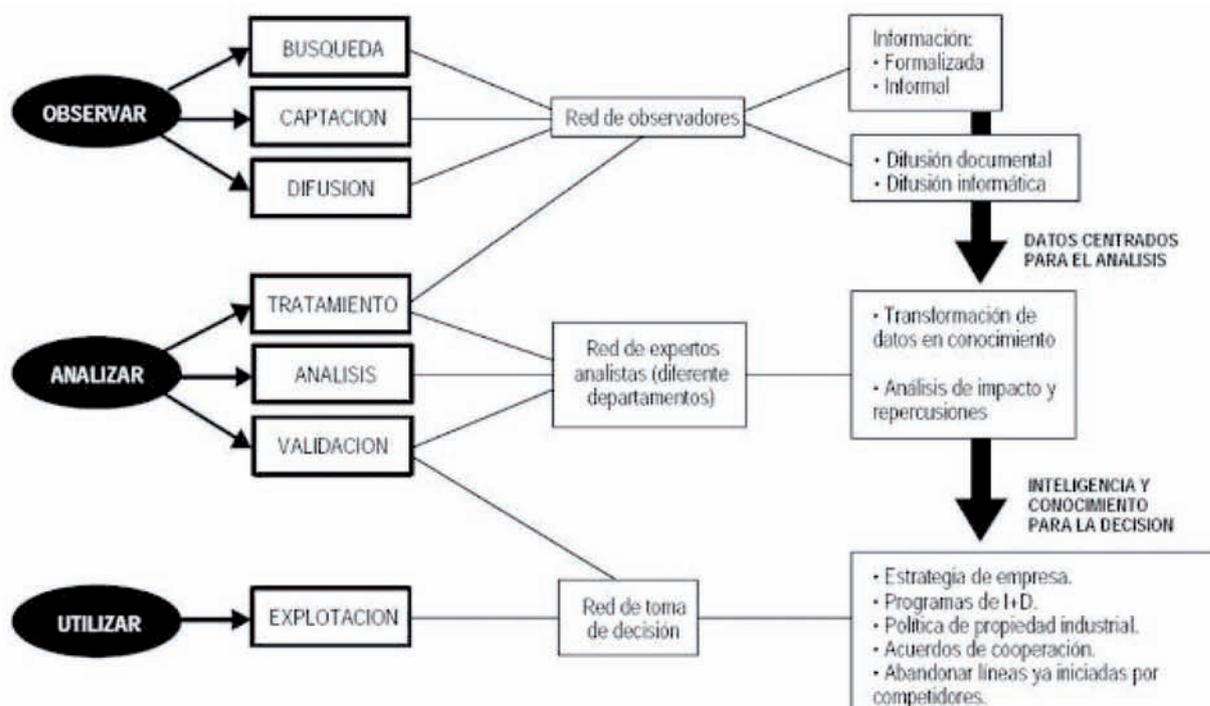
Es imprescindible crear una “central de análisis” que reúna, procese y distribuya toda la información. Así como de que centralice la realización de Informes de VT y desarrolle el canal de comunicación de la información así como de las peticiones de búsqueda.

El verdadero reto y valor añadido está en la selección de la información valiosa: ¡evitar la infoxicación! Para ello, se deben definir “criterios de búsqueda”.

Evitar la “miopía sectorial”: las mayores amenazas pueden venir de otros sectores.

La vigilancia cuesta dinero, por tanto, necesita un presupuesto específico para que se pueda planificar y realizar de modo ordenado.

Hay que estimular las redes externas de colaboradores para integrarlos al sistema.



# Las fuentes de información

## Conceptos previos

### Literatura gris

El concepto de literatura gris es relativo e incluye cualquier tipo de documento que, por determinadas circunstancias, no se difunde por los cauces ordinarios y, en consecuencia, conlleva problemas de consulta.

La literatura gris también es conocida como literatura semipublicada, fugitiva, informal, invisible, subterránea, efímera, no permanente y no convencional, está integrada por una extensa gama de documentos cuyas características no están del todo definidas y cuyo valor informativo es bastante dispar.

### Tipos de literatura gris

- Informes públicos y privados con información científica, técnica, económica, social, etc.
- Tesis no publicadas en el mercado editorial.
- Comunicaciones a congresos no publicadas en actas disponibles en el mercado editorial.
- Normas y recomendaciones de carácter técnico, marcas y patentes.
- Traducciones inéditas.
- Algunas publicaciones periódicas no publicadas comercialmente (p.e. de sociedades, periódicos locales, etc...).
- Algunos documentos oficiales de ámbito restringido o muy limitado.
- Documentación técnica publicitaria.
- Informes de laboratorios, estudios de mercado.
- Programas de ordenador de circulación restringida entre grupos de personas.
- Otros documentos que se ajusten a las características indicadas.

- Algunos autores incluyen pasquines, hojas volanderas, propaganda comercial o política, planos, almanaques, agendas, etc...

### Web Invisible

Hay estimaciones que calculan en 500 veces más grande que el total de la información indizada por los buscadores, la información que permanece invisible en la World Wide Web, lo que se ha denominado el inmenso océano de la Internet profunda.

Aunque los buscadores generalistas no suelen indizar archivos no textuales, sí existen una serie de buscadores especializados que indizan imágenes, vídeo, audio, archivos pdf, archivos comprimidos o ejecutables. Sin embargo, muchísimos datos quedan fuera de los buscadores tradicionales, ya sean estos generalistas o especializados, puesto que indizar cierto tipo de informaciones contenidas en enormes bases de datos numéricas o textuales, exige gastar muchos recursos y resulta muy costoso para los buscadores almacenar en sus bases de datos este tipo de formatos. Por otro lado, los buscadores tampoco indizan muchos de los datos que se generan de forma dinámica en tiempo real, puesto que se convierten en obsoletos en un brevísimo lapso de tiempo y no merece gastar recursos en informaciones tan fugaces; y a esto se une que muchas de estas bases de datos dinámicas han de rastrearse desde su propia ubicación o sitio web, y con sus propias herramientas de búsqueda personalizadas, puesto que precisan de pasarelas o contraseñas especiales para acceder a ellas. Si a esto unimos las páginas sin conexión o enlaces aparentes, vemos que una enorme masa de información no es accesible desde los principales buscadores existentes en la World Wide Web.

A toda esta gran masa de información es a la que se ha denominado Internet oculta<sup>9</sup>.

### Contenidos invisibles

Bases de datos bibliográficas: se incluyen en este grupo los catálogos de biblioteca accesibles a través de una pasarela (OPAC) web, otras bases de datos de referencias bibliográficas (de acceso público o restringido –registro previo gratuito o de pago-) y entidades similares tales como los catálogos de librerías (ej.: Amazon.com).

Bases de datos alfanuméricas: definidas por exclusión del grupo anterior, son todas las bases de datos que no tienen carácter bibliográfico. Comprendería, además, los recursos llamados de referencia que requieren algún tipo de pasarela de acceso para su consulta (ej.: Enciclopedia Británica).

Una situación ligeramente distinta es la planteada por las páginas generadas dinámicamente (asp, jsp, php o similares). Dichas páginas sólo existen en virtud de una consulta puntual, imposible de realizar por los robots de los motores de búsqueda y cuyo contenido puede alcanzar un considerable grado de personalización. Desde un punto de vista documental, los contenidos que explotan están en una base de datos y, por tanto, se consideran dentro de esta categoría.

Archivos y revistas electrónicas: se trata de bases de datos que incluyen documentos a texto completo y que sólo se pueden recuperar previa identificación de la referencia, labor para la que se requiere utilizar una pasarela web simple (formulario de consulta) o doble (palabra de acceso y formulario de consulta).

Ficheros no HTML o textuales: el (relativo) fracaso de HTML original a la hora de generar páginas con una maquetación muy rica, ha permitido que algunos formatos de diseño más elaborado hayan adquirido popularidad en la web (pdf, ps, ppt, doc). Estos formatos no textuales no son indizados correctamente por los robots de los motores de búsqueda (excepción: Google.com ya indiza documentos pdf y los convierte en HTML) y, por

tanto, constituyen una parte del webespacio invisible que ha ido adquiriendo cada vez más importancia.

*“La información, recurso estratégico indispensable para el desarrollo económico y social de un pueblo, muchas veces no logra llegar al público a través de los canales normales o convencionales de divulgación; un caso particular es la llamada literatura gris que sirve de apoyo a los procesos de investigación y que frecuentemente proporciona información no localizada en las fuentes convencionales. La literatura gris –que fue rápidamente aceptada por la comunidad científica, bien como fuente o como instrumento de difusión– trajo consigo un nuevo modo de presentar y de dar a conocer los resultados de investigación, una nueva forma de trabajar y otra manera de entender el proceso mismo de la investigación; en resumen, se trata de una nueva forma de comunicación científica. Actualmente, sus usuarios se han ido incrementando gracias a las nuevas TIC propicias para producirla y difundirla; como ejemplos encontramos la elaboración constante de informes y proyectos previos para cualquier toma de decisión y también el desarrollo de ediciones de documentos electrónicos accesibles por medio de redes de comunicación, que por su parte incrementan el número de publicaciones que caen en esta área gris. Su uso se da entre los mismos que la generan: científicos, investigadores, estudiantes, etc.; es decir, ellos mismos tienen necesidades de este tipo de literatura. Entre las entidades que más la generan se ubican las universidades, empresas públicas y privadas, laboratorios, bibliotecas, museos, sociedades, organismos de investigación, etc.”<sup>10</sup>*

### Internet

Como es bien sabido, la fiabilidad y el rigor de las fuentes de Internet deben ser tenidas en cuenta a la hora de citarlas y la red de redes no es una excepción, así que María Rubio Lacoba nos ofrece en el libro *Teoría y Práctica de la Documentación In-*

<sup>9</sup> Fuente: María Jesús Lamarca Lapuente. *Hipertexto: El nuevo concepto de documento en la cultura de la imagen* [http://www.hipertexto.info/documentos/web\\_invisib.htm](http://www.hipertexto.info/documentos/web_invisib.htm) [enlace visitado el 11-09-2009].

<sup>10</sup> SORIA RAMÍREZ, Verónica: La literatura Gris y los e-prints, BIBL. UNIV., NUEVA ÉPOCA, JULIO-DICIEMBRE 2003, vol. 6, n.º. 2.

*formativa* (ISBN: 84-344-1293-4) algunas pistas para poder identificar lo que puede ser considerado como fuente, o no, en Internet.

1. **La autoría de la información.** Se dan muchos casos en los que se recupera información que no puede ser atribuida a ninguna persona o institución ya que no viene firmada. En Internet, cualquiera puede producir información pero el hecho de que se encuentre en la Red no quiere decir que pueda considerarse como una fuente válida. Por ello, debemos de escoger aquellas fuentes que, por su trayectoria analógica o digital, ya vengán avaladas con cierto prestigio o reconocimiento.
2. **La caducidad de la información.** Si la presencia en la Red de muchas de las informaciones no vienen firmadas, deberíamos preguntarnos cuántas de ellas vienen datadas. La fecha de cualquier documento es imprescindible para conocer el grado de actualidad o de actualización de los datos que ofrece. Sin embargo, en la Red muchas informaciones no poseen fecha de publicación, por lo que debemos ser cautelosos a la hora de recoger sus datos.
3. **La Internet Profunda.** A pesar de los grandes esfuerzos que realizan los motores de búsqueda para la indización de la mayor cantidad de páginas posibles, el hecho es que existe una parte de la Red a la que estos no alcanzan a registrar. A esta parte de la Red, se la conoce como la Internet Oculta o Profunda que condensa miles de millones de páginas. La *Internet Invisible* contiene tanto a bases de datos cuyos contenidos no son accesibles sin registro previo como otro tipo de fuentes de información como catálogos de bibliotecas, bases de datos bibliográficas, revistas electrónicas (en las que es necesario un registro previo y las que solo se puede recuperar la información mediante búsquedas en su base de datos), documentos en formatos no indizables, obras de referencia: enciclopedias, diccionarios... en las que es necesario interrogar a la base de datos para acceder al contenido, etc.
4. **Información-Conocimiento.** El propio hecho de la existencia de cierta información sobre un tema no quiere decir que sea conveniente-

mente aprovechada. Su peso específico lo otorga el uso, la gestión, la explicación en unas coordenadas – espacio temporales que hagamos de ella, de hecho, la información malentendida puede degenerar en justamente lo contrario: En la desinformación, por un lado, y también en la infoxicación.

5. **El ruido documental.** Ante la cantidad de información que se genera a cada segundo en Internet, no es infrecuente que a la hora de recuperar documentos nos topemos con algunos que nos resultan completamente irrelevantes o poco ajustados a nuestras necesidades informativas. Estos documentos son los que se denominan *ruido documental*, dependiendo de la cantidad de ruido que recuperemos deberemos plantearnos nuestras estrategias informativas y las fuentes de información que estamos utilizando<sup>11</sup>.

El acceso a Internet como fuente de información ha disminuido su porcentaje de hace dos años cuando se situaba en un 40%, además aquéllos que se informan a través de la web de manera suficiente siguen consumiendo los medios de comunicación tradicionales.

Según un estudio publicado en Estados Unidos, por la consultora Pew Research Centre, Internet es ahora por detrás de la televisión, la fuente más importante para estar informado sobre temas de actualidad. El estudio señala que un 27% de los norteamericanos, compra el periódico un día a la semana, cerca del 37% se informa regularmente a través de Internet y un 39% acude a los boletines informativos ofrecidos por la televisión.

Como dato interesante señalar que el acceso a Internet como fuente de información ha disminuido su porcentaje de hace dos años cuando se situaba en un 40%, además aquéllos que se informan a través de la web de manera suficiente siguen consumiendo los medios de comunicación tradicionales.

<sup>11</sup> Fuente: Qué es fuente de información en Internet, post del blog El Documentalista Enredado, Marcos Ros-Martín — Domingo, 30 de Abril de 2006 [consulta 04-09-2009].

## ¿Cómo evaluar la fiabilidad de una web?

1. Mirar detenidamente la URL
  - Es una web personal o un sitio “oficial”
  - Dominio. ¿Es apropiado para su contenido?
  - Coincide el dominio con la entidad.
2. Mirar los lados de la página
  - ¿Quién la escribe?
  - Nombre, mail
  - Fecha.
3. Examinar el contenido
  - Fuentes bien documentadas
  - Usa información de 2ª mano? Esta alterada o falseada? Esta referenciada?
  - Enlaces a otros recursos? Funcionan?
  - Es evidente su parcialidad?
4. ¿Qué es lo que otros cuentan de ella?
  - ¿Quién la enlaza? Pej. En Google.- link:toda.la.url
  - Es enlazada por algún blog? Si es así, que es lo que dice de ella?
  - Busca al autor en Google.
5. Todo encaja?
  - Que fin tiene la página? Comercial, educativo... Posible ironía, parodia?

Fuente: Teaching Library, UC Berkeley, Spring 2008 (Traducción propia).

[http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/evaluation\\_checklist\\_2008\\_spring.pdf](http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/evaluation_checklist_2008_spring.pdf)

Aunque no es posible establecer reglas estrictas de fiabilidad en los contenidos de Internet, el profesional de la información, y por extensión cualquier internauta, debe disponer de una serie de directrices que le permitan evaluar la información para determinar su calidad.

Las herramientas que se proponen para aplicar mínimos criterios de evaluación son: autoridad, credenciales, inteligibilidad del mensaje, independencia, usabilidad, imparcialidad, temporalidad, utilidad y fuentes de procedencia del documento<sup>12</sup>.

Lluís Codina (2000a, 2000b, 2003, 2004) ha ido creando y depurando una herramienta de trabajo que como él mismo comenta es tan discutible y buena como otra cualquiera y está pensada para

ser aplicada a páginas web con un propósito principalmente informativo.

Lo importante es que diferencia y establece categorías: nos habla de parámetros, indicadores y procedimientos y los engarza creando un sistema coherente.

Los parámetros responden a la pregunta de qué queremos evaluar: “son las propiedades o dimensiones del recurso que van a ser analizadas o consideradas” (Codina, 2003, p. 3). Son 9 y se agrupan en 4 secciones: contenido, navegación (dentro de la propia página web o micronavegación), visibilidad (inserción en la web o macronavegación) y usabilidad.

1. Autoría/Fuente: identificación y solvencia de la fuente
2. Contenido: calidad y cantidad de la información
3. Acceso a la información: navegación y recuperación
4. Ergonomía: comodidad y facilidad de utilización
5. Luminosidad: número de enlaces hacia otras sedes web
6. Ubicuidad: facilidad relativa que presenta una sede web para ser localizada y encontrada en el contexto de la Web
7. Procesos: dónde nos encontramos y en qué momento
8. Errores: respuesta (y alternativas) ante fallos
9. Adaptación.

Los indicadores nos dicen cómo evaluamos los parámetros (por lo que cada uno de ellos está asociados a uno de los parámetros e identificado por un número que incluye el del parámetro) y llevan asociados uno o varios procedimientos en los que se detallan los pasos concretos que se dan para analizar cada indicador.

Establece cuáles son los indicadores de cada parámetro.

Realiza una ficha para cada indicador en la que incluye (como se ve en el ejemplo):

- definición: cuál es el indicador
- examen: qué preguntas debe hacerse el evaluador para determinar su calidad

<sup>12</sup> FORNAS CARRASCO, Ricardo: Criterios para evaluar la calidad y fiabilidad de los contenidos en Internet, Rev. Esp. Doc. Cient., 26, 1, 2003. En <http://redc.revistas.csic.es/index.php/redc/article/viewFile/226/282> [consulta: 07-09 2009].

- ejemplos: qué recursos digitales ayudan a entender lo que se debe evaluar o aclaraciones sobre lo que se debe evaluar. También puede incluir ejemplos de lo que no debe hacerse.
- procedimiento: orientaciones sobre cómo proceder
- puntuación: escala de puntuación recomendada para cada indicador: cuando se mide si algo se cumple o no: 0 ó 1; si no, recomienda una escala del 0 (error grave) al 3 (excelente)<sup>13</sup>.

## Búsquedas y buscadores: más allá de Google

Una vez identificado y precisado el tema y la tecnología a vigilar, en una segunda etapa se elabora e implementa la estrategia de búsqueda y captación de la información. Para ello se realizan las siguientes actividades.

- a) Identificación de las palabras clave asociadas al tema a vigilar (tecnología), las cuales permitirán hacer la búsqueda en las diferentes bases de datos. Aquí se puede acudir a tesauros ya contruidos, a diccionarios propios y/o a expertos.
- b) Validación por expertos, en el tema específico estudiado, de las palabras clave seleccionadas. Este medio de validación garantiza que la información buscada mediante las palabras conducirá a los documentos (artículos científicos, patentes, etc.) en que se registran los avances específicos del conocimiento y/o las tecnologías relacionadas con el problema relevante estudiado.
- c) Identificación y selección de las fuentes de información relevantes. Estas pueden ser estructuradas (bases de datos de patentes y de artículos científicos), semi- estructuradas (websites, mail, foros, chat) y no estructuradas (documentos y reportes, noticias).
- d) Formulación de la ecuación de búsqueda. Una vez definidas las palabras clave y seleccionadas las fuentes de información de consulta, que tienen su propio lenguaje de interrogación –o lenguaje de comandos– para la recuperación de la información consignada en ellas y que el usuario debe aprenderse formula la ecuación

de búsqueda. Esta contiene, de manera rigurosa, los textos en los que se registran y articulan las palabras precisas y clave para buscar en las bases de datos (fuentes de información consultadas) la información requerida.

- e) Elaboración del contenido. A partir de la ecuación de búsqueda, definida y validada en el paso anterior, se realizan las consultas en las bases de datos seleccionadas y se extrae de ellas la información relevante. El producto de tales procesos es el corpus, es decir, el conjunto extenso y ordenado de los registros de las bases de datos que contienen la información útil para el estudio.

## Las búsquedas

### Búsqueda semántica

La extracción del significado para adquirir un mayor nivel de abstracción se presenta “como lo último” de las búsquedas. Cuando se puede extraer, abstraer y manejar el significado de los objetos indexados, las búsquedas pueden afinarse mucho más que con los mecanismos convencionales. Sin embargo esa es sólo una parte del problema: también hay que ser capaces de extraer el significado (y aún diríamos que la “intención”) de lo que quiere decir el usuario.

Aunque la tecnología semántica se presenta como el elemento clave en la evolución de la Internet del futuro (incluso se propone como el sustrato sobre el que se asentaría la llamada Web 3.0), aún es difícil encontrarla en servicios reales. Todos aquellos que claman basarse en el uso de tecnología semántica, se dedican a mercados verticales, muy orientados a aplicaciones muy concretas como la búsqueda de personas (como Spock <<http://www.spock.com>>), sin que los resultados sean tan espectacularmente precisos como cabría esperar.

Los buscadores verticales, especializados, y centrados en un tipo concreto de contenidos están teniendo un gran desarrollo, aunque por su propia naturaleza es difícil que supongan una alternativa generalista.

<sup>13</sup> Fuente: La evaluación de páginas web, Andoni Calderón. <http://observatorio.cnice.mec.es/modules.php?op=modload&name=News&file=article&sid=334> [consulta 11/09/09].

La Web Semántica<sup>14</sup> se puede definir como un marco que permite publicar, compartir y reutilizar datos y conocimiento tanto en la red como a través de aplicaciones. Dentro de este marco, se encuentran.

1. XML, que provee una sintaxis elemental para estructura el contenido dentro de los documentos, pero sin asociar ningún tipo de semántica al mismo.
2. RDF, un lenguaje que permite expresar modelos de datos, tanto las descripciones de los objetos (recursos) como las relaciones entre los mismos. Los modelos basados en RDF se pueden representar en sintaxis XML.
3. OWL, que añade más vocabulario para describir propiedades y clases como las relaciones entre clases (por ejemplo que sean disjuntas), cardinalidad (por ejemplo que sean exactamente 3), igualdad, características de las propiedades (por ejemplo simetría) y clases enumeradas.
4. SPARQL, un protocolo y lenguaje de consultas para recursos de la Web semántica.
5. Ontologías, que definen conceptos y relaciones entre los mismos, como FoaF (Friend of a Friend)<sup>15</sup>, que es una ontología basada en RDF que permite modelar la información de personas y las relaciones entre las mismas.

#### Búsqueda personal

En el otro extremo está la tendencia hacia incorporar un matiz más humano en las búsquedas. En realidad es basa en que el análisis y clasificación semántica se realicen por medio de personas, o al menos complementado con su visión. Se trata de una acción que tiene relación con la llamada Inteligencia Artificial Artificial, que pone a personas reales a realizar procesos generalmente mecanizados insatisfactoriamente: desde clasificar imágenes, hasta interpretar el sentido de la búsqueda de un usuario. Con las limitaciones propias de un sistema con una fuerte componente manual (por ejemplo, una menor escalabilidad), ejemplos como Mahalo <<http://www.mahalo.com>> están para demostrar las carencias de los algoritmos automáticos.

El etiquetado es una primera acción que ayuda a una clasificación más precisa. La anotación, el comentario sobre contenidos (como el reciente ejemplo de Google Reader) amplía los horizontes del análisis automático partiendo de acciones personales.

El matiz personal en las búsquedas está siempre presente en el análisis de la interacción de los usuarios que da lugar a la generación automática de perfiles, y la inclusión de recomendaciones. La corrección de resultados en función de la reacción y preferencias abre la puerta a otra modalidad que combina uno de los movimientos más poderosos en la red.

#### Búsqueda social

Los principios de esta modalidad son la identificación de perfiles afines a partir de la red social (amigos, conocidos, otras personas con intereses similares) y modular los resultados de acuerdo con las acciones previas de los miembros de esa red. Se trata, como se ve, de una forma de complementar los algoritmos con un matiz personal, si bien en este caso el toque humano tiene una procedencia más diluida. Google, aparentemente investigando en esta línea, no ha creado una red social alternativa, pero sí trabaja en herramientas de identificación y análisis de la red de un usuario. Otras herramientas tratan de construir automáticamente (y no explícitamente) esa red, partiendo del análisis del intercambio de correo, por ejemplo. Esas redes sociales implícitas, y generadas por medio de algoritmos, pueden ser la pieza que ayude a complementar a los sistemas automáticos y de gran capacidad aunque aún con pobre calidad.

#### Búsqueda y referencia

Es difícil encontrar un término que defina esa combinación de Wikipedia y Google que para muchos tiene el potencial de cambiar el mundo de las búsquedas. Wikipedia tiene una enorme cantidad de información, generada por sus propios usuarios, y que pese a polémicas, disputas y controversias, es asumida como una fuente de referencia por mu-

<sup>14</sup> La Búsqueda Semántica [http://weblogs.madrimasd.org/sistemas\\_inteligentes/archive/2008/05/02/90703.aspx](http://weblogs.madrimasd.org/sistemas_inteligentes/archive/2008/05/02/90703.aspx) [consulta 30-08-09].

<sup>15</sup> FOAF, 2007. The Friend of a Friend (FOAF) Project. Sitio Web accesible en: <http://www.foaf-project.org/>. [consulta: 3/09/09].

chísimas personas. ¿Y si las búsquedas mostraran no miles de resultados que exigen una labor adicional de investigación y depuración, si no que apuntaran a resultados completos, elaborados y sintetizados como los de la Wikipedia? Para muchos usuarios este escenario define a la herramienta que están esperando.

Wikia Search <<http://answers.wikia.com>> o Wikiasari sería un movimiento en este sentido, proporcionado un sistema de búsquedas apoyado fundamentalmente en los contenidos de la Wikipedia. El servicio Knol <<http://knol.google.com>> de Google, aún en construcción, sería una respuesta en forma de una Wikipedia con autores reconocidos y contenidos fiables. La principal limitación está en el volumen de información, ya que es imposible competir con expertos a la legión de voluntarios que contribuyen en la Wikipedia. Cuando se haga público podremos valorar lo acertado de este concepto, nada novedoso si consideramos otras “enciclopedias” en la red, como Encarta o la versión electrónica de la Britannica.

Otros frentes en el desarrollo de las búsquedas tienen que ver con el **lenguaje natural** y la **traducción automática** de contenidos. En ambos Google está haciendo una fuerte inversión dada su vocación de servicio global. La traducción automática ampliaría notablemente los contenidos disponibles, y un lógico movimiento de futuro ahora que el país con más usuarios conectados es China, y cabe esperar una reducción en el peso relativo de los contenidos en inglés. Sin embargo, la traducción automática tiene un largo y frustrante historial y ha engullido enormes inversiones sin alcanzar resultados satisfactorios. El factor diferencial introducido por Google (el análisis comparativo de una enorme base de textos en varios idiomas) es muy novedosa y siendo realistas algo que sólo está en manos de esta empresa.

El lenguaje natural mejora las posibilidades de interacción con el usuario, y cuando se trata de la interpretación vocal amplía los horizontes de las aplicaciones preparando su salto al móvil y entornos donde el uso de pantallas está muy restringido, como en el caso del coche. A día de hoy, los proyectos de lenguaje natural y traducción automática

son los que están recibiendo más atención en Google Research<sup>16</sup>.

## Los buscadores

### Buscadores especializados y temáticos

Ofrecen resultados de un ámbito determinado de actividad.

Son más útiles que los buscadores genéricos cuando buscamos una tipología determinada.

Elimina el “**ruido**” que introduce otro tipo de páginas web.

Poseen un **alto grado de parametrización**, por ejemplo, en Google Scholar podemos buscar sólo artículos recientes o en Technorati blogs con un determinado nivel de “autoridad”.

- Noticias. Google News <<http://news.google.es>>
- Material Académicos. Google Scholar <<http://scholar.google.es>>
- Blogs. Technorati <<http://technorati.com>>

### Buscadores que aportan innovación en la forma de presentar los resultados de búsqueda

Estos buscadores presentan el resultado con imágenes, con árboles de conocimiento relacionados entre sí, dando más información en el enlace que proporcionan, etc. Entre estos buscadores nos han llamado la atención los siguientes:

- Oskope <<http://www.oskope.com>>: Buscador muy intuitivo, que ofrece varias formas para mostrar los resultados de búsqueda (lista, grid, gráfico, etc). No hace una búsqueda global, permite buscar en amazon, ebay, flickr.com, youtube, etc.
- Picitup <<http://www.picitup.com>>: Este buscador directamente sirve para buscar fotos en la red, por lo que no es tan innovador. La parte interesante es que la búsqueda se puede ir refinando: buscar fotos similares, paisajes, colores, caras, etc.
- Searchme <<http://beta.searchme.com>>: De momento esta versión beta muestra una captura de cada página y permite filtrar la búsqueda por temas que te ofrece el propio buscador.
- Grokker <<http://www.grokker.com>>: Ofrece en una columna un listado de resultados agrupa-

<sup>16</sup> Fuente: La Cofa <http://www.lacofa.es/index.php/tecnologias/futuro-de-internet/espectativas-de-evolucion-en-los-buscadores> [Consulta: 08-09-2009].

dos en bloques según subtemas que considera el buscador y en otra columna detalla los resultados. También ofrece una vista de un mapa de resultados. La contrapartida es que parece algo lento.

- Search Monkey <<http://developer.yahoo.com/searchmonkey>>: Esta es una apuesta de Yahoo para permitir a los desarrolladores y dueños de sitios modificar las búsquedas de Yahoo!, pudiendo mostrar los resultados de la forma más útil o modificando el aspecto visual de los mismos.
- KartOO <<http://www.kartoo.com>>: Co-fundada en Francia en 2001, tla compañía homónima lanzó en 2002, KartOO un meta motor de búsqueda implementado en Flash que proporciona un interfaz visual para explorar los resultados. Una vez entrada la búsqueda obtienes un “mapa” en vez de una lista. Este mapa visualiza los resultados como entidades y las palabras clave como relaciones. Los sitios web se representan mediante una captura de pantalla en baja resolución con la URL. Además, el tamaño se utiliza como atributo visual para representar la relevancia del sitio dentro de la interrogación.

Las palabras relevantes se usan también para relacionar diferentes resultados que se muestran en el mapa y, además, para permitir la posibilidad de redefinir la búsqueda añadiéndolas a la interrogación inicial. La propia información relacional produce curvas de nivel utilizando la metáfora del paisaje que se encuentra en los mapas cartográficos. La herramienta utiliza la clausura como vehículo para permitir al usuario ver cuáles son las relaciones entre objetos de una forma más clara que con enlaces.

Aquí la interacción involucra al usuario en el proceso de búsqueda y le proporciona medios de exploración.

Kartoo tiene otros interesantes proyectos de búsqueda visual como es Ujiko <<http://www.ujiko.com>>. Su display muestra los resultados en un menú oval con una sección central abierta donde el usuario puede marcar determinados resultados como relevantes. La sección central se usa para añadir información a los resultados. Mediante bloques coloreados se ayuda al usuario a identificar agrupaciones (clusters) nominales.

Otra interesante característica de este motor de búsqueda es que cuanto más lo usas más funciones es capaz de proporcionarte. Cada vez que visitas un sitio web ganas un punto de experiencia. Con cada 10 puntos promocionas al siguiente nivel consiguiendo nuevas funcionalidades. Esta innovadora característica demuestra el deseo de Kartoo de obtener información sobre la interacción del usuario introduciendo su aplicación como si de un juego se tratase.

A pesar de estas mejoras, las nuevas interfaces heredan los mismos viejos y básicos problemas sin proporcionar en cambio avances significativos en la interacción. La explosión de los multimedia y las nuevas necesidades de información no hace más que incrementar la necesidad de disponer de interfaces visualmente ricas, más sofisticadas y complejas capaces de buscar en contenidos multimedia interrelacionados. Hace ya bastante tiempo han aparecido algunos navegadores visuales que intentan superar las limitaciones de la presentación y la interacción en los motores de búsqueda tradicionales.

Hasta ahora estos tipos de navegadores y motores de búsqueda visual se habían considerado aplicaciones online más cercanas a lo artístico que a lo utilizable. Puede ser útil el artículo “Motores de búsqueda visuales” de la revista Infovis: <http://www.infovis.net/printMag.php?num=198&lang=1>

**Buscadores que buscan innovación en la manera de solicitar la búsqueda,** que tradicionalmente se ha hecho mediante cadenas de “palabras clave”:

Entre estos buscadores están los buscadores vocales, que permiten introducir el input a través de la voz. Entre estos ya existen iniciativas de Yahoo con mejoras en oneSearch y de Microsoft con mejoras en Microsoft Live Search.

Otro tipo de buscadores interesantes serían los que tienen como input una imagen. Vodafone ya tiene lista una propuesta, con su buscador Otello. El funcionamiento de este buscador será el siguiente: haces una foto con tu terminal, la envías por MMS al número de servicio correspondiente a Otello y recibes como respuesta información relevante asociada a esa imagen.

Este tipo de buscadores pueden ser muy interesantes, especialmente cuando no se usa un ordenador para realizar las búsquedas. En concreto,

parecen especialmente interesantes cuando la búsqueda se realiza desde un móvil y de hecho las iniciativas analizadas van claramente encaminadas a su uso desde terminales móviles. Hay que tener en cuenta que escribir en un móvil es tedioso (aunque incorpore un teclado qwerty) por lo que poder introducir la información de manera vocal o mediante una foto puede ser una opción muy atractiva para usuarios con este tipo de terminales. Si se consigue que estos usuarios utilicen este tipo de buscadores, se les podrá ofrecer otros servicios combinados que junto con el buscador permitan el uso de otras capacidades de telecomunicaciones.

### Las fuentes: Gratuitas vs Pago

No hace tanto tiempo la búsqueda de información para hacer un Informe de Vigilancia requería estar suscrito a Bases de Datos muy caras y exclusivas, donde estaba todo lo que había disponible.

Bases de Datos como por ejemplo Dialog (ahora Grupo Thomson) eran las únicas fuentes donde encontrar determinadas cosas.

Hoy en día, ya no es así. Las patentes, las tesis doctorales, las hemerotecas de los periódicos (la última incorporación es la hemeroteca de ABC que incluye la publicación Blanco y Negro <http://hemeroteca.abc.es/>), los movimientos de productos en las aduanas, la legislación, son en su gran mayoría información gratuita en Internet. Lo único que hay que tener es paciencia y método y realizar periódicamente un barrido de fuentes para tenerlas todas localizables.

Ejemplo de ello son el BOE, los Boletines de las CCAA, las BD del ICEX, la BD de patentes de los diferentes países, etc.

Los principales proveedores de información (Thomson, Elsevier, EBSCO, Chemical Abstract...) están incorporando a sus bases de datos herramientas que permiten el tratamiento básico de la información (recuentos de la frecuencia de coaparición de palabras), el análisis más complejo de la información a través de co-ocurrencias de palabras, clusterización, y visualizaciones, que aportan mayor

valor a los resultados de una búsqueda, ofreciendo de esta manera un valor añadido que justifique la suscripción.

### Reflexiones

Las búsquedas complejas están condenadas a desaparecer. Las grandes Bases de Datos ya no tienen la exclusiva de la información. Hay otros métodos para obtener la información que deseamos que antes eran impensables.

Por otro lado las pymes no necesitan un gran software de Vigilancia Tecnológica. Lo que hoy ofrece Internet de forma más o gratuita o a un coste muy bajo es más que suficiente para apoyo de un servicio de estas características.

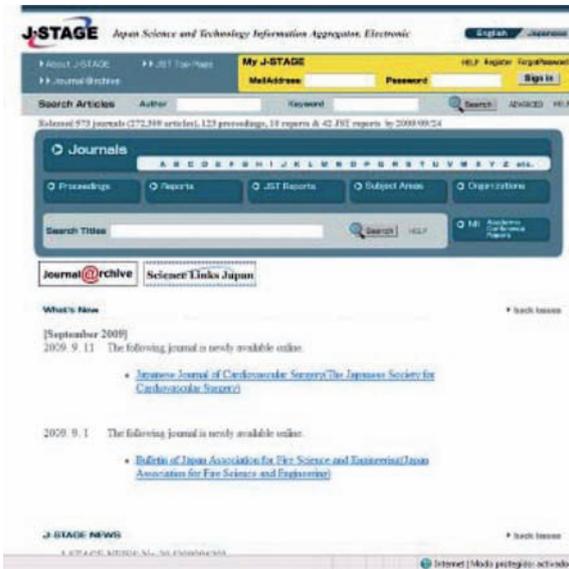
A día de hoy, y si la tendencia sigue como hasta ahora, será la unión de diferentes herramientas las que conformen el sistema.

Muchas de estas herramientas nos ayudan a aglutinar una red de contactos que sirven para añadir un tipo de información más subjetiva, más allá de las grandes búsquedas. Esta red de contactos a través de las redes sociales, unido a otro tipo de herramientas valorizan las búsquedas y dan valor a los Informes de Vigilancia, así como te mantienen al día de nuevas fuentes. Si se sabe montar el puzle de todos estos servicios (twitter, linkedin, lector de rss, buscadores especializados, rastreadores de webs, alertas de noticias, etc...) se crearía un servicio de Vigilancia a un precio más que razonable.

<b>Thomson:</b> Es el líder mundial como proveedor de recursos y productos de información alta calidad, en casi todas las áreas de conocimiento, asociados a servicios de alto valor añadido, orientados a todo tipo de usuarios: investigadores, empresarios, inventores, estudiantes, políticos, economistas.	
ISI Web of Knowledge	Asocia a los resultados de las búsquedas, los recuentos simples de los campos más importantes.
Thomson Data Analyzer y Aureka	Thomson Data Analyzer -derivado del software original Derwent AnalyticsSM - trabaja con bases de datos de patentes y publicaciones científicas (compatible con las de Thomson y con otras) y permite caracterizar compañías (sus carteras de productos actuales y potenciales) así como evaluar áreas tecnológicas.  Aureka 9.2, es una herramienta de análisis y gestión de la Propiedad Industrial e Intelectual a través de un entorno web.
Delphion Snapshot y Clustering	La base de datos tiene aplicaciones que se basan en técnicas de minería de datos: PatentLab-II, Snapshot y Clustering.  Snapshot es una herramienta analítica que resume la información bibliográfica seleccionada de la patente, a partir de un corpus obtenido de una búsqueda cualquiera o de una carpeta de trabajo guardada. Una ayuda de este tipo permite al usuario centrar la investigación en un grupo de patentes más aproximado a sus necesidades.  Clustering es una herramienta que se apoya en tecnologías relacionales y en la lingüística para el análisis de documentos de patentes asociados a un corpus o a un archivo de trabajo. Los agrupa de acuerdo a las palabras claves identificadas durante el análisis en el título y el resumen de cada patente. Las patentes son asignadas únicamente a un cluster.
<b>Engineering Village:</b> Este proveedor también ha invertido en ofrecer un mejor servicio a través de sus bases de datos: Compendex, Inspec... El recuento simple de los campos más importantes para los resultados de una búsqueda determinada.	
<b>SciFinder:</b> Software asociado a la base de datos Chemical Abstracts (CAS). Por ejemplo, SciFinder es una herramienta que permite la búsqueda en lenguaje natural.	
<b>STNAnavist 2.0:</b> STNAnavist 2.0 es una herramienta asociada a STN.	

Ejemplos de Bases de Datos de contenido completo y gratuito

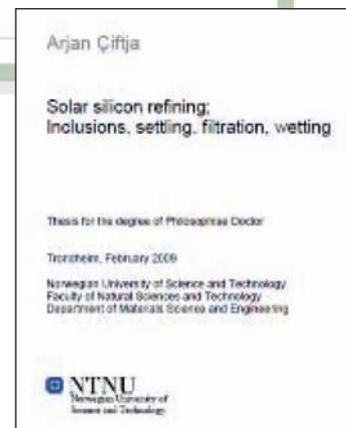
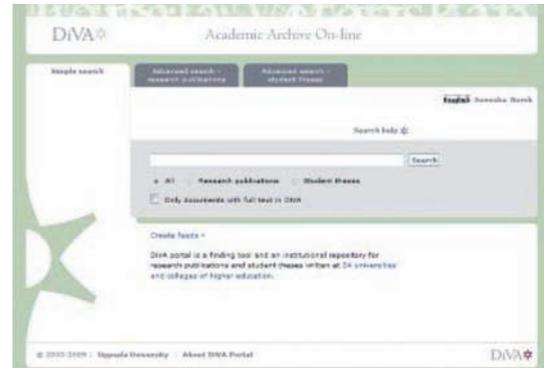
- Japan Science and Technology Information Aggregator  
<http://www.jstage.jst.go.jp/browse/>



- Canadian Film Encyclopedia Search  
<http://www.filmreferencelibrary.ca/index.asp?navid=46>



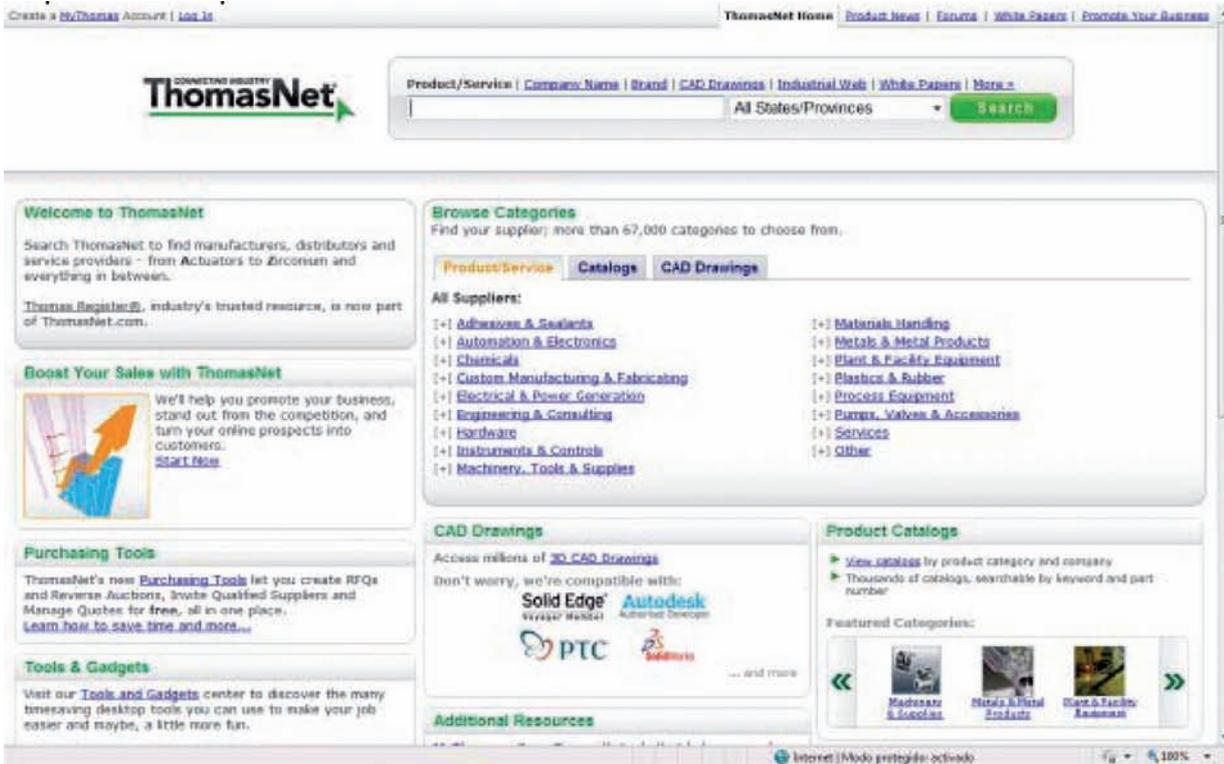
- Academic Archive On-line  
<http://www.diva-portal.org/index.xsql?lang=es>  
 Búsqueda de tesis doctorales y algún otro tipo de publicación del tipo de tesinas, informes...  
 Limitado a las universidades que aparecen en el menú. Muchos resultados están a texto completo.



• Thomas Register of America Manufacturers

<http://www.thomasnet.com/>

Información sobre productos, empresas y marcas de EEUU y Canadá. Gratuito, pero requiere suscripción.



• Estadísticas de Extremadura

<http://www.estadisticaextremadura.com/>

• Distribución territorial de apellidos (Datos del Padrón a 1-1-2006)

<http://www.ine.es/fapel/FAPEL.INICIO>



• **Bibliothèque nationale de France**  
[http://www.bnf.fr/pages/zNavigat/frame/version\\_espagnole.htm?ancre=espanjol.htm](http://www.bnf.fr/pages/zNavigat/frame/version_espagnole.htm?ancre=espanjol.htm)



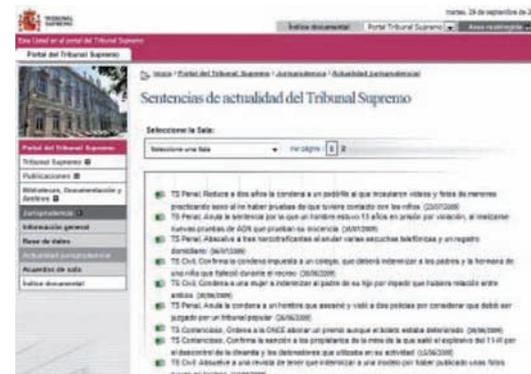
• **Patent Scope**  
 Search International Patent Applications.  
<http://www.wipo.int/pctdb/en/>



• **Fuentes de búsqueda de Tesis doctorales:**  
 Tesis de la Universidad de Barcelona. <http://www.bib.ub.edu/recursos-informacio/tesis-doctorals/>



• **Sentencias del Tribunal Supremo.**  
<http://www.poderjudicial.es/eversuite/GetRecords?Template=cgpi/ts/principal.htm>



• **Project MUSE.**  
<http://muse.jhu.edu/>  
 Project Muse ofrece a texto completo, el acceso asequible a los contenidos actuales de las humanidades y las ciencias sociales de prestigiosos diarios. MUSE es la fuente más fiable de los títulos de muchas de las editoriales universitarias más importantes del mundo y sociedades científicas, revistas con artículos aclamada por la crítica de los académicos más respetados en sus campos.



# Las patentes

*“Cuando encuentras un diamante que no es de nadie, es tuyo. Cuando encuentras una isla que no es de nadie, es tuya. Cuando eres el primero en tener una idea, la haces patentar, es tuya. Yo poseo las estrellas porque jamás, nadie antes que yo soñó poseerlas”*

“El Principito” Antoine de Saint-Exupéry (fuente de la cita: OEPM).

## Las patentes como proceso/trámite

Las patentes constituyen uno de los elementos jurídicos dirigidos a promover el desarrollo tecnológico.

Permiten a los creadores de nuevas tecnologías apropiarse de los resultados de sus esfuerzos in-

ventivos, y los protege de la competencia de potenciales usuarios de las mismas tecnologías.

Es un poderoso instrumento comercial para que las empresas obtengan exclusividad sobre un nuevo producto o proceso que da una sólida posición en el mercado y ventaja competitiva con un aumento de los beneficios o del rendimiento de las inversiones.

Participación en nuevos mercados.

Reducción de los riesgos de infracción.

Es una imagen positiva para una empresa ya que patente se asocia con i+d o con innovación.

Incremento de la capacidad para obtener financiamiento.

The screenshot shows the 'Ciencia e Investigación Hechos y cifras' page on the L'Oréal España website. The page features a list of key statistics about their R&D efforts:

- 3.268 colaboraciones de 60 nacionalidades distintas trabajando en 33 disciplinas diferentes.
- 604 millones de € destinados a investigación dermatológica y cosmética en 2016.
- 10 centros de investigación en todo el mundo y 12 centros de evaluación.
- 1/3 del presupuesto de I+D dedicado a Investigación Avanzada.
- 5.000 hitos nuevos desarrollados cada año.
- 608 patentes registradas en 2016.
- 100 acuerdos de cooperación activos con instituciones académicas y de investigación.

The page also includes a navigation menu with categories like PIEL, CABELLO, OJOS, and INNOVACIÓN RESPONSABLE.

La empresa L’Oreal utiliza el número de patentes como reclamo publicitario para demostrar la capacidad de investigación que tienen y lo punteros que son en su sector.

Algunos datos a tener en cuenta:  
Solicitud de Patentes en España.

Patentes del año 2006: 21.486 patentes.  
Patentes del año 2007: 26.621 patentes.

Ranking de inventores con más inventos patentados.
<ul style="list-style-type: none"><li>• RODRIGUEZ FERRE, JOSE MANUEL (172 patentes)</li><li>• GUERET, JEAN-LOUIS (163 patentes)</li><li>• THEURER, JOSEF (115 patentes)</li><li>• MUR GIMENO, EMILIO (112 patentes)</li><li>• AMMERMANN, EBERHARD, DR. (107 patentes)</li><li>• DAUMAL CASTELLON, MELCHOR (94 patentes)</li><li>• PEDAIN, JOSEF, DR. (92 patentes)</li><li>• JANE CABAGNERO, RAMON (86 patentes)</li><li>• LORENZ, GISELA, DR. (85 patentes)</li><li>• RENARD, PIERRE (84 patentes)</li></ul>
<a href="#">Listado completo de inventores</a>

Ranking de empresas con más inventos patentados.
<ul style="list-style-type: none"><li>• BASF AKTIENGESELLSCHAFT (3722 patentes)</li><li>• SIEMENS AKTIENGESELLSCHAFT (3254 patentes)</li><li>• BAYER AKTIENGESELLSCHAFT (2957 patentes)</li><li>• HOECHST AKTIENGESELLSCHAFT (2421 patentes)</li><li>• THE PROCTER &amp; GAMBLE COMPANY (2238 patentes)</li><li>• L'OREAL (2212 patentes)</li><li>• BAYER AG (2045 patentes)</li><li>• WESTINGHOUSE ELECTRIC CORPORATION (2028 patentes)</li><li>• ROBERT BOSCH GMBH (1942 patentes)</li><li>• GENERAL ELECTRIC COMPANY (1712 patentes)</li></ul>
<a href="#">Listado completo de empresas</a>

Fuente: <http://patentados.com/>

## Las Patentes como fuente de información

“Las patentes no sólo son útiles como instrumento de protección, sino también como fuente de información técnica para las empresas y grupos de investigación que poseen capacidad para desarrollar tecnología propia. Los informes tecnológicos de patentes (ITP) también son una buena fuente de información para todo tipo de empresas que quieren evitar infringir patentes anteriores cuando introducen sus productos en el mercado, para aquellas otras que optan por trabajar con tecnología del dominio público, o para las que prefieren detectar tecnologías de las que ser licenciarios. Pueden además proporcionar información sobre las actividades de la competencia. Frecuentemente son solicitados para analizar la singularidad de una tecnología concreta, ya que la información técnica contenida en las patentes permite conocer el estado de la técnica existente en un sector técnico concreto y puede, por tanto, ser determinante a la hora de comenzar o replantear un determinado proyecto de investigación, evitando así costes innecesarios. Esta es la razón por la que tales informes tecnológicos deberían resultar de gran interés para las empresas e imprescindibles para universidades y centros públicos de investigación en las fases incipientes de planificación de nuevos proyectos de investigación”. (...).

Interpretar la información técnica y jurídica que contiene una patente no es tarea sencilla. Aunque una patente, ya sea alemana, inglesa, norteamericana o de cualquier otro país, siempre presenta la misma estructura, y esto claramente simplifica su consulta, su contenido frecuentemente no resulta claro para alguien que no está habituado a utilizar este tipo de documentos. Por otra parte, localizar aquellas patentes que por un motivo u otro estén relacionadas con una cuestión técnica que le puede surgir a una empresa en un momento dado, a menudo puede resultar más complejo de lo inicialmente previsto, ya que requiere que quien haga la búsqueda, además de entender el problema técnico que se plantea, ha de conocer bien el contenido y funcionamiento de las bases de datos existentes en el mundo sobre patentes. Los informes tecnológicos de patentes pretenden ahorrar a la empresa estas tareas de búsqueda e interpretación de patentes y ofrecer una información suficientemente clara con relación a la cuestión técnica planteada.

Esta es la razón de que tales informes contengan un análisis detallado de las patentes y otros tipos de documentos publicados en el mundo relativos al tema en cuestión, incluyendo además las copias de las patentes y otros documentos considerados por el técnico responsable del mismo especialmente significativos; también incorporan los listados de referencias bibliográficas de todas las patentes recuperadas, ya que pueden ofrecer información adicional de interés. Hay que tener presente que las patentes, desde que se solicitan hasta que se publican, al menos dieciocho meses más tarde, no son consultables.

Los Informes Tecnológicos de Patentes, Fundación Cotec, pp. 34-35.  
Disponible desde <http://sites.google.com/site/lararey/patentes>

No obstante, este uso de las patentes lleva asociado de forma implícita un conjunto de dificultades que es necesario tener en cuenta:

- Sólo una parte de las innovaciones se patentan, ya que hay otros mecanismos diferentes para proteger las invenciones, por ejemplo, el secreto industrial.
- Algunas invenciones no son patentables, o porque la legislación las excluye expresamente o porque no cumplen alguno de los requisitos necesarios (novedad, actividad inventiva, aplicación industrial).
- En determinados sectores el ciclo de vida de los productos es muy corto y no es interesante patentar<sup>17</sup>.

#### **Ventajas de patentar:**

- Promueve la transferencia de tecnología y la inversión;
- Fomenta la investigación y el desarrollo en las universidades y en los centros de investigación.
- Genera nuevas tecnologías y empresas;
- Las empresas acumulan patentes y las utilizan en actividades de concesión de licencias, empresas conjuntas y otras iniciativas de generación de ingresos.

#### **Ventajas de utilizar las patentes como fuentes de información:**

- Identificación de socios técnicos comerciales.
- Identificación de mercados tecnológicos.
- Retribución económica para sus desarrolladores.
- Identificación de competidores comerciales.
- Protección de patrimonios tecnológicos.
- Evitar invasión de derechos adquiridos.
- Dinamismo en el proceso de transferencia de tecnología.
- Propiciar licencias y sus consiguientes royalties de derechos tecnológicos.
- Resolver problemas técnicos en las empresas.
- Encontrar soluciones tecnológicas en sectores productivos.
- Apoyar proyectos de investigación y mejoras tecnológicas ya existentes.

- Evitar esfuerzos innecesarios en la solución de problemas ya resueltos.

El análisis de los datos bibliográficos extraídos de un documento de patente puede proporcionar información muy valiosa. Por ejemplo, un análisis de la relación entre la tasa de denegación de patentes frente al número de solicitudes realizadas y sobre las causas que provocaron dicha denegación (qué requisito de patentabilidad falló), puede aportar una excelente información acerca de hasta qué punto las empresas, las universidades y los centros de investigación llevan a cabo labores de vigilancia tecnológica y cómo utilizan esta información antes de aportar fondos a proyectos de investigación<sup>18</sup>.

#### **Trucos para entender un poquito mejor las patentes**

No todas las patentes tienen el mismo valor económico. Las consideradas de mayor valor comercial son las patentes triádicas, es decir, las concedidas con efectos conjuntos en las oficinas europeas, estadounidenses y japonesas.

Está generalmente admitido que el número de solicitudes de patentes originadas en un país constituye un indicador bastante significativo de la situación de su sistema de ciencia-tecnología. Al mismo tiempo, existe entre los medios especializados el convencimiento de que el número de solicitudes de patentes depositadas por españoles o residentes en España es menor de lo que se correspondería con su nivel de desarrollo económico o industrial (Fuente Informe Cotec 2007 sobre Tecnología e Innovación en España <<http://www.cotec.es>>).

#### **Estructura de los documentos de Patente**

##### **IET (Informe del Estado de la Técnica)**

Estudio que indica el grado de novedad y de actividad inventiva de la solicitud de patente. Nunca valora la utilidad o posibles mejoras técnicas que la invención suponga.

Se hace comparando la solicitud con otras patentes (nacionales y extranjeras) y con literatura no

<sup>17</sup> HIDALGO-NUCHERA, Antonio; IGLESIAS-PRADAS, Santiago y HERNÁNDEZ-GARCÍA, Ángel: "Utilización de las bases de datos de patentes como instrumento de vigilancia tecnológica". *El profesional de la información*, 2009, septiembre-octubre, v. 18, n. 5, pp. 511-519.

<sup>18</sup> Ver nota anterior.

patente (técnica y científica) que reflejen el estado de la técnica.

Hace referencia al sector técnico de la invención según lo establecido en la C.I.P. (Clasificación Internacional de Patentes).

Valoración de la novedad y la actividad inventiva respecto las reivindicaciones:-Documento A: la invención reivindicada es nueva y tiene actividad inventiva-Documento X: la invención reivindicada no es nueva-.

Documento Y: la invención reivindicada no tiene actividad inventiva (normalmente con dos documentos).

El informe sobre el estado de la técnica es obligatorio. En él se recogen una serie de documentos relevantes clasificados según ésta categoría. Esta es la manera en que los documentos citados afectan al documento estudiado:

- X. El documento por si sólo afecta a la novedad o a la actividad inventiva.
- Y. varios documentos Y combinados afectan a la actividad inventiva (si dicha combinación es evidente para un experto en la materia)
- A. Documento relacionado con el estudiado, que no afecta a la actividad inventiva.
- O. Divulgación no escrita: descripción oral, utilización o cualquier otro medio.
- P. Documento publicado antes de la presentación de la solicitud, pero después de la fecha de prioridad del documento estudiado. Se indica si es X, Y o A.
- E. Documento publicado después de la presentación de la solicitud del documento estudiado, aunque solicitado antes que éste, y que afecta a la novedad.

Las categorías O, P y E son raramente utilizadas.

**Adición.** La adición de patente es gratuita. Expira con la patente principal. Compensa publicarla como patente propiamente dicha, cuando la primera lleva mucho tiempo publicada.

**Familia de patentes** Una misma invención puede estar patentada, p.ej., en 20 países, lo que da lugar a 20 títulos diferentes de propiedad industrial, es decir, 20 documentos de patente que son parte de la misma “familia”.

**Nomenclatura** que aparece al lado del número de Solicitud de patente o al lado del número de la patente.

- A1. Solicitud de patente con informe sobre el estado de la técnica.
- A2. Solicitud de patente sin informe sobre el estado de la técnica.
- R. Informe sobre el estado de la técnica publicado separadamente.
- B1. Patente con informe sobre el estado de la técnica.
- B2. Patente con examen previo.
- A6. Patente sin informe sobre el estado de la técnica.
- T1. Traducción de las reivindicaciones con dibujos de solicitudes europeas (antiguo A4).
- T2. Traducción revisada de las reivindicaciones con dibujos de solicitudes europeas.
- T3. Traducción de patente europea concedida (antiguo B3).
- T4. Traducción revisada de patente europea concedida.

#### Fecha de prioridad

Quien hubiere presentado regularmente una solicitud de patente de invención, de modelo de utilidad, de certificado de utilidad o de certificado de inventor en alguno de los países de la Unión para la Protección de la Propiedad Industrial o sus causahabientes gozarán, para la presentación de una solicitud de patente en España para la misma invención, del derecho de prioridad establecido en el Convenio de la Unión de París para la Protección de la Propiedad Industrial.

Es decir, se mantiene como fecha de prioridad la del país de origen.

#### Búsqueda de citas

Permiten localizar documentos relacionados entre sí, y del que pueden considerarse distintos tipos de citas:

- Patentes citadas en la memoria.
- Patentes de la misma familia y sus informes.
- Citas del examinador en el informe sobre el estado de la técnica.
- Patentes que citan una patente.

## Patentes-fuentes de información tecnológica

¿Por qué las patentes son una excelente fuente de información tecnológica?

1. Los documentos de patente tienen una estructura uniforme y armonizada internacionalmente.
2. La información que contienen es inédita.
3. Referencian tanto al investigador como el centro o empresa donde se ha realizado el proyecto que materializa la patente.
4. Muchas de ellas tienen un Informe sobre el estado de la técnica (sólo si el sistema de Concesión lo contempla), donde se recoge el grado de novedad y de actividad inventiva de la solicitud de patente en función de las reivindicaciones (las reivindicaciones es la parte donde se describe lo que realmente se quiere proteger).
5. Una clasificación internacional con alrededor de 70 000 códigos tecnológicos.
6. Fecha de prioridad en la que se recoge cuándo se ha solicitado la protección de la innovación (lo que permite estudiar evoluciones).
7. País de prioridad: Permite conocer dónde se están generando las innovaciones.
8. Países de extensión: Se localizan así los países en los que se están protegiendo las invenciones.

Otros documentos relacionados con un documento de patente

- Documento de patente propiamente dicho.
- Patentes que cita el solicitante en la memoria.
- Patentes de la misma familia y sus informes.
- Citas del examinador en el Informe del Estado de la Técnica.
- Patentes que citan a la que estamos consultando.

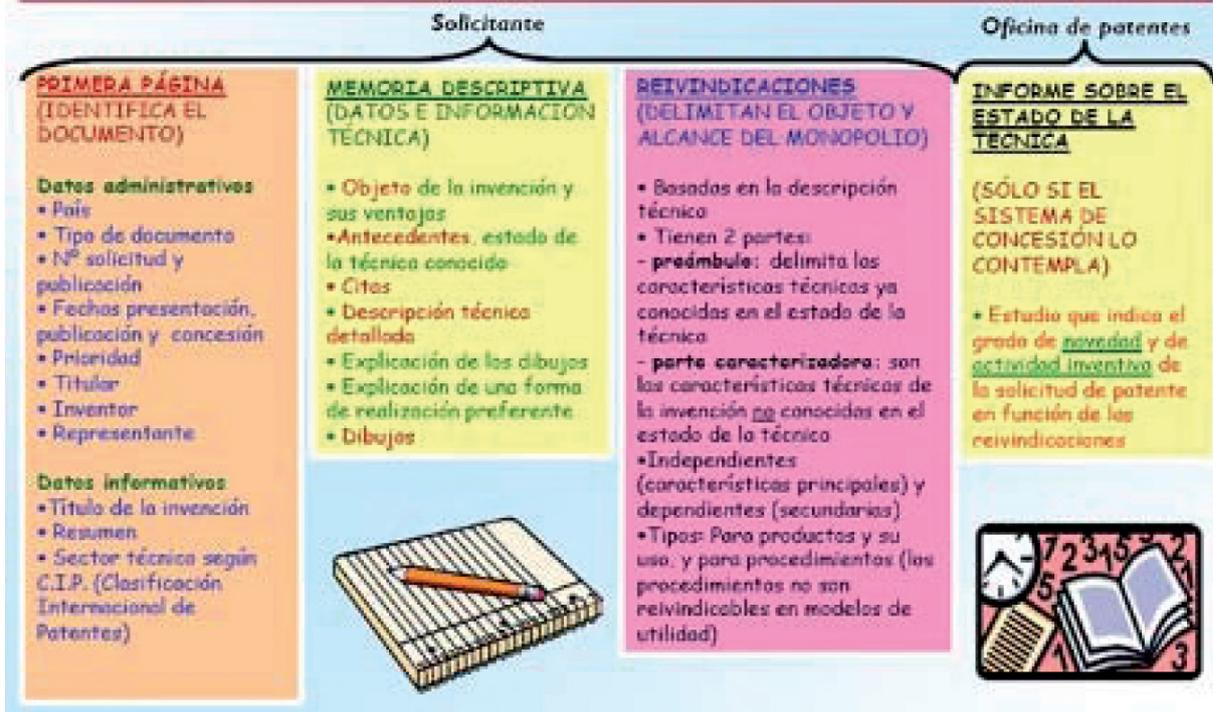
La gestión de la innovación tecnológica además de la actividad patentadora requiere una constante actualización de conocimientos a través de los cuales podemos conocer y vigilar el comportamiento innovador de los competidores. Estos objetivos justifican una búsqueda de patentes (Morin 1989).

- Conocimiento del entorno tecnológico en el que se mueve la empresa.
- Vigilancia de la capacidad de avance tecnológico de la competencia.
- Planificación de la innovación tecnológica: Según un estudio de la DG XIII de la CE, el 30% del gasto de I+D en el mundo se destina a proyectos de investigación sobre innovaciones ya patentadas.
- Resolución de un problema técnico concreto.
- Análisis de la patentabilidad de los desarrollos propios.
- Análisis de riesgos de infracción de patentes.
- Valoración de la tecnología sobre la que se discute en contrato de licencia.



Fuente: OEMP

## LA ESTRUCTURA DE LOS DOCUMENTOS DE PATENTE





Fuente: OEMP <http://www.oepm.es/cs/OEPMSite/contenidos/ponen/gip06/miercoles/miercoles4.pdf> [vinculo consultado por última vez el 11 de septiembre de 2009].

Para saber más sobre este tema

- Sobre patentes.- <http://wikipolitecnica.pbworks.com/Patentes>.
- Sobre las patentes. Patentes y transferencia de tecnología, Numero 49, octubre 2008.- <http://www.madrimasd.org/revista/revista49/editorial/editorial.asp>.
- Díaz Pérez M, Orea Igarza U, Cordero Machado E. Los análisis de patentes como base para la toma de decisiones en los proyectos de investigación.
- Estudio de un caso. Acimed 2007;16(2).
- Las patentes como fuente de informacion (OEPM).
- Borja González-Albo Manglano; Ma Ángeles Zulueta: Estudio comparativo de bases de datos de patentes en internet, en Anales de Documentacion, N° 10, 2007, pp. 145-162.
- López López, Pedro; García-Escudero, Patricia: Literatura de patentes: utilidad y fuentes de información, En: Boletín de la ANABAD, ISSN 0210-4164, Tomo 50, N° 1, 2000, pp. 107-128.

#### Sistemas gratuitos de búsqueda de patentes

**Espacenet:** Base de la Oficina Europea de Patentes. Contiene más de 60 millones de documentos de 71 países y 10 campos de búsqueda.

**Depatisnet:** Base de la Oficina Alemana de Patentes. Contiene más de 20 millones de documentos de 10 países y 30 campos de búsqueda. Permite buscar en el texto completo en algunas colecciones.

**Free Patents On-line:** Sistema que consulta de manera conjunta las colecciones de patentes US, EP, JP y PCT.

**Patents Lens:** Sistema que consulta de manera conjunta las colecciones de patentes US, EP, PCT y AU.

**Prior smart:** Sistema que consulta de manera conjunta de los servidores ESPACENET, LATIPAT, USAPAT, PATENTSCOPE; PATENT LENS.

**Patent scope:** Base de la OMPI con solicitudes PCT.

#### Sistemas de pago

**Derwent Patent Index:** Base de datos con información de patentes de más de 40 países. Es pro-

iedad del grupo Thomson y accesible a través de diversos distribuidores.

**Delphion:** Sistema integrado de búsqueda de patentes, incluye el acceso al Derwent Patent Index y a las colecciones de patentes US, EP; DE, JP, CH, WO e INPADOC. Es Propiedad del grupo Thomson.

**Thomson Innovation:** Sistema integrado de búsqueda de patentes, que incluye además de las colecciones que se indican en DELPHION el acceso a publicaciones científicas.

**Total Patent:** Sistema integrado de la empresa LEXIS-NEXIS. Las colecciones son algo diferentes a las de Thomson Innovation.

**SumoBrain:** Sistema integrado de consulta, incluye textos completos. Desarrollado por el mismo equipo que el servicio gratuito Free Patents on line.

Cada uno de estos sistemas posee unas características técnicas diferentes, por lo que a la hora de evaluar cuál es el mejor habría que ponderar toda una serie de criterios: cobertura, campos de búsqueda, posibilidad de acceder a los textos completos, exportación de resultados, RSS, etc.

# Portales, software, buscadores y otras herramientas

## Herramientas de búsqueda avanzada

La herramienta IP Century <<http://www.sales4ipcentury.com>> ofrece resultados automáticos del estado de la técnica, la nulidad, la libertad para realizar búsquedas. Es posible que en un futuro próximo sea posible obtener mejores resultados con nuevas tecnologías en materia de búsqueda y con ordenadores más potentes.

Es en el ámbito universitario donde pueden encontrarse la mayoría de esos nuevos motores. Los datos Medline se utilizan en ambos casos en los ejemplos siguientes.

GoPubMed <<http://www.gopubmed.org>>, desarrollado por la Universidad de Dresden, es una herramienta de búsqueda de documentación basada en la ontología que permite encontrar los artículos en materia de investigación biomédica. La búsqueda se presenta a PubMed y los resúmenes resultantes se clasifican utilizando la Gene Ontology (ontología de genes) y la Medical Subject Headings (MeSH). Las herramientas gráficas suplementarias mostrarán todas las informaciones estadísticas posibles.

MEDIE <<http://www-tsujii.is.s.u-tokyo.ac.jp/medie>>, creado por una universidad japonesa, es un motor de investigación inteligente desarrollado para encontrar correlaciones biomédicas a partir de MEDLINE. Se pueden localizar resúmenes/ frases en MEDLINE al especificar los campos semánticos de las correlaciones, por ejemplo "What activates p53" (qué activa el p53) o "What causes colon cancer" (qué causa el cáncer de colon). Esta es una de las herramientas de búsqueda científica semántica más impresionante.

Estas tres herramientas representan únicamente una pequeña parte de los esfuerzos llevados a cabo para desarrollar más fácilmente la utilización de las herramientas de búsqueda. Aplicaciones de búsqueda más poderosas pueden aparecer antes de lo esperado.

Fuente: Una nueva generación de herramientas de búsqueda de patentes y de informaciones científicas

<http://www.ipr-helpdesk.org/newsletter/35/html/ES/howToFindAPatent.html>

## Software de patentes

### Decopa de IPCentury

El programa DECOPA de la empresa IPCentury. Este programa realiza un estudio en profundidad de los documentos de patentes utilizando técnicas de minería de la información. Se trata de un programa de pago pero dispone de un acceso gratuito (opción demo) con el que es posible, dado un número de patente concreto o un texto que puede ser el de una reivindicación de una patente, realizar cuatro tipos de análisis que denominan: novedad, invalidez, patentabilidad e infracción.

El sistema hace el análisis comparando el texto o la patente que incluye el usuario con la colección de documentos disponible en su base de datos, y ofrece una lista de resultados ordenados por relevancia. Por tanto, este sistema permite localizar documentos que estén relacionados de alguna manera con la patente que se estudia, lo que permite automatizar parcialmente el proceso de búsqueda de patentes en un sector.

### PROInTec: un software para el tratamiento inteligente de datos sobre patentes

El software, al estar diseñado y desarrollado en el marco de un proyecto de investigación presentado, desarrollado y liderado por el grupo PROGINTEC, no se comercializa aún. En estos momentos sólo brinda servicios gratuitos a profesores investigadores de la provincia Pinar del Río, vinculados con las prioridades de la ciencia y a miembros del Polo Científico. Se accede por medio de la *Web tecnológica de la UPR*, desarrollado por el mismo grupo de investigación y que promueve información de corte científico tecnológico. PROInTec permite la búsqueda, recupera-

ción, procesamiento, análisis y presentación de información contenida en documentos de patentes procedente de bases de datos internacionales.

### Matheo Patent

Matheo Patent <http://www.matheo-patent.com> es un software de uso personal (monopuesto) diseñado para automatizar las tareas de búsqueda, recuperación y análisis de patentes de las bases de datos de la Oficina de Patentes de los EE.UU. (USPTO) y de la Oficina Europea de Patentes (EPO). Ambos sitios dan información gratuita y de calidad, pero requieren mucho tiempo y paciencia si se desean hacer las búsquedas “a mano”.

Matheo Patent ha desarrollado las siguientes funciones:

- Recuperación automática de los resultados de una búsqueda sin límite en el número de patentes descargadas.

- Descarga de todas las informaciones relativas a una patente (ficha, resumen, reivindicaciones, estado legal, gráficos, documento pdf) y ordenación mediante pestañas.

- Construcción de una base de datos local con todos los resultados obtenidos en varias estrategias de búsqueda.

- Almacenamiento de las estrategias de búsqueda y actualización selectiva de cada una de ellas para detectar novedades.

- Sistema para puntuar las patentes de acuerdo con las preferencias de la persona usuaria según cuatro ejes (inversiones, formación, riesgo de infracciones e impacto en el negocio).

- Análisis estadísticos y generación de representaciones gráficas (histogramas, matrices y redes).

- Agrupación de los miembros de la misma familia de patentes en un solo registro.

- Generación de informes personalizados, etc.

- Exportación de resultados en formato texto y XML. Matheo Patent se utiliza en los procesos de búsqueda, evaluación y comparación de tecnologías, detección de innovaciones, vigilancia tecnológica y de la competencia, evaluación de negocios, inteligencia competitiva, etc.

### Mimosa

- Software que agrupa las BD de patentes de USA, Japon y Europa.

## Comparativa entre las tres herramientas del Matheo: Analyzer, Patent y Web

MATHEO ANALIZER	MATHEO PATENT	MATHEO WEB
<ul style="list-style-type: none"> <li>Minería de Datos textuales.</li> </ul>	<ul style="list-style-type: none"> <li>Minería de datos textuales.</li> </ul>	<ul style="list-style-type: none"> <li>Minería de la Web.</li> </ul>
<ul style="list-style-type: none"> <li>Base datos estructurada (patentes, artículos...).</li> </ul>	<ul style="list-style-type: none"> <li>Base de datos de patentes de USPTO y EPO (limitación).</li> </ul>	<ul style="list-style-type: none"> <li>Diversas fuentes en Internet.</li> </ul>
<ul style="list-style-type: none"> <li>Necesita etapa de pretratamiento (obtención de corpus e importación).</li> </ul>	<ul style="list-style-type: none"> <li>No necesita de pretratamiento de la información.</li> </ul>	<ul style="list-style-type: none"> <li>No necesita de pretratamiento de la información.</li> </ul>
<ul style="list-style-type: none"> <li>Tratamiento semiautomático de la información, lo que le concede mayor flexibilidad.</li> </ul>	<ul style="list-style-type: none"> <li>Tratamiento automático de las patentes recuperadas.</li> </ul>	<ul style="list-style-type: none"> <li>Uso de métricas estadísticas.</li> </ul>
<ul style="list-style-type: none"> <li>Uso de métricas estadísticas.</li> </ul>	<ul style="list-style-type: none"> <li>Uso de métricas estadísticas.</li> </ul>	<ul style="list-style-type: none"> <li>Visualización: histogramas, grafos y matrices.</li> </ul>
<ul style="list-style-type: none"> <li>Visualización: histogramas, grafos y matrices.</li> </ul>	<ul style="list-style-type: none"> <li>Visualización: histogramas, grafos y matrices.</li> </ul>	<ul style="list-style-type: none"> <li>Publicación automática de informes y reportes.</li> </ul>
<ul style="list-style-type: none"> <li>El analista debe organizar y editar el informe.</li> </ul>	<ul style="list-style-type: none"> <li>Publicación automática de informes y reportes.</li> </ul>	

Fuente: [http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35\\_INTEC.pdf](http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35_INTEC.pdf)

## Otras herramientas

Entre las herramientas existentes para el análisis de información de los principales convertidores de formatos disponibles en el mercado, son: *Bibliolink* y *CONVI*, entre los más utilizados; los gestores de bases de datos más populares hoy, el *Procite* y *Endnote*. Y entre los paquetes de procesamiento estadístico el: *Excel*, *Statistica*, *SPSS*, *XLSTAT*, etcétera. Y específicos de información de patentes, el *PATS-TAT+* (desarrollado por *Derwent Publications*) y *BATELLE* (*Battelle Development Corporation*). Estos últimos presentan limitaciones para su uso debido a que sólo procesan datos de la bases de datos *Derwent* y de la *United States Patent and Trademark Office* (USPTO) <<http://www.uspto.gov>>.

Sin dejar de mencionar a *INFOCAM*, diseñado por el Instituto de Documentación Científica y Tecnológica-IDICT de Cuba; *CUIB-METRIC*, de la Universidad Nacional Autónoma de México-UNAM; *TOA-KNOWBOT*, del *Tech Technology Policy and Assessment Center*, Georgia; *NEURODOC*, del *L'Institut Scientifique et Technique*, Francia y el poderoso *BIBEXCEL*, mayoritariamente asociado al *Science Citation Index-SCI*, desarrollado por el profesor *Olle Persson* de la Universidad de Umea, Suecia.



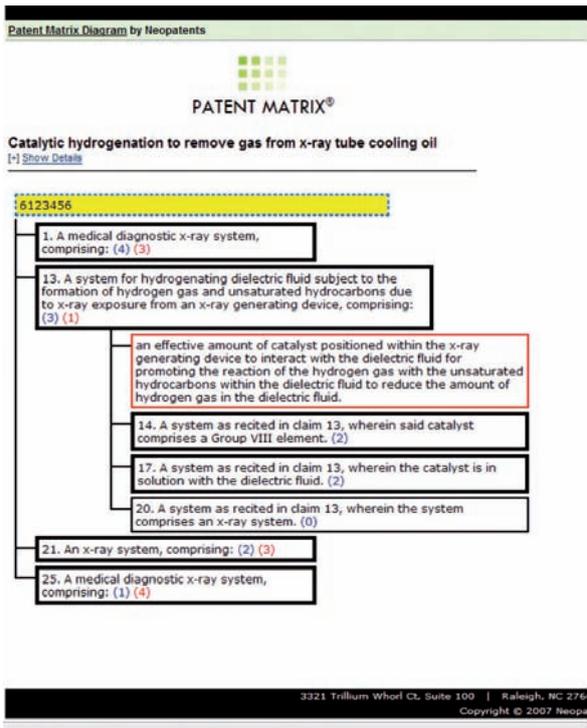
Entre los sistemas que permiten las representaciones de datos y son promovidos por el *Patent Information Users Group* (PIUG) <<http://www.piug.org>>, se encuentra *BizInt Smart Charts*, una herramienta para la búsqueda, análisis y presentación de los resultados de los análisis. Permite la construcción de un árbol de referencias para determinar el ciclo y trayectorias de una tecnología, así como las posibles estrategias de patentamiento en un sector técnico. Y *Wisdomain*, desarrollado por *Delphion* para complementar su servicio de búsqueda de patentes, que ofrece informes y gráficos sobre los campos de información bibliográfica de la patente.

Entre las herramientas más utilizadas en los últimos años para realizar estudios métricos y que permiten usar información de patentes, se encuentran productos europeos como: *Leximape*, el *Dataview*, el *Tretalogie*, *BibTechMon*, entre otros; y norteamericanos como *PatentLab-II Eureka* de *Aurigin*, el *VantagePoing* (TOAK), *Matheo-Patent*® y el *InfoViz* [[http://www.omniviz.com/files/Omniviz\\_Webinar\\_nov2008.pdf](http://www.omniviz.com/files/Omniviz_Webinar_nov2008.pdf)] desarrollado por *Battelle*, entre otras aplicaciones no menos importantes.

Esta empresa se ha fusionado con la empresa *BioWisdom Ltd.*, e inicialmente se centró en el desarrollo de métodos de visualización de los resultados del análisis. Actualmente, ofrece un paquete avanzado de software que provee visualización de datos numéricos, información categorizada, secuencias genómica, estructuras químicas, y documentos textuales, como publicaciones y patentes, en un mismo espacio visual. El software combina avanzados algoritmos de análisis textual y estadísticos, con visualizaciones de alta calidad, que facilitan el descubrimiento de nuevo conocimiento. Como fortaleza tiene que permite el tratamiento y visualización (*Galaxy*, *CoMet*) de todo tipo de información en forma conjunta. Las aplicaciones iniciales han estado en el campo de la Química, las cuales luego se extendieron a áreas de gobierno, educación y mercados.

de productos comerciales. Las áreas de aplicaciones incluyen la I+D, ECM, Legal, Finanzas, Mercado, Ensayos Clínicos<sup>19</sup>.

<sup>19</sup> intec Factor clave para la toma de decisiones estratégicas en las organizaciones; madri+d [[http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35\\_INTEC.pdf](http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35_INTEC.pdf)].



La utilización de mapas conceptuales constituye una estrategia de aprendizaje para producir aprendizajes significativos, ya que su objetivo básico es relacionar conceptos de forma ordenada. Recordamos que, en general, un aprendizaje es significativo cuando podemos construir conocimiento en base a conocimientos anteriores.

Pero además de este objetivo básico, existen otros objetivos del mapa conceptual, que son:

- Estructurar un conjunto de conceptos y/o ideas.
- Establecer relaciones entre ellos generando nuevos conceptos más complejos.

Los motivos por lo cuales es interesante utilizar los Mapas Mentales son:

- Facilidad: realizar los gráficos con herramientas específicas es muy sencillo.
- Eficiencia: al organizar las ideas se puede mejorar su posterior utilización.
- Utilización: esta técnica es muy utilizada en diferentes ámbitos profesionales, como son las empresas, universidades, etc.

## Software de visualización de la información

La técnica del Mapa Conceptual o Mental fue creada por el psicólogo británico Tony Buzan, investigador en el campo de la inteligencia. Su importancia radica en que expresan una forma de pensamiento, que es el pensamiento irradiante. Esto significa que las personas no tenemos una forma de pensamiento completamente 'lineal', tendemos a expresar nuestros pensamientos de forma más bien estrellada y luego establecer relaciones entre ellos. Si queremos que no 'se pierdan' hay que plasmarlos de alguna forma y luego organizarlos y relacionarlos. A partir de aquí estaremos en disposición de sacar conclusiones y/o resultados.

Un mapa conceptual es una herramienta para la organización y representación del conocimiento ([www.es.wikipedia.org](http://www.es.wikipedia.org)). Los mapas conceptuales constituyen un mecanismo muy útil para transmitir de forma clara y concisa mensajes o ideas complejas.

## TouchGraph

Es un programa escrito en java que permite visualizar y navegar interactivamente por redes de información interrelacionada. Existen dos modalidades, una versión liberada en sourceforge y una versión de pago propiedad de Google.

## FreeMind: mapas conceptuales

FreeMind<sup>20</sup> es una herramienta para la elaboración y manipulación de mapas conceptuales. Es decir, una herramienta para organizar y estructurar las ideas, los conceptos, su relación entre ellos y su evolución. Y, como hemos comentado, puede ser utilizada en cualquier área del ámbito educativo y como mecanismo o forma de plasmar tormentas de ideas de todo tipo para su posterior reutilización.

Está basado en Java y es Open Source, bajo licencia GPL. La página del proyecto es <http://freemind.sourceforge.net>. Es un software ligero, sencillo de instalar, configurar y utilizar. Al ser Java es multi-

<sup>20</sup> Mas información: <http://observatorio.cnice.mec.es/modules.php?op=modload&name=News&file=article&sid=716&mode=thread&order=0&thold=0&POSTNUKESID=2d1e46becd8a7678fcb32bbadf58dd3c> [ultimo acceso 12/09/09].

plataforma y, por lo tanto, disponible en Windows y GNU/Linux.

Su requerimiento más importante es tener instalado el entorno de ejecución Java JRE 1.4 o superior. Los proyectos generados los guarda como XML, garantizando así su reutilización.

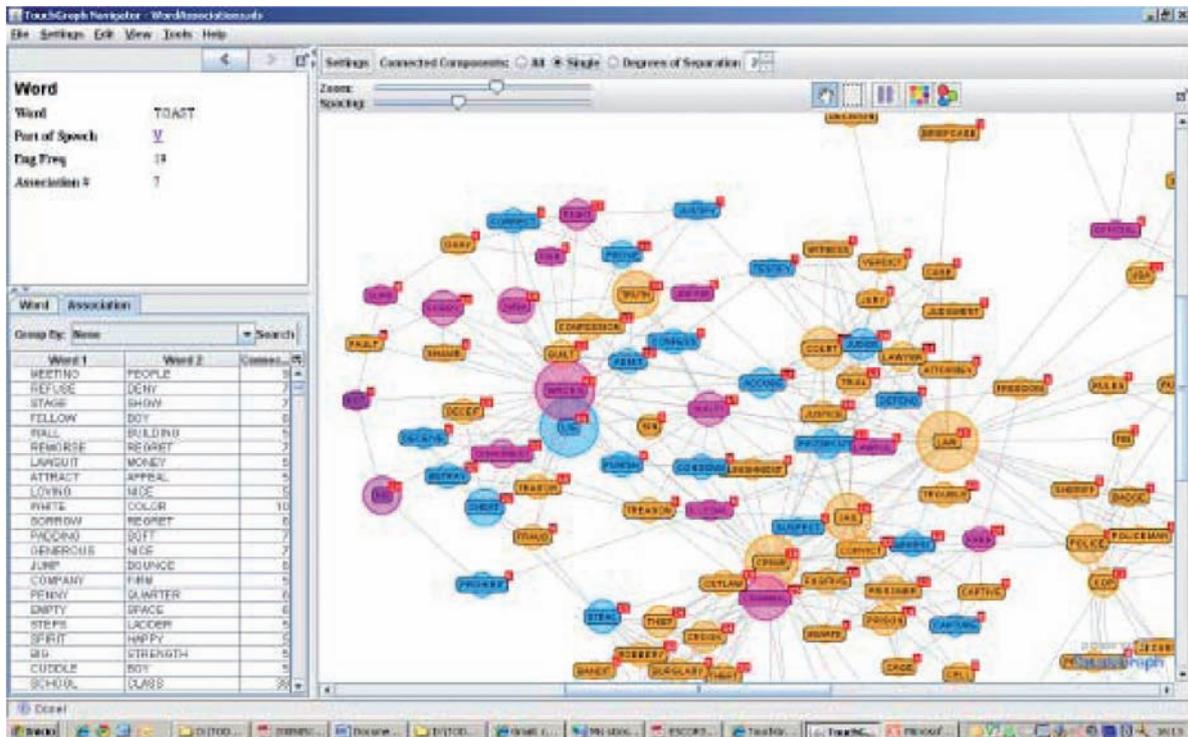
El resultado es un grafo en el que se representan las ideas relacionadas con un tema y se pueden ir refinando y extendiendo cada una de las ramas hasta capturar todas las ideas relativas al tema que se pretende analizar.

Las principales características de FreeMind son las siguientes:

- Permite la utilización de enlaces HTML en los nodos.
- Soporta el plegado y desplegado de los nodos.
- Arrastre/Copia inteligente, es decir, el usuario puede mover los nodos y su información y estilo asociado. También para múltiples nodos seleccionados, arrastrar textos, enlaces y archivos asociados.

- Permite publicar los mapas en Internet como páginas HTML o XHTML si queremos incluir la imagen.
- Tiene la opción de incluir iconos en el mapa conceptual para diferenciar ideas/conceptos específicas o relacionadas. También puede incluir iconos simplemente como decoración.
- Utiliza formato XML lo que facilita exportar los mapas conceptuales a otras aplicaciones.
- Es multilingaje, con lo cual soporta español. Otras aplicaciones con funcionalidad similar a FreeMind son:

- Xmind, basado en eclipse framework y también Open Source. <<http://www.xmind.net>>
- MindManager, propiedad de la empresa MindJet. <<http://www.mindjet.com>>
- Compendium (Open Learn). <<http://projects.kmi.open.ac.uk/compendium>>
- CmapTools, creado por el Institute for Human and Machine Cognition (IHMC), de la Universidad de West Florida (Estados Unidos), <<http://cmap.ihmc.us>>

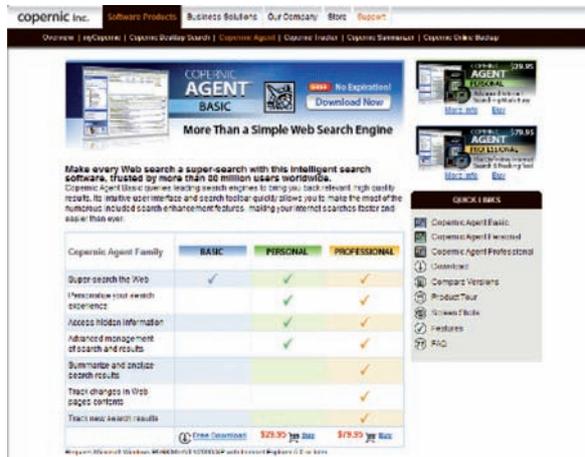


Para saber un poco más de TouchGraph:  
<http://www.barriblog.com/index.php/2008/09/22/como-usar-touchgraph-navigator-y-no-morir-en-el-intento/>

## Agentes de búsqueda

### Copernic

<http://www.copernic.com/>



### Botbox

<http://www.botbox.com/>



### WebSite-Watcher

<http://www.aignes.com/>

Es un software especializado en sistematizar la monitorización de páginas al día y cambios de especial interés en páginas web.

WSW tiene un menú sencillo para personas con poca experiencia y un menú avanzado para quienes desean programar scripts o definir la detección de cambios muy especiales.

WSW puede monitorizar todo tipo de informaciones o ficheros accesibles en la red local, en la intranet o en internet.

Permite sistematizar el chequeo de cambios en cualquier tipo de páginas web, incluyendo aquellas que están protegidas por un identificador y password, una función javascript, un formulario de

búsqueda, etc. También permite monitorizar páginas con número de sesión, etc. WSW dispone de un asistente para la creación de macros en casos complicados, que permiten que una persona con un breve entrenamiento programe a WSW para que llegue a la página que hay que monitorizar.

Posibles usos de WebSite Watcher (fuente Grupo CDE <<http://www.cde.es>>):

- Monitorizar foros para detectar nuevos temas de discusión y sus respuestas.
- Monitorizar canales de noticias RSS en los que aparece al menos un término de un listado predefinido.
- Monitorizar nuevos resultados de una búsqueda en una base de datos.
- Detectar cambios en una especificación de un producto, servicio, contrato, etc.
- Detectar una versión nueva de un fichero binario (p.ej. ficheros .zip o .exe).
- Vigilar cuando hay una nueva versión o un nuevo precio de un producto.
- Vigilar la evolución de la competencia, detectando los cambios que se van produciendo en su sitio web.
- Vigilar cambios en un sitio web, detectando nuevos contenidos en los que aparece un término de interés.

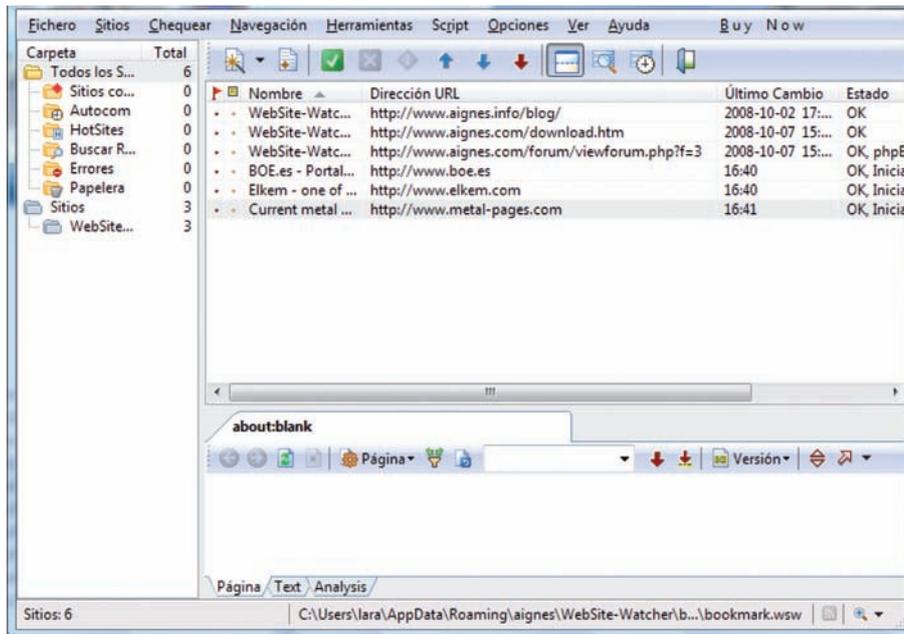
Website-Watcher permite establecer filtros para considerar un cambio como "significativo":

- que esté situado en un área concreta de la página.
- que figure al menos una palabra clave, de un listado de términos predefinidos.

También permite ignorar cambios en contenidos no significativos, p.ej. día y hora, número de personas usuarias conectadas, nº de documentos en una base de datos, tiempo de respuesta a una pregunta, mensajes de error de un servidor, etc.

Una vez detectado un cambio, WSW permite activar diferentes sonidos, alertas o ejecutar diferentes programas en función de la página que ha cambiado. Por ejemplo, es capaz de publicar una página en la intranet siempre que haya cambiado. También es posible archivar ciertas páginas en combinación con el programa WebSite Archive.

Por último, es posible definir diferentes formatos de informes y generarlos a voluntad.



## Naveganza

<http://www.isoco.com/pdf/naveganza.pdf>

*Cuando buscamos información con un buscador tradicional, encontramos multitud de respuestas a nuestras preguntas, algunas pueden contener la respuesta que buscamos y otras no guardar ninguna relación con nuestra necesidad de información. Naveganza es la solución de iSOCO que permite obtener respuestas precisas a las preguntas que se le realizan en lenguaje natural basándose en los conceptos y no solo en las palabras.*

## Software de Vigilancia Tecnológica

Es muy importante hacer una diferenciación entre las plataformas integradoras que hay que configurar, como por ejemplo la plataforma Denodo, de otros tipo de software que son paquetes cerrados y sólo hay que configurar el tipo de información o como se quiere visualizar. No hay una solución definitiva que nos evite realizar las tareas de Vigilancia Tecnológica. En el mercado hay soluciones, muchas plataformas, algún que otro software, pero ninguno es la solución definitiva. Lo que quiero decir es que por ejemplo, la Plataforma

VIXIA, la plataforma de Vigilancia del ITA (Instituto Tecnológico de Aragón), utilizan por debajo el software de DENODO technologies.

## Thera

Thera Centre de Llenguatge i Computació S.L. está constituida por un grupo de trabajo dinámico y versátil, que reúne jóvenes con diferentes perfiles profesionales con un alto nivel de especialización: lingüistas, filólogos, traductores e ingenieros informáticos.

Thera, spin-off de CLiC, Centre de Llenguatge i Computació de la Universitat de Barcelona, cuenta con la experiencia de mas de 20 años de investigación en el área de la Lingüística Computacional, y con mas de 6 años en el desarrollo de aplicaciones en las tecnologías del conocimiento y en la gestión de la información basadas en la ingeniería lingüística <<http://www.thera-clic.com/site/Empresa/index-ES.html>>.

## Vigiale

Plataforma Web para la gestión de la Vigilancia Tecnológica y del Entorno, diseñada bajo los principios de la segunda generación de aplicaciones Web o “Web 2.0?”, que permite gestionar recursos

de información de distinta naturaleza, ordenarlos, clasificarlos, actualizarlos y recibir notificaciones de cambio. Vigiale es además, una red social en torno a un sector específico, permitiendo la interacción y colaboración entre los usuarios en foros de discusión sobre un determinado recurso, y en proyectos para el desarrollo de estudios de Vigilancia Tecnológica e Inteligencia Estratégica.

## LUXID

(www.temis.com/index.php?id=217&sel=1)

Aplicación líder en el tratamiento de informaciones en texto libre –artículos científicos y patentes principalmente– utilizando Minería de Textos (Text Mining). Analiza diversas clases de información en una sola aplicación. La mayoría de las aplicaciones se concentran en un área de conocimiento,

como la interacción proteína-proteína, pero Temis ha creado una poderosa solución integradora que permite analizar, por ejemplo, información química conjuntamente con información biológica.

## Diana Technologies

XerKa es una solución software que vigila automáticamente la aparición de nueva información de interés para la empresa. Automatiza los procesos de búsqueda, análisis, clasificación y difusión de dicha información para que cada usuario disponga de información de valor en el momento oportuno y la empresa rentabilice el proceso de vigilancia.

## Vixía

El CIS-Ferrol desarrolla el sistema «Vixía» en el año 2003. «Vixía» es un Sistema de Vigilancia Tec-



Tabla 4. Análisis comparativo de algunas herramientas de Software utilizadas en VT Documentos: WEB (

Producto	Costo promedio versión profesional	Documentación que procesa	Apoyo Ciclo VT				Estadística	Actores de uso			
			Planificación	Búsqueda	Análisis	Inty Com.		Pymes	CDT's	Universidad	Org. Ntl
<b>Herramientas enfocadas en la búsqueda de información a través de Internet</b>											
C-4-U Scout 1.3	Gratis	WEB	Poco	Poco	No	No	No	X	X		
CI Spider 1.2.1	No Def.	WEB	Poco	Poco	Poco	Poco	Básica		X	X	
Copernic 6.0	USD\$79.95	WEB	Poco	Si	Poco	No	Básica	X	X	X	X
Seekip	Gratis	WEB	Poco	Si	No	No	No	X	X	X	X
Strategic Finder 2.0	€\$475	WEB	Poco	Si	Poco	No	Básica		X	X	
WebFerrer 5.0	USD\$29.95	WEB	Poco	Poco	Poco	No	Básica	X			
WebSeeker 5.0	USD\$29.95	WEB	Poco	Si	No	No	Básica	X			
<b>Herramientas enfocadas al procesamiento y análisis de información</b>											
MindModel 2.195	USD\$220	WEB-Doc TXT	No	Poco	Poco	No	Básica		X	X	
Sonar Professional	USD\$795	Doc TXT	Poco	Si	Poco	No	Básica			X	
XIStat 7.5.3	USD\$395	Estructurada	Poco	No	Si	Poco	Avanzada			X	
Hamlet 2.0 Beta	No Def.	Doc TXT	Poco	No	Si	Poco	Avanzada	X	X		
TextPack	€500	Doc. TXT	Poco	No	Si	Poco	Básica			X	
Matheo 3.0	€3450	Doc TXT – Est.	Si	No	Si	Poco	Básica			X	X
T-Lab 4.1	€1150	Doc TXT	Poco	No	Si	Poco	Básica	X	X	X	
SPSS 14	USD\$1499	Estructurada	Poco	No	Si	Poco	Avanzada			X	X
Tetralogie 6.0	USD\$12000	Estructurada	Si	No	Si	Poco	Avanzada			X	X
<b>Herramientas enfocadas al procesamiento y análisis de información</b>											
Matheo Pathent 3.0	€600	Patentes	Si	Si	Si	Poco	Básica	X	X	X	X
Aurekal 9.2	No Def.	Patentes	Si	Si	Si	Si	Básica			X	X
<b>Herramientas enfocadas al procesamiento y análisis de información</b>											
GoldFire 2.5	USD\$16000	Todo	Si	Si	Si	Si	Básico			X	X

Fuente: Valoración, selección y pertinencia de herramientas de software utilizadas en Vigilancia Tecnológica. En Ingeniería e Investigación, abril, 2006, vol. 26, numero 001, Colombia.

nológica (VT) e Inteligencia Competitiva (IC) cuyo aspecto más innovador es la asociación de una metodología y una plataforma tecnológica desarrollada exprofeso según los requisitos operativos definidos por la Unidad de Vixilancia e Transferencia Tecnológica do CIS Galicia. «Vixía» permite a las empresas disponer de información permanentemente actualizada sobre los cambios que se producen en su entorno y que les afectan en su actividad diaria. El Sistema «Vixía» facilita además los procesos de gestión de la información que genera la propia empresa, y dinamiza la comunicación interna de forma multidireccional: entre los departamentos (horizontal) y entre los distintos niveles de la jerarquía organizativa de la empresa (vertical).

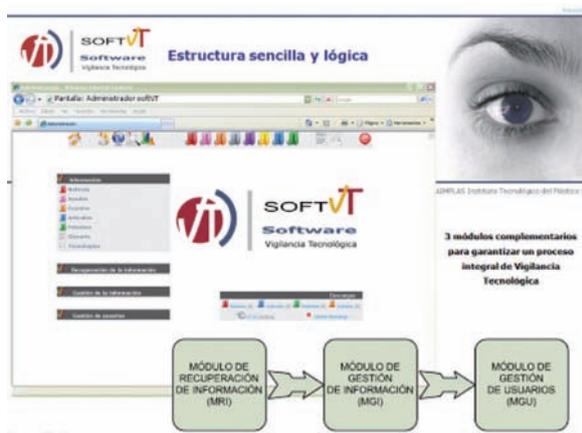
### AIMPLAS SOFTVT

<http://www.softvt.com>

*softVT* es un software de vigilancia tecnológica desarrollado en 2003 por el Departamento de Información Técnica de AIMPLAS Instituto Tecnológico del Plástico, formado por un equipo de profesionales de la Documentación e Informática con varios años de experiencia en el campo de la vigilancia tecnológica.

La aplicación *SoftVT* permite no sólo el control de un número indefinido de *fuentes de información*, sino también la generación de varios productos y servicios de información a través de la información recuperada, todos ellos orientados a la *detección, gestión y diseminación de la información tecnológica* más reciente y actualizada, que en muchos casos puede ser crucial en la toma de decisiones estratégicas por parte de una empresa.

*SoftVT* se trata de una herramienta “escalable” y versátil, por lo que ya ha sido utilizada como solución en diferentes sectores (plástico, juguete, aeronáutica, automoción, envase-embalaje, etc.). Entre las posibilidades que ofrece este *software para la vigilancia tecnológica*, pueden enumerarse diferentes aplicaciones: desde la publicación web de un boletín de novedades tecnológicas extraídas de un *número ilimitado de fuentes de información*, hasta una implantación total que permita crear y gestionar una base de datos propia y el envío de boletines y *alertas tecnológicas personalizadas*, con el fin de establecerse como una solución integral en cuanto a *inteligencia competitiva* se refiere.



### Vicubo

<http://www.vicubo.es/>

e-Intelligence tiene su Origen en la **Universidad Politécnica de Madrid** como grupo de I+D.

Se creó como *spin off* en el año 2002 formada por un **equipo multidisciplinar** de (informáticos, ingenieros en telecomunicaciones, documentalistas...) con amplia experiencia en el ámbito de las Tecnologías de la Información y la gestión de la Innovación Tecnológica.

Sus principales áreas de negocio son:

- Vigilancia Estratégica.
- Consultoría y Desarrollo Web.
- Administración de Sistemas.

### Denodo

Denodo <http://www.denodo.com> es una empresa especializada en la extracción e integración de información en tiempo real sobre sistemas distribuidos.

Denodo trabaja con información procedente tanto de entornos corporativos (bases de datos, hojas de cálculo, aplicaciones, sistemas de ficheros y cualquier otro repositorio de información no estructurada) como de Internet (sitios web de entidades socias, empresas proveedoras, clientes y competidoras e información web accesible sobre datos y noticias de mercado, regulaciones, ayudas, oportunidades, etc). El resultado es la alimentación de los procesos de negocio con datos integrados en tiempo real.

## Digimind

DIGIMIND <http://www.digimind.fr> es una empresa francesa que ha diseñado y desarrollado una plataforma de gestión del ciclo de la Inteligencia Estratégica denominada DIGIMIND Evolution. Esta plataforma permite a las empresas desplegar equipos de vigilancia estratégica y dinamizarlos.

DIGIMIND Evolution responde a las necesidades crecientes de las empresas de desarrollar sus capacidades anticipativas y reactivas ante la rápida evolución de su entorno y sobre todo ante las acciones de la competencia.

Este software está concebido para facilitar que las empresas diseñen equipos de vigilancia distribuidos pero a la vez coordinados, ágiles y eficaces para detectar amenazas y oportunidades. Digimind Evolution soporta todas las actividades de estos equipos, de modo que la recopilación, la difusión y el análisis de las informaciones estratégicas se convierten en tareas fáciles (Fuente: Grupo CDE).

## Portales de organismos públicos que ofrece servicios de Vigilancia Tecnológica

### Obtesol

[http://www.obtesol.es/index.php?option=com\\_content&task=view&id=5&Itemid=35](http://www.obtesol.es/index.php?option=com_content&task=view&id=5&Itemid=35)

### Aimme

<http://observatorio.aimme.es/>

### Instituto Tecnológico de Aragón

<http://www.ita.es>

### Madri+d

<http://www.madrimasd.org/informacionidi/biblioteca/Publicacion/Vigilancia-tecnologica/default.asp>

### Zaintek

<http://apli.bizkaia.net/apps/danok/w2/ebizkaia/L1/index.asp>

### CIS-Centro de Innovación y Servicios

<http://www.vixia.info>

## Webs, blogs y otros portales interesantes

La Cofa: <http://www.lacofa.es/>

Este es un weblog colectivo dedicado a la Vigilancia Tecnológica sobre el Futuro de Internet, la Banda Ancha y la Movilidad y otros temas de las Tecnologías de la Información y las Comunicaciones. Y siempre primando un enfoque más técnico que de negocio o cultural. Todo esto con criterio y ánimo flexible. El contenido, estructura y orientación no es fijo ni inamovible, al igual que estos principios, por lo que estaremos muy atentos al desarrollo y a la respuesta que tenga esta actividad para modificarla buscando la máxima utilidad de forma que sea una herramienta útil para sus lectores.



### collective intellect

<http://www.collectiveintellect.com>

## La Inteligencia Competitiva: factor clave para la toma de decisiones estratégicas en las organizaciones

<http://www.madrimasd.org/Inteligencia-Competitiva/introduccion/default.asp>



# Análisis de la información recuperada

En esta fase se realiza el procesamiento y análisis de la información que hemos buscado. Es el paso más importante, porque es transformar datos en conocimiento.

Para fortalecer estos procesos se utilizan herramientas informáticas (software) especializadas y se hacen consultas que permiten obtener resultados convencionales.

Herramientas informáticas especializadas.

Dado el volumen de información que en general contienen las fuentes, para su manipulación se han desarrollado diversos software especializados, que se encuentran disponibles en el mercado (Excel, access, SPSS...).

Consultas convencionales a las fuentes seleccionadas. Estas consultas difieren de acuerdo con la naturaleza de los temas, es decir, según sean de índole científica o tecnológica y, por ende, de las fuentes de información: bases de datos de artículos científicos o de patentes. En el caso de las bases de datos documentales, que proporcionan información de carácter científico, algunas consultas genéricas que se realizan en los distintos estudios son:

Número de artículos por año registrados en la base de datos: este indicador de frecuencia permite concluir si se está en presencia de un área de interés científico con amplia trayectoria, madura, o de reciente aparición.

Fuentes potenciales de conocimiento: estas se identifican a partir del recuento simple de los países, revistas (fuentes de publicación), autores, que investigan en los temas (establecidos a través de palabras clave) y su nivel de importancia, medido por la frecuencia.

Temas de interés científico para el país líder o algún país competidor en los últimos años: permite establecer el perfil científico del líder. Su cálculo se basa en la coocurrencia o aparición conjunta de los temas y el país.

Temas o áreas de investigación de los autores líderes: a partir de la coocurrencia de estos dos campos de información Cambios en los temas a través del tiempo: permite observar las trayectorias de desarrollo del conocimiento.

Redes de colaboración entre autores y países: para lo cual se analiza la coocurrencia de los campos autor y país.

En el caso de las bases de datos de patentes, para la obtención de información sobre los avances tecnológicos algunas de las preguntas (convencionales) de los equipos de investigación son:

Número de patentes concedidas a través del tiempo: con este indicador se identifican los temas que para las (diferentes) empresas han sido de creciente interés y en los cuales han desarrollado sus capacidades tecno-lógicas.

Número de patentes concedidas y solicitadas por año de publicación: la diferencia entre las patentes concedidas y las solicitadas permite establecer el número de años que transcurren entre la solicitud y la concesión para un área de conocimiento.

Identificación de líderes tecnológicos: para ello se realizan recuentos simples de algunos campos de información, como las patentes concedidas por instituciones titulares o las otorgadas por país.

Número de patentes por área tecnológica según la clasificación internacional de patentes: A través de este indicador se identifican las principales áreas tecnológicas y los temas de mayor interés conducentes a la patentabilidad.

Número de patentes por año: con este indicador se identifica la dinámica de las tecnologías, es decir, aquellas áreas que se han fortalecido a través del tiempo y aquellas más emergentes, definidas a partir de los temas que han aparecido en los últimos años.

Selección de las áreas tecnológicas emergentes: para ello se descomponen las áreas tecnológicas

según la clasificación internacional de patentes, analizando las patentes concedidas y solicitadas.

Coocurrencia de las áreas tecnológicas de mayor frecuencia para un año específico: con este indicador se identifican las relaciones entre las tecnologías Perfil tecnológico de las empresas líderes o pioneras en áreas emergentes: para establecerlo se analizan las relaciones más fuertes entre las empresas y las áreas tecnológicas identificadas.

Áreas tecnológicas de mayor competencia, de nichos especializados y/o de poca concurrencia. La posibilidad de elaborar mapas tecnológicos depende del software empleado en los ejercicios de vigilancia tecnológica. Dichos mapas se construyen a partir de información estructurada (es decir, bases de datos de publicaciones científicas o de patentes) y permiten obtener una representación visual fácil de captar y de recordar a partir de una gran cantidad de información; también proporcionan un panorama detallado de las líneas de investigación y desarrollo tecnológico mediante el examen de lo que se está publicando o patentando.

## Informes de Vigilancia

La vigilancia tecnológica aprovecha un volumen extraordinario de información científico-técnica (propiedad intelectual, investigación, productos, estándares, etc...) y permite realizar una detección temprana de amenazas y oportunidades, aportando a las organizaciones ventajas competitivas.

Metodológicamente, los informes de vigilancia se pueden desarrollar en cuatro fases, que se detallan a continuación.

La primera fase involucra la definición de la temática y de los factores críticos de vigilancia; esta actividad se hace conjuntamente entre el equipo de VT y el cliente.

Una vez cumplida esta etapa se decide el equipo de trabajo, que son los expertos que realizan el informe de vigilancia, y el equipo de seguimiento formado por personal con experiencia en la temática, cuya labor es definir, seguir y evaluar el trabajo de vigilancia.

Tras la formación de los equipos, se procede a la reunión de lanzamiento del trabajo, cuyo objetivo es aclarar el enfoque idóneo y las líneas prioritarias del estudio. Con las ideas resultantes de la reunión, se inicia la segunda fase, donde el equipo de trabajo reúne la información solicitada y considerada de interés, concretando la primera versión del informe que se envía al equipo de seguimiento.

La tercera fase involucra al equipo de seguimiento que, tras analizar el informe, aporta su opinión y sugerencias sobre el avance del trabajo y, si es el caso, procede a la redefinición y concreción de algún aspecto referido a los objetivos y perfil de la vigilancia tecnológica establecida. En la cuarta y última fase, el equipo de trabajo elabora la versión final del informe, añadiendo y completando los comentarios aportados por el equipo de seguimiento y concluyendo de este modo el trabajo.

Esta metodología permite la existencia en todo momento de una fluida comunicación entre el personal que realiza el trabajo y la asociación que lo demanda, obteniéndose de ese modo un informe ajustado a las necesidades del cliente.

## Ejemplos en Internet

Servicios y aplicaciones en movilidad para el sector turístico, 2007.

[http://www.n-economia.com/informes\\_documentos/pdf/sintesis\\_documentos/SINTESIS\\_NE\\_01-2008.PDF](http://www.n-economia.com/informes_documentos/pdf/sintesis_documentos/SINTESIS_NE_01-2008.PDF)

Informes de vigilancia tecnológica de Madri+d.

<http://www.madrimasd.org/informacionidi/biblioteca/Publicacion/Vigilancia-tecnologica/default.asp>

Nuevo informe de Vigilancia Tecnológica sobre Polímeros Inteligentes.

<http://e-archivo.uc3m.es/dspace/handle/10016/2194>

Informes de Vigilancia Tecnológica.

<http://www.oepm.es>. La OEPM en colaboración con instituciones como el OPTI y la SETSI elabora trimestralmente y de forma gratuita los boletines de vigilancia tecnológica.

Sector Metal Mecánico 2009 [http://www.fiagro.org.sv/systemFiles/VT9\\_detecci%C3%B3n\\_gluten\\_alimentos%5B1%5D.pdf](http://www.fiagro.org.sv/systemFiles/VT9_detecci%C3%B3n_gluten_alimentos%5B1%5D.pdf)

Procesos empresariales bajo entornos colaborativos: Tecnologías y aplicaciones.

[http://innovacion.ita.es/obsetec/archivos/Informe\\_VT\\_Entornos%20Colaborativos.pdf](http://innovacion.ita.es/obsetec/archivos/Informe_VT_Entornos%20Colaborativos.pdf)

Informe de Vigilancia sobre Técnicas analíticas para detectar el gluten en los alimentos.

[http://www.fiagro.org.sv/systemFiles/VT9\\_detecci%C3%B3n\\_gluten\\_alimentos%5B1%5D.pdf](http://www.fiagro.org.sv/systemFiles/VT9_detecci%C3%B3n_gluten_alimentos%5B1%5D.pdf)

DNI electrónico. Informe de vigilancia tecnológica sobre eID realizado por CITIC

<http://www.consultoras.org/frontend/aec/DNI-Electronico—Informe-De-Vigilancia-Tecnologica-Sobre-EID-Realizado-Por-CITIC-vn6208-vst98#>

# Vigilancia tecnológica 2.0

## Definiciones

La Web 2.0 es la representación de la evolución de las aplicaciones tradicionales hacia aplicaciones web enfocadas al usuario final. El Web 2.0 es una actitud y no precisamente una tecnología.

Para entender de donde viene el término de Web 2.0 tenemos que remontarnos al momento en que Dale Dougherty de O'Reilly Media utilizó este término en una conferencia en la que compartió una lluvia de ideas junto a Craig Cline de MediaLive. En dicho evento se hablaba del renacimiento y evolución de la web.

## La Web 2.0 con ejemplos

Entender la evolución que ha llegado con la Web 2.0 puede realizarse con ejemplos, con proyectos. Podemos comparar servicios web que marcan claramente la evolución hacia el Web 2.0 con una nueva forma de hacer las cosas:

- Web 1.0 → Web 2.0.
- Doubleclick → Google AdSense (Servicios Publicidad).
- Ofoto → Flickr (Comunidades fotográficas).
- Akamai → BitTorrent (Distribución de contenidos).
- mp3.com → Napster (Descargas de música).
- Britannica Online → Wikipedia (Enciclopedias).
- Sitios personales → Blogs (Páginas personales).
- Especulación con dominios → Optimización en motores de búsqueda SEO.
- Páginas vistas → Costo por click.
- CMSs → Wikis (Administradores de contenidos).
- Categorías/Directorios → Tags.

## Vigilancia 2.0

Vigilancia 2.0 consiste en utilizar aplicaciones y sitios 2.0 para la búsqueda y filtrado de información pero también, en aprovechar la inteligencia colectiva y los efectos de red para vigilar el entorno siguiendo un modelo distribuido que permita el acceso a contenidos actualizados y relevantes para la organización.

La Vigilancia 2.0 implica entonces un cambio cultural. No se puede circunscribir a un departamento, sino que tiene que “socializarse” en la organización, y abordarse como una tarea donde la inteligencia colectiva puede mejorar los resultados.

En resumen, la detección de señales o “patrones” de interés estratégico para la empresa ya no es una tarea exclusiva del “gran capitán” (incapaz de mirar a tantos lugares a la vez), sino que depende del trabajo colaborativo de muchas *antenas conectadas en red*.

Una primera ventaja de la Vigilancia 2.0 es el sustancial abaratamiento de costes que puede significar el uso de estas herramientas.

Cualquier profesional puede crear su sistema de alertas o de sindicación RSS a medida y estar informado sobre los mercados, las nuevas tecnologías que afecten a su sector, la presencia de sus competidores en la Red y las publicaciones de interés, de un modo fácil y automatizado.

*Hoy en día* va siendo obsoleta la práctica de navegar por Internet a la caza y captura de un posible nuevo contenido *en alguna de las webs que ya tenemos identificadas como interesantes*.

*Para eso al abrir nuestro agregador dispondremos ya de todos esos contenidos sin haber invertido tiempo en su localización, así la tarea se convierte en altamente eficaz, permitiendo al experto (trabajador o directivo) centrarse en el aná-*

lisis y aplicación del conocimiento para el provecho de la empresa.

Otra de las ventajas, igual de importante, es la mejora de la productividad gracias a los filtros, sistemas de alerta y sobre todo aprovechando la inteligencia colectiva existente en la red.

La información vital para muchas organizaciones ya no se encuentra sólo en las publicaciones científicas o sectoriales, sino que esta fluyendo libremente, y con mucha mas velocidad, por las redes sociales, en los blogs y en el microblogging.



<http://www.cea.es/herramientas/page/Vigilancia.aspx>

Complementos/Páginas para facilitar y agilizar la búsqueda

- Pdf Search Engine: <http://www.pdf-search-engine.com/>



Funcionalidades que se pueden integrar en el servicio

- Return My Pants: <http://www.returnmypants.com>

Web que ofrece un útil recurso para todos los que andamos siempre prestando cosas, y la mayor parte de las veces, perdiéndolas por no recordar a quién se la hemos dejado. La interfaz es sencilla y clara y cumple perfectamente su objetivo: gestionar lo que prestas y te prestan (imaginemos que el servicio de VT de una empresa tiene una biblioteca de apoyo).

- Masterbranch: <http://www.masterbranch.com>

Agrega toda nuestra actividad en proyectos de código abierto y en sitios de preguntas y respuestas y la muestra en una página. Es una mezcla de red social y gestor de curriculums, puesto que podemos seguir a otros usuarios, actualizar nuestro estado,...

- Google fusion tables: <http://tables.googlelabs.com>

Google ofrece un nuevo servicio en fase beta llamado Google Fusion Tables, se trata de una herramienta web, enfocada a trabajar de forma colaborativa en la nube con un conjunto de datos relativamente grande puesto que dejará subir archivos tanto .csv como en formato Excel hasta un máximo de 100MB. ¿Y qué podemos hacer con esos archivos? Pues básicamente se pueden consultar sus datos en formato tabla paginada, aplicar nu-

meros filtros que permiten un grado de personalización muy elevado, podemos incluso definir agregaciones o bien pintar gráficos con ellos, unos gráficos realmente bien presentados, con muchas opciones de visualización e interactivos: siempre podremos clicar en cualquier valor del gráfico para saber el dato relacionado. Incluso es capaz de presentar un Mapa del mundo y colores correspondientes a alguna magnitud.

**Google fusion tables**  
 Share and discuss your data online

Upload small or large data sets from spreadsheets or CSV files.  
 Visualize your data on maps, timelines and charts.  
 Pick who can access your data, hide parts of your data if needed.  
 Merge data from multiple tables  
 Discuss your data with others. Track changes and discussions.  
 Take a tour

©2009 Google - [Google Home](#) - [About Fusion Tables](#) - [Privacy Policy](#) - [Terms of Service](#)

**Population Density** Population Density

File View Edit Visualize Merge

Current view: All - Show table (use columns)

Geographic Area	7/1/2000	7/1/2001	7/1/2002	7/1/2003	7/1/2004	7/1/2005	7/1/2006	7/1/2007
United States	304050724	301290332	298362973	295560548	292802127	290210914	287726647	285328803
Northeast	54824779	54761693	54627987	54531286	54458795	54319451	54134361	53910648
Midwest	69561448	69342592	69047830	68785283	68560340	68299248	68058431	67835632
South	111718549	110395133	109716622	109244182	108745280	104339187	103125430	101838852
West	79854948	68880344	68870634	67999838	67120712	66253188	65408425	64484471
Alabama	4681900	4628995	4587964	4537239	4506574	4488288	4480908	4482832
Alaska	686293	691111	678301	668625	660975	650426	642291	633180
Arizona	6580180	6253421	6178251	5981238	5750475	5585512	5449195	5333632
Arkansas	2855380	2830557	2804199	2769918	2740191	2717889	2701899	2689601
California	36756666	36377534	36121296	35885415	35629666	35307388	34916495	34597930
Colorado	4938456	4842779	4751474	4662734	4600050	4548339	4503156	4431918
Connecticut	3581252	3489988	3467896	3478714	3475351	3467932	3446261	3428206
Delaware	873092	861953	850366	838519	826802	814282	803174	794499
District of Columbia	581633	587988	585419	582048	578621	577371	578112	577678
Florida	18328340	18199526	18116093	17702476	17313811	16937337	16652679	16348734

- Rememberthemilk: <http://www.rememberthemilk.com>
- Delicious: <http://delicious.com>

del.icio.us es un servicio de gestión de marcadores sociales en web. Permite agregar los marcadores que clásicamente se guardaban en los navegadores y categorizarlos con un sistema de etiquetado denominado folcsonomías (*tags*). Pero no sólo puede almacenar sitios webs, sino que también permite compartirlos con otros usuarios de del.icio.us y determinar cuántos tienen un determinado enlace guardado en sus marcadores.

- Slideshare: <http://www.slideshare.net>
- Scribd: <http://www.scribd.com>

Slideshare es un espacio gratuito donde los usuarios pueden enviar presentaciones Powerpoint u OpenOffice, que luego quedan almacenadas en formato Flash para ser visualizadas online. Es una opción interesante para compartir presentaciones en la red. Admite archivos de hasta 20 Mb de peso, sin transiciones entre diapositivas.

Scribd es un nuevo servicio que combina red social con el alojamiento de documentos ofimáticos.

- Acepta los formatos que más utilizamos (doc, pdf, ppt) y otros.
- Podemos obtener el código para insertarlo en nuestros blogs.
- Podemos bajárnoslo en los formatos pdf, word, txt y archivo de audio mp3.
- Nos lee el documento de viva voz. Por ahora es sólo curiosidad divertida pues lee los documentos como lo haría un inglés sin conocimientos de castellano, es decir, casi incomprensible.
- Cualquier usuario lo podrá añadir a favoritos, comentarlo, enviarlo por correo electrónico, votarlo, etc.
- No es necesario registrarse para subir documentos, pero sin ello no podremos acceder a otros servicios.

### Ejemplos 2.0

- Metal 2.0: <http://www.metal20.org>
- Biblioteca Digital Mundial: <http://www.wdl.org/es/>
- Biznar: <http://biznar.com>

### Enlaces

*Último acceso a los enlaces 29-09-2009.*

Google Received 72 Percent of U.S. Searches in January 2009.

[http://image.exct.net/lib/fe1c1774726706/d/1/SearchEngines\\_Jan09.pdf](http://image.exct.net/lib/fe1c1774726706/d/1/SearchEngines_Jan09.pdf)

10 Steps to Selecting the Right CI Software: A Practitioner's Perspective.

<http://brasil.ciphersys.com/10%20Steps%20to%20Selecting%20the%20Right%20CI%20Software-%20A%20Practitioner's%20Perspective.pdf>

MKM Consulting's weblog.

<http://blogmkmconsulting.wordpress.com/>

Informação é um ativo valioso para você e sua empresa?

<http://inteligenciacompetitivainforma.blogspot.com/>

Alcoa alinha Inteligência de Mercado à área de Procurement.

[http://www.metaanalise.com.br/inteligenciademercado/index2.php?option=com\\_content&do\\_pdf=1&id=2111](http://www.metaanalise.com.br/inteligenciademercado/index2.php?option=com_content&do_pdf=1&id=2111)

Pesquisa revela que 88% dos internautas realizam compras pela web.

<http://www.metaanalise.com.br/inteligenciademercado/inteligencia/pesquisas/index.php>

Herramienta informática para vigilancia Tecnológica — Vigtech.

<http://pisis.unalmed.edu.co/avances/archivos/ediciones/Edicion%20Avances%202007%201/15.pdf>

Software para la vigilancia tecnológica de patentes: evaluación desde la perspectiva de los usuarios.

<http://www.elprofesionaldelainformacion.com/contenidos/2006/noviembre/06.pdf>

Herramientas de Software especializadas para Vigilancia Tecnológica e Inteligencia Competitiva.

[http://www.intempres.pco.cu/Intempres2006/Intempres2006/Evaluacion%20de%20trabajos/Jenny%20Marcela%20S%20E1nchez%20Torres\\_Herramientas%20de%20Software...pdf](http://www.intempres.pco.cu/Intempres2006/Intempres2006/Evaluacion%20de%20trabajos/Jenny%20Marcela%20S%20E1nchez%20Torres_Herramientas%20de%20Software...pdf)

Vigilancia Tecnológica de las Herramientas basadas en Software Libre – 2006.

[http://www.alterpime.net/documentos/Vigilancia\\_Tecnologica\\_de\\_las\\_Herramientas\\_basadas\\_en\\_Software\\_Libre.pdf](http://www.alterpime.net/documentos/Vigilancia_Tecnologica_de_las_Herramientas_basadas_en_Software_Libre.pdf)

# Bibliografía

- Blog *La Cofa* <http://www.lacofa.es/index.php/tecnologias/futuro-de-internet/espectativas-de-evolucion-en-los-buscadores> [Consulta: 08-09-2009].
- BRESSÁN G. *Almacenes de datos y minería de datos*. 2003. Disponible en: <http://www/banners/interstitial.html?http://exa.unne.edu.ar/> [Consultado: 29 de julio de 2009].
- BRESSÁN G. *Vigilancia Tecnológica e Inteligencia Competitiva: usos y aplicaciones en la empresa navarra*, 2005. En: <http://www.navactiva.com/web/es/documentacion?tag=210> [consultado el 3 de septiembre de 2009].
- CALDERÓN, Andoni. *La evaluación de páginas web*, <http://observatorio.cnice.mec.es/modules.php?op=modload&name=News&file=article&sid=334> [consulta 11/09/09].
- CALDERÓN, Andoni. *La Búsqueda Semántica* [http://weblogs.madrimasd.org/sistemas\\_inteligentes/archive/2008/05/02/90703.aspx](http://weblogs.madrimasd.org/sistemas_inteligentes/archive/2008/05/02/90703.aspx) [consulta 30-08-09].
- CAMPAÑA AGUILERA, Fermín. Diseño y desarrollo de un sistema de vigilancia tecnológica, proyecto fin de carrera, UPC, 2005 [<http://biblioteca.upc.es/PFC/arxiu/migrats/36419-2.pdf>].
- CANCIO MARTÍNEZ, Elena I.; OVIEDO RIVERO, Iradia; ROMEO LAMEIRAS, Eva y LABRADA LABRADA, Jorge. *Vigilancia Tecnológica herramienta en la gestión de información y el conocimiento en estudios energéticos. Aplicación práctica en investigaciones sobre el uso del hidrógeno como fuente de energía*, Cuba, 2006.
- CODINA, Lluís. *Conferencia inaugural de los Estudios de Información y Documentación de la UOC del segundo semestre del curso 2001-2002* (15 de septiembre de 2001): [http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202\\_imp.html](http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202_imp.html) -[vínculo visitado en agosto de 2009].
- CODINA, Lluís. *Internet invisible y web semántica: ¿el futuro de los sistemas de información en línea?*, revista tradumàtica Traducció i Technologies de la Informació i la Comunicació Número 2: La Documentació: Novembre 2003 <http://www.fti.uab.es/tradumatica/revista>: ISSN 1578-7559.
- Competitive Intelligence at Procter & Gamble. John E. Pepper. *Competitive Intelligence Review*, vol. 10 (4) 4-9 (1999) [<http://lara.rey.googlepages.com/11.pdf>].
- DELGADO FERNÁNDEZ, Mercedes e INFANTE ABREU, Marta. *Vigilancia tecnológica como factor clave para el éxito en la I+D+i: aplicación en el ámbito universitario*, 2009.
- FOAF, 2007. *The Friend of a Friend (FOAF) Project*. Sitio Web accesible en: <http://www.foaf-project.org/>. [consulta: 3/09/09].
- FORNAS CARRASCO, Ricardo. *Criterios para evaluar la calidad y fiabilidad de los contenidos en Internet*, Rev. Esp. Doc. Cient., 26, 1, 2003. En <http://redc.revistas.csic.es/index.php/redc/article/viewFile/226/282> [consulta: 07-09-2009].
- GIMÉNEZ TOLEDO, Elea y ROMÁN ROMÁN, Adelaida. "Vigilancia tecnológica e inteligencia competitiva: conceptos, profesionales, servicios y fuentes de información". En: *El profesional de la información*, 2001, mayo, v. 10, n. 5, pp. 11-20.
- HIDALGO-NUCHERA, Antonio; IGLESIAS-PRADAS, Santiago y HERNÁNDEZ-GARCÍA, Ángel. "Utilización de las bases de datos de patentes como instrumento de vigilancia tecnológica". *El profesional de la información*, 2009, septiembre-octubre, v. 18, n. 5, pp. 511-519.
- intec Factor clave para la toma de decisiones estratégicas en las organizaciones; madri+d* [[http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35\\_INTEC.pdf](http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/35_INTEC.pdf)].
- La inteligencia competitiva: factor clave para la toma de decisiones estratégicas en las organizaciones, colección mi+d nº 35.
- LAMARCA LAPUENTE, María Jesús. *Hipertexto: El nuevo concepto de documento en la cultura de la imagen* [http://www.hipertexto.info/documentos/web\\_invisib.htm](http://www.hipertexto.info/documentos/web_invisib.htm) [enlace visitado el 11-09-2009].
- LEÓN ALIZ, T. *Diseño e implementación de un Sistema de Vigilancia Tecnológica en una empresa de escasos recursos*, Congreso Internacional de Información info 2004, en <http://www.bibliociencias.cu/gsd/cgi-bin/library?e=d->

- 000-00—0eventos—00-0-0—0prompt-10—4——0-11—1-es-50—20-about—00031-001-1-0utfZz-8-00&a=d&c=eventos&cl=CL1.4&d=HASH63935b5e7ca3232eaa7606. [Consultado el 3 de septiembre de 2009].
- LEÓN, Andrés M.; CASTELLANOS, Óscar F. y VARGAS, Freddy A. Valoración, selección y pertinencia de herramientas de software utilizadas en Vigilancia Tecnológica, *Revista de Ingeniería e Investigación*, vol. 26, nº 1, abril 2006.
- MALAYER RODRÍGUEZ, Florentino y VARGAS PÉREZ, Marisela (ed.). *Vigilancia tecnológica y competitividad sectorial: lecciones y resultados de cinco estudios*, ISBN: 978-958-44-1156-3, Bogotá, 2007.
- MALAYER RODRÍGUEZ, Florentino y VARGAS PÉREZ, Marisela (ed.). *Vigilancia tecnológica y competitividad sectorial: lecciones y resultados de cinco estudios*, ISBN: 978-958-44-1156-3, Bogotá, 2007.
- MARTÍNEZ USERO, José Ángel y SANZ MARTOS, Sandra. *A la caza de las patentes: Inteligencia competitiva a través de agentes inteligentes*, 2000.
- MASSÓN-GUERRA, J. L. *Inteligencia Competitiva: Bases Teóricas y Revisión de Literatura* (2005).
- MSc. BOTTA-FERRET, ELEAZAR y Lic. CABRERA-GATO, JANIA E. *Minería de textos: una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital*, *Acimed* 2007; 16 (4) [http://bvs.sld.cu/revistas/aci/vol16\\_4\\_07/aci051007.html](http://bvs.sld.cu/revistas/aci/vol16_4_07/aci051007.html), [consulta: 3 de agosto de 2009].
- MUÑOZ DURÁN, Javier; MARÍN MARTÍNEZ, María y VALLEJO TRIANO, José (2006). La vigilancia tecnológica en la gestión de proyectos de I+D+i: recursos y herramientas. *El Profesional de la Información* 15 (5): pp. 411-419.
- REY VÁZQUEZ, Lara. "Ferroatlántica I+D y la vigilancia tecnológica". En: *El profesional de la información*, 2006, noviembre-diciembre, v. 15, n. 6, pp. 420-425.
- REY VÁZQUEZ, Lara. *Ferroatlántica I+D: Experiencia de una Unidad de Vigilancia Tecnológica*, en: PUZZLE - Año 3, Edición nº 11 mayo-junio 2004, ISSN 1696-8573.
- ROS MARTÍN, Marcos. *Qué es fuente de información en Internet*, post del blog El Documentalista Enredado, Domingo, 30 de abril de 2006 [consulta 04-09-2009].
- SÁEZ DOMINGO, Daniel y ANTOLÍN FDEZ., María. *La vigilancia tecnológica aplicada al sector de tecnologías de la Información y la comunicación: Observatorio tecnológico del ITI*, Fesabid 2009.
- SAGRADO OLIVENZA, I.: *Minería de texto. Recuperación y organización de la información*. 2007. Disponible en: <http://mineriainformacion.50webs.com/recuperacion-informacion.html> [Consultado: 29 de julio de 2009].
- SORIA RAMÍREZ, Verónica. *La literatura Gris y los e-prints*, BIBL. UNIV., NUEVA ÉPOCA, JULIO-DICIEMBRE 2003, vol. 6, nº. 2.
- STATE OF THE ART: COMPETITIVE INTELLIGENCE. A Competitive Intelligence Foundation Research Report 2005-2006. Executive Summary. Edited by Dale Fehringer, Bonnie Hohhof, and Ted Johnson.
- TENA MILLÁN, Joaquín y COMAI, Alessandro. *La inteligencia competitiva en la planificación estratégica y financiera.*, ed. Harvard Deusto [<http://lara.rey.googlepages.com/13.pdf>].
- TENA MILLÁN, Joaquín y COMAI, Alessandro. *Los propósitos de la inteligencia en la empresa: competidora, cooperativa, neutral e individual*. En "El profesional de la información, 2001, mayo, pp. 4-10.
- VELASCO, C. A. y QUINTANA GARCÍA, C. "Inteligencia competitiva, prospectiva e innovación La norma UNE-166006 EX sobre el sistema de vigilancia tecnológica". *Boletín ICE*, 2006.



Lara Rey Vázquez es licenciada por la Universidad de Santiago de Compostela en Geografía e Historia, especialidad archivos y bibliotecas (1997). Ha trabajado en Archivos históricos, bibliotecas y Centros de Investigación. Cabe destacar la Biblioteca Pública de Pontevedra, el Archivo Histórico del Museo de Pontevedra, el Centro de Investigaciones Forestales de Lourizán y el Centro de Innovación y Servicio de Ferrol. Desde el año 2002 es la persona responsable del Dpto. de Vigilancia Tecnológica y la Responsable del Sistema de Gestión de I+D+i de la empresa Ferroatlantica I+D. Ha impartido diversos cursos de Vigilancia Tecnológica e Innovación en varios organismos (ANABAD-Galicia, Ingafor...). Ha publicado algunos artículos y colaboraciones sobre este tema.

# INFORMES DE APEI

1. *Informe APEI sobre web social* por Dídac Margaix Arnal. 2008
2. *Informe APEI sobre acceso abierto* por Julio Alonso Arévalo, Imma Subirats Coll y María Luisa Martínez Conde. 2008
3. *Informe APEI sobre usabilidad* por Yusef Hassan Montero y Sergio Ortega Santamaría. 2009
4. *Informe APEI sobre vigilancia tecnológica* por Lara Rey Vázquez. 2009



<http://www.apeiasturias.org>

[info@apeiasturias.org](mailto:info@apeiasturias.org)

2009