

LA INFORMACION ALIMENTA Y AHOGA

Información si, pero, ¿en qué condiciones?

En el Capítulo 2 del libro propusimos un modelo ideal para el sistema de información de una organización (el *modelo híbrido*). Allí se vió que para que la información *aporte valor* a la organización es preciso que sea información *adecuada a los objetivos* de la organización, que esté estructurada para que pueda ser *convertida en conocimiento*, de manera que ese conocimiento se pueda *aplicar a la toma de decisiones* que lleven a *mejores resultados*.

En este capítulo se tratará con algo más de detalle la importancia de obtener información de utilidad (*auditoría de la información*), de manera sistemática (*inteligencia empresarial*), así como cuales son los instrumentos hoy existentes para la localización y obtención de esa información (*cadena de valor* de la información). Se tratará tanto de *herramientas* como de *métodos*. Y siempre con la idea de que, si bien antes las organizaciones podían quejarse de tener *poca* información, y, por tanto, de no poder tomar decisiones de manera racional, hoy el verdadero problema es que hay *demasiada* información. Y gestionar con demasiada información es quizás peor que gestionar con poca.

Así pues, en el libro se habrá tratado de cómo la información se transforma en conocimiento, y de cómo el punto crítico en todo sistema de información consiste, en el fondo, en conseguir que la información que circula por él sea *relevante* (relevancia y utilidad), esté bien *organizada* (arquitectura de la información), sea *fácil de entender* (visualización y usabilidad), de manera que *la organización la pueda usar* (cultura de la información).

Y es que, si bien es cierto que la información *alimenta* a las organizaciones, si no es suministrada adecuadamente las puede *ahogar*.

Auditoría de la información

Ante la gran cantidad de información que se produce en el mundo, las organizaciones tienen dos importantes retos:

- Primero, conseguir que no entre en la organización más que información que pueda ser de utilidad (más exactamente, que sea *relevante*).
- Segundo, que se obtenga cuanta más información relevante mejor.

En otras palabras, que el *ruido* (obtención de información no relevante) sea mínimo, y que el *silencio* (no obtención de información relevante) sea también mínimo. Conseguir estos dos objetivos no es fácil; para hacerlo se precisa, ante todo, llevar a cabo un *análisis de los requerimientos de información* de la organización (o sea, de la información que precisa para poder funcionar correctamente), o, en otra terminología, una *auditoría de la información*.

Una auditoría de la información es básicamente *un proceso de identificación y evaluación de los recursos de información necesarios para cumplir con los objetivos de la empresa*. Se trata de un paso previo a la determinación de una estrategia de gestión de la información. Puesto de manera muy simple: el objetivo de la auditoría de la información es asegurar que la información que *circulará* por el sistema es la que más conviene a la organización.

La auditoría de la información tiene tres componentes principales:

- Identificación de la información que la organización *precisa* para cumplir sus *objetivos*.
- Comparación de esas necesidades con la información que la organización *ya* está utilizando, de forma que se desvelen necesidades no satisfechas y se valore si los recursos actuales aportan realmente valor a la organización.
- Análisis de los dos puntos anteriores, que permite *definir una estrategia* sobre la información que debe circular por la organización, y puede llegar a detallar de qué forma debe circular.

Así pues, la *auditoría* termina justo dónde empezaría la definición de una *estrategia* de gestión de información.

Hay dos formas básicas de llevar a cabo una auditoría de la información. Por un lado, quién la esté elaborando puede limitarse a *preguntar a los miembros de la organización* en cuestión (más exactamente, a aquellos miembros que la auditoría considere que es preciso contactar; en algunos casos serán sólo los directivos, mientras que en otros se contactará con niveles inferiores en la organización) sobre sus "necesidades de información". Esta táctica conlleva diversos problemas, entre los que podemos citar:

- La pregunta sobre "necesidades" es normalmente confundida, aunque sea a nivel psicológico, con la pregunta "deseos". O sea, el interlocutor no responde sobre *lo que necesita* sino sobre *lo que cree que necesita*.
- El acceso a determinados recursos de información no se solicita en términos de eficacia, sino por una cuestión de *status*. Alguien puede solicitar el acceso a un servicio muy especializado de información

financiera, por ejemplo, con el simple objetivo de *poder decir* que dispone de él, aunque acabe no dedicándole ni un segundo de su atención.

- Cuando alguien es preguntado sobre lo que necesita, tiende a *responder por exceso*; en otras palabras, en esa circunstancia lo más fácil es responder que uno necesita “todo” lo que se produce en su campo de interés, aunque, obviamente, mucho de ese “todo” acaba no sirviéndole para nada.
- *Uno no sabe lo que no sabe*; o sea, si el interlocutor no es un experto en recursos de información en su campo, se limitará a solicitar los recursos de los que tiene constancia que existen, por ejemplo porque ya los ha utilizado con anterioridad. De hecho, la situación presente es una en la que desconocemos más información que nos *podría* ser de utilidad (*ignorancia profunda*) que conocemos información que, aunque no la tengamos, *sabemos* que tenerla nos sería de utilidad (*ignorancia conocedora*).

No hay un método generalmente aceptado para llevar adelante una auditoría de la información. De hecho, las metodologías existentes se adaptan a las características de la empresa a auditar, o sea, se personalizan. Pero, entre los modelos existentes, hay dos que parecen destacar: el *modelo de Orna*, y el *modelo de Burk/Horton*.

El modelo de Orna (1990) tiene cinco partes principales:

- *Motivación de los miembros de la organización*. En esta etapa se trata de presentar cuáles son los objetivos de la auditoría de información, con el fin de eliminar suspicacias y conseguir la máxima participación en el proceso de detección y análisis de necesidades.
- *Retrato de la organización*. Se trata aquí de identificar claramente cuáles son los *objetivos* (prioridades, factores de éxito, etc), la *estructura* (tanto formal, o sea, el organigrama, como la informal, es decir los grupos fácticos o tácitos existentes), y la *cultura* (el “cómo hacemos aquí las cosas”) de la organización. Con este análisis se determina *qué* precisa la organización, y cuál es el *contexto* en el que esas necesidades pueden ser o no cubiertas.
- *Auditoría de recursos*. Este es el verdadero núcleo del estudio, en el que se identifican los recursos de información *existentes* (por ejemplo, a través de la confección de un mapa: qué información entra en la organización, quién genera información, quién usa información, etc), así como se determinan las *prácticas de uso* de esa información (cómo se estructura, cómo se organiza, cómo se difunde, cómo se procesa, etc), y se precisan los *costes implicados* tanto en su adquisición como en su uso.

- *Balance informacional.* En esta parte se utiliza una metáfora contable para ilustrar que se trata de comparar los recursos existentes con los objetivos y con la estructura de la organización (o sea, ¿tenemos la información que necesitamos?, y ¿puede la organización metabolizarla como es preciso?), y también comparar los costes de adquisición y tratamiento de esos recursos con el valor que aportan (o sea, *análisis costes-beneficios*).
- *Plan de acción.* Finalmente, del análisis de todo lo anterior se deriva un plan de *acciones concretas*, tanto en lo que se refiere a *qué* información introducir en el sistema, sobre a *cómo* utilizarla.

Por su parte, el modelo de Burk/Horton (1988) se basa en la identificación de *entidades informacionales primarias* (lo que ellos denominan IRE, *information resource entities*), la evaluación de costes/beneficios de estas entidades, el análisis de la relación entre éstas y la estructura y funciones en la organización, y una síntesis final sobre las ventajas y desventajas de la información existente con respecto a los objetivos de la organización. Todo ello realizado a través de una metodología muy precisa de *mapificación* de recursos de información.

Un modelo simplificado, a la manera de síntesis entre ambos métodos, sería el siguiente. Una auditoría de la información sería un proceso *sistemático* de determinación de la información que la organización requiere *para satisfacer sus objetivos*. Y consistiría en tres procesos básicos:

- *Determinación de los objetivos de la organización.* Este análisis puede hacerse a escala organizacional, departamental o personal. En los tres casos se conseguirá un diferente nivel de concreción. Obviamente, los objetivos de la organización en genérico no tienen porque coincidir con los de las personas que trabajan en ella. Por otro lado, los objetivos de un directivo a nivel estratégico (director general, por ejemplo) no son los mismos que los de un directivo a nivel operacional (director de fábrica, por ejemplo), o que los de un técnico o empleado en cualquier nivel inferior. Sea cual sea el nivel escogido (en el fondo, que nivel de *granularidad* se escoja en la auditoría depende de los recursos, en dinero y tiempo, se hayan destinado a la auditoría), el análisis deberá ser coherente con el mismo; o sea, las propuestas que se hagan deberán ser aplicables al nivel en cuestión. En cualquier caso, se debe notar que no se pregunta qué información *desean*, o cual *creen* que necesitan, sino que la pregunta se limita a sus *objetivos*.
- *Identificación de los factores críticos de éxito.* Se trata aquí de identificar qué cosas deben pasar necesariamente para que se cumplan los objetivos detectados en la fase anterior. Para cada objetivo se puede limitar a tres el número de factores críticos de éxito (FCE). Estos nos ayudan a entender qué acciones son prioritarias para la consecución de los objetivos.

- **Especificación de la información crítica.** Para cada factor crítico de éxito se especifica aquella información de la que es preciso disponer, tanto para poderlo cumplir como para poder hacer un seguimiento de cómo se está cumpliendo el factor. Esta información esencial para el cumplimiento de los FCE, y por tanto de los objetivos de la organización, recibe el nombre de *información crítica*. Puede que en la organización entre mucha otra información, y que la mayoría sea ruido o sea adquirida más por cuestión de *status* que de eficacia. Pero lo que la auditoría tiene que detectar es la información sin la que la organización no puede cumplir eficazmente sus objetivos.

Una vez realizada la auditoría, la organización debe haber mejorado su conocimiento acerca de qué información dispone y de cómo la usa. Esto resulta fundamental para poder diseñar, por ejemplo, una *Intranet* que resulte de utilidad. Sin contenidos útiles (orientados a la consecución de los objetivos), una *Intranet* tiene poco sentido.

Aunque un resultado deseado de la auditoría de la información es que sea posible llevar adelante una estrategia para separar la *información crítica* de la *accesoria* o *fatal*, es decir, para establecer un *filtro* que evite el ruido (Figura A2.1), también es cierto que siempre es preciso dar un cierto margen que permita la obtención de información que sin estar ligada a los objetivos presentes puede contribuir a cambiarlos.

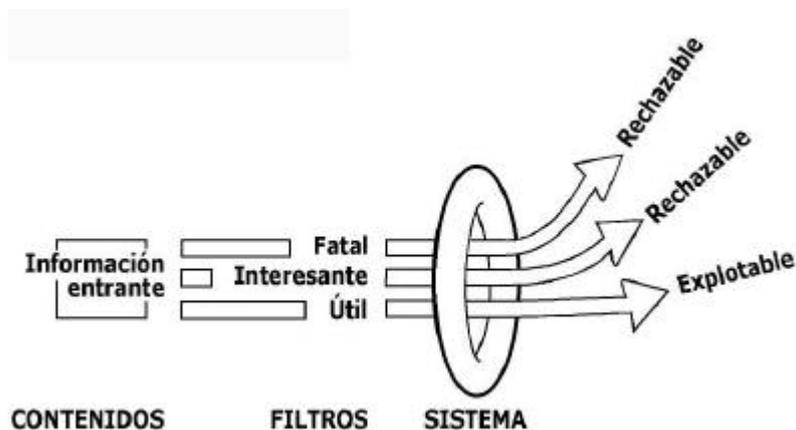


Figura A2-1. El objetivo del *filtro informacional* es que sólo entre en la organización aquella información que resulta de utilidad para el cumplimiento de los *objetivos*

Así, por ejemplo, está comprobado que mucha de la información que nos acaba siendo de utilidad la conseguimos por casualidad, a través de la *serendipia*: nos “tropezamos” con información de interés en muchas ocasiones en las que en

realidad estamos buscando otra información. La organización debe entender que este proceso de *encuentro casual* cumple un papel fundamental en la obtención de información útil. La auditoría debe también determinar hasta qué punto la *serendipia* es importante en la empresa en cuestión, y si lo es, deberá incorporarla de alguna forma en la estrategia final propuesta.

Finalmente, la auditoría de la información está muy relacionada con otras áreas ya tratadas en el texto. Ya se ha visto la relación con el diseño de una *intranet*. También constituye una parte importante en la construcción de sistemas de *inteligencia económica*, de *gestión del conocimiento*, y de gestión del *capital intelectual*.

En cualquier caso, es necesario entender el proceso de auditoría como un *diálogo* permanente entre los usuarios y los profesionales implicados en la realización de la misma, ya que las necesidades de información y los objetivos de las personas y de las propias organizaciones cambian con el tiempo. Así pues, debemos considerar el proceso de auditoría de información como un *proceso permanente*, responsabilidad de todos los miembros de la organización.

¿Espionaje industrial o inteligencia empresarial?

Según se ha visto más arriba, disponer de información sobre lo que ocurre en el entorno de la organización resulta fundamental para su éxito competitivo. Una auditoría de la información nos ayuda a determinar *qué* información es crítica para cumplir nuestros objetivos, y nos dice hasta qué punto esas necesidades objetivas de *información crítica* están satisfechas o no. La auditoría no nos dice, sin embargo, *cómo* debemos conseguir esa información crítica.

Muchas organizaciones creen que son capaces de obtener la información crítica *ad-hoc*. Es decir, que se deben movilizar recursos para obtener esa información sólo *cuando* se presenta la necesidad. Otras, en cambio, piensan que es preciso disponer de un procedimiento *sistemático* de observación del entorno, de manera que la información crítica esté siempre disponible cuando se necesita.

Por otra parte, algunas organizaciones creen que la única forma de obtener información relevante sobre su entorno es a través de procedimientos ilícitos o ilegales. En otras palabras, creen que para obtener información crítica hay que incurrir en el *espionaje industrial*.

Lo cierto es que el espionaje industrial ha sido usado durante siglos como fuente de información sobre la “competencia” (sea lo que sea esta competencia), y que el problema sigue siendo bastante actual.¹ Los datos indican, por ejemplo, que la

¹ Un ejemplo: en julio de 2000, la Bolsa norteamericana castigó a *Oracle* por contratar a una agencia de detectives para espiar, hurgando en sus basuras, a tres organizaciones presuntamente relacionadas con *Microsoft* (*El País*, 09/07/00, p60). El objetivo era demostrar, mediante documentos encontrados a través de

mitad de las empresas norteamericanas son cada año objeto de algún tipo de sustracción no deseada de información. Por ejemplo, tan sólo las pérdidas debidas a robos de información de las 1000 mayores empresas norteamericanas se estimaron en 1999 en unos 45.000 millones de dólares.² Y aunque se tiene tendencia a pensar en el espionaje industrial como una acción de empresas de un país contra empresas de otro, parece ser que la mayoría de acciones se producen entre empresas de un mismo país.³

El espionaje industrial es ya un delito importante en la mayoría de países occidentales. En los Estados Unidos es un delito federal desde octubre de 1996, castigado con penas de hasta 15 años y 10 millones de dólares de multas (véase la *Economic Espionage Act* de 1996). En España, el delito está recogido y penado en el *Código Penal* (artículo 278, 279 y 280).

Sin embargo, la línea entre lo que es lícito y lo que es *ilícito* o *ilegal* es a menudo muy fina. Por ejemplo:

- ¿qué actitud tomar cuando un empleado de una empresa competidora nos viene a vender información de alto interés?
- ¿podemos hurgar en las basuras de un competidor con el fin de indagar en qué están trabajando sus diseñadores?
- ¿es incorrecto "colarse" en una convención de ventas de un competidor, en la que se presenta un producto muy innovador?
- ¿podemos enviar a uno de nuestros empleados a una entrevista de trabajo con un competidor, con el objetivo de detectar nuevas líneas de trabajo?
- ¿podemos contratar de una sola vez unas cuantas decenas de los mejores empleados de nuestro principal competidor, con el fin tanto de absorber su *know-how* para nuestro beneficio como de dejarle sin recursos para responder a sus compromisos con sus clientes?

Por raros que parezcan los casos aquí citados, todos ellos se han producido en empresas muy conocidas en los últimos años. Algunas de estas acciones son delito en algunos países, mientras que no lo son en otros. Pero en todos los

este procedimiento de "arqueología de las basuras" (*Upside.com* 12/07/00) que esas organizaciones "independientes", que clamaban contra el juicio de *Microsoft* por monopolio, no eran en realidad más que "marionetas" del gigante del *software*.

² Según un estudio de la *American Society for Industrial Security* (<http://www.asisonline.org/>) y la consultora *PriceWaterhouseCoopers*. Véase *Business 2.0*, abril 2000, p329. Véase también <http://www.warroomresearch.com/>.

³ Sobre el impacto del espionaje industrial en España, véase el estudio de *Cotec* (<http://www.cotec.es>) titulado *Vigilancia tecnológica e inteligencia competitiva. Su potencial para la empresa española* (1999).

casos, aparte de su legalidad, está claro que muchas de esas acciones vulneran *principios éticos* que deberían ser muy básicos en las empresas.

Justamente, que exista una *ética de la información empresarial*, que determine *qué es lícito*, aparte de que además sea *legal*, en la obtención de información para la empresa resulta conveniente por dos razones:

- Primera, porque deja claras la reglas del juego: si caes en la tentación de aceptar la obtención de información por cualquier procedimiento, estableces un *hábito organizacional* que puede impedir el desarrollo de métodos más lícitos aunque también más complicados (comprar información ilícita es normalmente más fácil que desarrollar un sistema de inteligencia empresarial, como veremos), y además quién compra lo ilícito puede estar estimulando que los suyos también vendan lo ilícito (o sea, algunos empleados pueden pensar que si todo vale al comprar, también vale todo al vender, por ejemplo, la información crítica de *su* empresa).
- Y, segunda, impedir que entremos en el juego del espionaje industrial puede evitar muchos disgustos, en especial desde el momento que esta actividad está perseguida, y penada, y que ello puede generar importantes multas, con frecuencia superiores al potencial beneficio que obtendríamos con la información a adquirir.

Los mejores expertos en la obtención de información empresarial⁴ afirman que hoy es posible obtener la mayor parte de la información que necesita una empresa sobre su entorno a través de procedimientos *totalmente lícitos*. Esa información existe, y está disponible tanto en fuentes *públicas* (por ejemplo, registros mercantiles, registros industriales, de la propiedad, fuentes estadísticas, etc) como *privadas* (bases de datos especializadas, informes comerciales, directorios, etc).

Pero conseguirla no tiene por qué ser fácil: muchos de esos expertos resumen su experiencia diciendo que para obtener la información crítica sobre el entorno (y, en especial, sobre la competencia) “hay que hacer los deberes”. Con esta expresión sintetizan la idea fundamental en el campo de la inteligencia empresarial: los sistemas *ad-hoc* pueden funcionar en ocasiones, pero la mejor garantía de eficacia informacional se consigue cuando se diseña, mantiene y explota un procedimiento *sistemático* de obtención de la información crítica (o sea, la información que una auditoría de la información ha identificado como crítica para el cumplimiento de los objetivos de la organización).

En efecto, se podría definir la *inteligencia empresarial* como un programa *sistemático para recopilar y analizar información sobre el entorno, con el fin de*

⁴ Véase, por ejemplo, *Fuld & Co*, en <http://www.fuld.com/>.

*cumplir los objetivos de la organización.*⁵ La definición tiene tres componentes importantes:

- Primero, se trata de un proceso que debe hacerse de manera *sistemática*; no se trata de algo realizado ad-hoc. Obviamente, si es un proceso sistemático alguien tiene que llevarlo a cabo; en las empresas esta necesidad genera una nueva función, la *función inteligencia*, que puede ser responsabilidad de *un departamento* (el de inteligencia empresarial, por ejemplo) o puede estar esparcida por *distintos departamentos* de la organización (por ejemplo, marketing, diseño, producción, etc).
- Segundo, la función inteligencia no se reduce a obtener la información, sino que es fundamental que se haga con ella algún tipo de *análisis* con el fin de aumentar su valor para la organización.
- Y, tercero, que la información tenga valor depende del grado con que contribuye al cumplimiento de los *objetivos*. Si el procedimiento de obtención de información es legal y ético, hablamos de inteligencia empresarial; si no lo es, hablamos de espionaje industrial.

Frecuentemente se utiliza el caso del sistema de información industrial desarrollado por los conglomerados industriales y comerciales japoneses en los años 50-80 como ejemplo de un enorme sistema de inteligencia empresarial (en algunas ocasiones, se utilizó el término *aspiradora* para describir el alcance y potencia del mismo). Piénsese, por ejemplo, que se ha estimado que durante las décadas de los 70 y 80, los japoneses adquirieron patentes norteamericanas, en particular en las áreas de electrónica de consumo y semiconductores, por valor de un millardo (mil millones) de dólares, gracias a las cuales desarrollaron tecnologías que les generaron un mercado de un billón (un millón de millones) de dólares. O que, a principios de los 90, el 51% de los doctorados concedidos en los Estados Unidos en las disciplinas técnicas y científicas lo fueron a estudiantes procedentes de Asia y el Pacífico (Fialka 1996).

Un buen sistema de inteligencia empresarial aporta valor a la organización porque, entre otras cosas, permite *anticipar* cambios en el mercado, *prever* las acciones de los competidores, *descubrir* nuevos productos y procesos, *aprender* de la experiencia de otros, etc. De manera muy resumida, se podría decir que un sistema de inteligencia permite a la empresa disponer de información sobre lo que *pide* el mercado (el *market pull*, lo que el mercado está mostrando que desea), y sobre las *posibilidades* que las nuevas tecnologías aportan al cambio de productos o procesos (*technology push*, o sea, lo que la tecnología nos permite cambiar en lo que hacemos o en cómo lo hacemos).

⁵ La *SCIP* (*Society of Competitive Intelligence Professionals*) define la inteligencia empresarial (competitiva) como "un programa sistemático y ético para captar y analizar información sobre las actividades de los competidores y las tendencias generales de los negocios, con el fin de promover sus propios objetivos de empresa.

Un sistema de inteligencia empresarial consta de varias partes (que algunos expertos denominan el *ciclo de la inteligencia empresarial*)(Kahaner 1996):

- *Identificación* de la información crítica (que hemos visto que era el objetivo de la auditoría de la información)
- *Recopilación* de la información (o sea, obtención de la información crítica a partir de sus fuentes. Trataremos sobre las distintas fuentes de información más adelante)
- *Análisis* (es decir, interpretación de lo que nos dice la información recopilada, a través de la identificación de patrones, confección de modelos, evaluación de la calidad y relevancia de la información, etc)
- *Distribución* (para asegurarse que llega la información adecuada a todo aquel para el que sea relevante para la consecución de sus objetivos organizacionales)
- *Aseguramiento* (si se han invertido recursos en la obtención, análisis y distribución de información, que aporta así valor añadido, es deseable que esta información se proteja para que no sea utilizada por los competidores, por ejemplo)

Sobre los distintos componentes de este ciclo se ha tratado ya en el libro. Baste decir aquí algo sobre el último componente. Algunas empresas, en especial en el sector financiero, disponen de mecanismos para guardar en lugar seguro su información crítica, o para eliminar de manera segura (o sea, que no pueda ser utilizada por otros) aquella información que ha cumplido su *ciclo de utilidad*. Existen tanto empresas que se dedican a asesorar sobre cómo proteger la información de la organización (la mayoría de estas empresas tienen su origen en agentes procedentes de la *inteligencia militar*), como otras que se dedican a almacenarla en forma segura, y otras a destruirla también de forma segura. En estos casos, el procedimiento definido de seguridad y destrucción es el resultado de una *auditoría de seguridad informacional*.

Un sistema de inteligencia empresarial *puede* resultar de gran utilidad en una organización. Y se ha utilizado aquí la palabra “puede” porque el problema para esta nueva función consiste precisamente en ganarse el derecho a existir, o sea, en *demostrar* que la inversión en el sistema acaba rindiendo. Porque normalmente el beneficio que se deriva de la implantación del sistema es indirecto: mejora la calidad de la información que circula por la organización, se puede planificar mejor, se conoce mejor lo que piden los mercados, lo que hacen los competidores, lo que las tecnologías aportan de nuevo, etc. Pero *medir* económicamente este impacto, dar una cifra aunque sea aproximada, resulta muy difícil. De hecho, este problema no es más que una faceta del problema más general de la valoración del

rendimiento de cualquier inversión en un sistema de información. Los beneficios son normalmente *intangibles*, y por tanto, difíciles de cuantificar. Algunas empresas japonesas afirman, frente a esto, que no tiene sentido cuantificar la utilidad, el beneficio, de la inversión en un sistema de inteligencia empresarial. Para ellas, disponer de él es simplemente, un coste inevitable, incuestionable, para permanecer en el negocio. Es *un coste más del negocio*.

La función inteligencia empresarial requerirá en el futuro de profesionales de la información capaces de discriminar la mejor información existente en el entorno, capaces de adaptarla a las necesidades de la organización, capaces incluso de detectar si una información aparentemente inocua puede tener importantes consecuencias para la misma (para lo cual, debe comportarse más de manera *proactiva* que simplemente *reactiva*, y deberá conocer bien los *objetivos* y los *procesos* fundamentales de la organización). Adecuadamente formado, este tipo de profesional⁶ se convertirá en alguien muy útil a la organización, y su función será considerada un *centro de beneficio* más que un *centro de coste*.

La cadena de valor de la información

Mediante la *auditoría de la información* una organización identifica *cuál* es la información crítica de la que debe aprovisionarse. Mediante un sistema de *inteligencia empresarial* pone en funcionamiento un método *sistemático* de captura, análisis y distribución interna de esa información crítica. En otras palabras, la auditoría define *qué* información necesitamos, y el sistema de inteligencia *cómo* debemos capturarla y procesarla. En este módulo trataremos *dónde* obtener la información, o sea, de las *fuentes de información*.

La información que entra en una organización se podría clasificar de acuerdo con al menos tres criterios:

- Según la fuente de *origen*: la información que utiliza una organización puede provenir tanto de la fuente que la origina (información *primaria*) como puede haber sido elaborada por algún tipo de intermediario (información *secundaria*).
- Según el grado de *formalización*: mucha información entra a través de documentos (información *formal*), mientras que otra gran parte entra en forma de contactos humanos, como conversaciones, signos de todo tipo, comunicación humana no verbal, etc. (información *informal*).
- Según el *sopORTE*: la información puede llegar en *papel*, o en cualquier formato *digital* (texto, sonido, imagen fija, imagen en movimiento, programas, etc).

⁶ Los profesionales que ya se dedican a esta función forman distintas *asociaciones* en el mundo. Véanse, por ejemplo, la *SCIP* (*Society of Competitive Intelligence Professionals*) (<http://www.scip.org/>), la más importante en número de miembros.

Normalmente, la transmisión de información entre la fuente (el «productor») y el usuario (el «consumidor») no se lleva a cabo directamente sino que, como ocurre con el intercambio de cualquier producto o servicio, hay una *cadena de valor* entre la materia prima y el producto final, una cadena a través de la que se va *añadiendo valor* a la información bruta inicial hasta la información con las características y el formato que más se adecuan a lo que precisa el usuario.

Los componentes principales de la cadena de valor de la información son (Figura A2-2): los creadores de información primaria, los productores de bases de datos (y, en la actualidad, los buscadores en el Web), los suministradores de documentos, los intermediarios de información y los servicios de valor añadido.

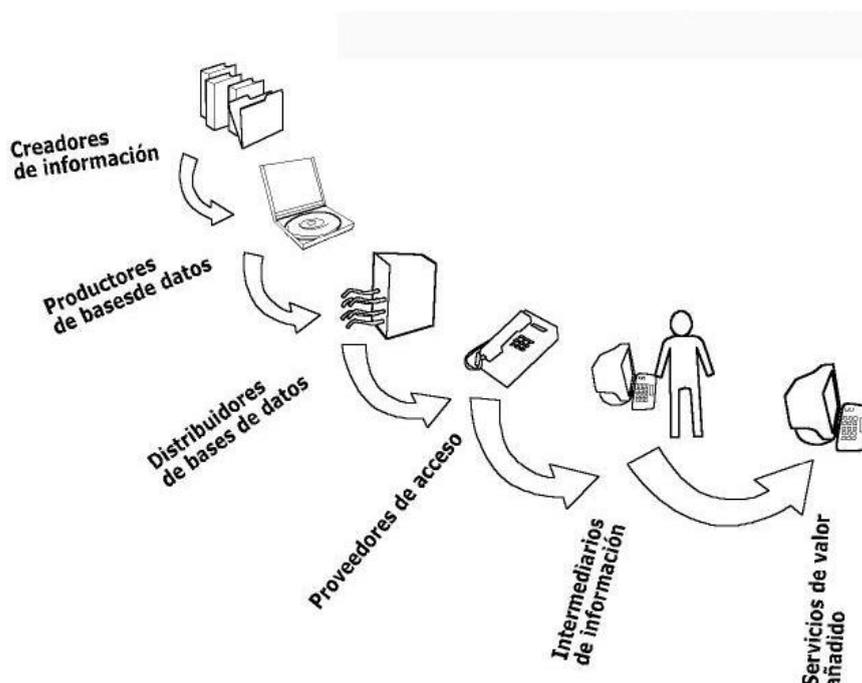


Figura A2-2. Los componentes principales de la cadena de valor de la información

Creadores de información primaria

Son millones los puntos en el mundo dónde se *crea* información. Ya se vió en el Anexo 1 que, por ejemplo, la producción de literatura científico-técnica aumenta de forma exponencial, y cómo el Web crece cada día en millones de nuevas páginas.

En la actualidad, los creadores de información son tanto personas como máquinas. Una persona escribe un artículo, por ejemplo; un satélite captura datos de forma ininterrumpida. Las personas siempre han creado información, pero no ha sido hasta la llegada del Web que un particular ha podido pasar a ser *editor* de sus propios contenidos. Hasta ese momento, “publicar” implicaba que alguien con

capacidad para hacerlo aceptaba difundir tus contenidos en algún medio formal (un libro, un artículo de revista, etc). En la actualidad, cada individuo con acceso y espacio en el Web puede ser editor de sus propios contenidos. Pero, que cualquiera pueda publicar no quiere decir que todo lo publicado sea de calidad, obviamente. Precisamente, el *control de calidad* sobre lo obtenible a través del Web debe ser una de las responsabilidades de los profesionales de la información implicados en el sistema de inteligencia de la organización.

Normalmente, un criterio de calidad de la información consiste en utilizar información que procede de alguna *fuerza de prestigio*. Por ejemplo, de una publicación (impresa o digital) que se comprometa a garantizar la calidad de sus contenidos. Resulta, en efecto, curioso, que la proliferación de información en el Web no haya acabado con las fuentes tradicionalmente consideradas como fiables, sino que incluso éstas han visto reforzado su papel. Es cierto que un autor de prestigio puede constituirse en fuente primaria.⁷ Pero lo normal es que los autores publiquen a través de fuentes que garantizan calidad. En los campos científicos y técnicos, por ejemplo, esto se consigue mediante *revistas arbitradas*, es decir, revistas en las que un colectivo de expertos⁸ determina si un artículo puede ser publicado o debe ser rechazado.

Entre las fuentes de información primaria más frecuentes cabe citar los libros, las revistas y periódicos, los congresos y conferencias, los informes y estudios, las patentes y normas, etc. En rigor, cabría distinguir entre *libros*, *publicaciones periódicas* y *literatura gris*. Se denomina literatura gris a lo que se publica de forma irregular o en fuentes no habituales; tal es el caso de tesis doctorales, informes de investigación, técnicos o comerciales, estudios de mercado, etc. Incluso dentro de un conjunto de fuentes, podríamos distinguir algunas más primarias que otras. Por ejemplo, las estadísticas que emite regularmente un organismo público son más primarias que el resumen que de las mismas publica una revista o un periódico. Los datos o las conclusiones presentes en una tesis doctoral son más fiables que los que se pueda publicar en algún resumen realizado por alguien que no es el autor. Por lo que, en general, es recomendable acercarse lo más posible a la fuente más primaria de la información, para evitar errores o para filtrar interpretaciones realizadas por cualquier intermediario entre la fuente y nosotros.

La principal consecuencia de la espiral de producción de información primaria en el mundo es que resulta cada día más difícil estar *bien informado*, o mejor dicho, *estar completamente informado*. Se ha documentado que un profesional que quisiera estar al día sobre todo lo que ocurre en el mundo acerca de su campo debería leer unas 50 páginas de documentos técnicos al día (y eso antes del fenómeno Internet), lo cual resulta, obviamente, imposible. Es muy frecuente, como consecuencia, estar a la vez *sobrecargado de información* (incluso con consecuencias de angustia y ansiedad) y *desinformado*; hay más papeles que nunca sobre la mesa y menos tiempo para leerlos, mientras que el entorno es progresivamente más complejo y es más difícil proceder por intuición o sin una base firme de información de calidad, suficientemente metabolizada. La

⁷ Algunos lo hacen, como, por ejemplo, Jakob Nielsen, experto en usabilidad, que publica sus propios artículos a través de su página, <http://www.useit.com>

⁸ Denominados *referees* (árbitros) en inglés.

experiencia de muchos profesionales coincide en que con frecuencia resulta difícil, si no imposible, encontrar la información que se necesita en el momento adecuado, y que uno de los mecanismos típicos de obtención de información es una combinación de *espiguelo* (*browsing*: hojear documentos) y *casualidad* (*serendipity*: se encuentran informaciones de interés cuando no se esperaban).

Todo este fenómeno de multiplicación de la cantidad de información existente en el mundo se ha venido a llamar la *explosión de la información*, aunque más bien debería llamarse la *explosión de la desinformación*, porque resulta peor disponer de exceso de información, indigerible y confundidora, que disponer de escasez. Es un problema grave que empieza a afectar incluso a la salud de las personas, en la forma de lo que se denomina *ansiedad de la información*: el deseo de leer todo lo que está a nuestro alcance sin darnos cuenta de la imposibilidad de tal tarea.

Para resolver el problema de la sobrecarga de información y la consecuente desinformación, el mercado ha respondido ideando una serie de instrumentos que facilitan la búsqueda de información que no de su metabolización, que sigue siendo una tarea fundamentalmente humana. Algunos de estos instrumentos se describen a continuación.

Productores de bases de datos

Que exista mucha información en un campo no resulta de mucha ayuda si alguien no se dedica a *organizarla*. Este no es un problema nuevo. Históricamente, la necesidad de disponer de mecanismos para buscar entre una cantidad ingente de información primaria, se evidenció hace décadas en disciplinas que se caracterizan por una gran producción de información, como la medicina, la química, la ingeniería, las patentes, etc. A finales del siglo pasado ya existían en algunos de estos campos organizaciones que se dedicaban a *recopilar* una gran parte de lo que se publicaba sobre una determinada área en los principales idiomas, por lo menos-, y a crear instrumentos que permitieran localizar una información concreta entre todo lo recopilado, a partir de algunas claves, como, por ejemplo, su temática, su fecha, sus autores, el medio en que se publicó, etc.

Su objetivo era, pues, crear fuentes de *información secundaria*, es decir, fuentes que no contenían información nueva sino que servían para buscar y localizar información primaria. Por lo general, estos instrumentos consistían en *boletines de sumarios o índices*, cuya utilidad radicaba en la forma en la que los documentos estaban clasificados, de acuerdo con *códigos* o *palabras claves* (términos que describían el contenido de un documento). Un buen ejemplo de este tipo de índices es el *Chemical Abstracts*, cuya publicación se inició en el siglo pasado y que todavía sigue editándose.

Más tarde, el avance de la informática permitió el desarrollo de *bases de datos*. Estos instrumentos permitían *almacenar* una gran cantidad de información en muy poco espacio, a la vez que ofrecían herramientas para *buscar* (un *lenguaje de interrogación*). Aprovechando esta posibilidad, han ido apareciendo centenares de

bases de datos, relativas a prácticamente todas las disciplinas.⁹ Algunas de estas bases de datos son de origen público y gratuitas, mientras que otras son de uso restringido y/o de pago.

La calidad de una base de datos se puede medir por cuatro factores:

- *Focalización* temática: cuanto más especializada, más relevancia podrá obtenerse potencialmente (véase, más adelante, el módulo dedicado a relevancia y pertinencia).
- *Exhaustividad*: hasta qué punto contiene lo publicado en su área de focalización.
- *Estructura*: cómo están organizados los contenidos, es decir, cuantos campos se utilizan para describir los registros, lo que determina directamente la “buscabilidad” de los mismos.
- Mecanismos de *búsqueda*: o sea, simplicidad de uso y, simultáneamente, potencia del lenguaje de interrogación con el que cuenta.

Todos estos factores son independientes. Así, por ejemplo, una base de datos puede ser muy exhaustiva en su campo, pero las deficiencias de la estructura de los registros o del lenguaje de interrogación, puede hacer que el mecanismo de búsqueda sea muy ineficiente.

Sistemas de búsqueda en el Web

En la actualidad, a las bases de datos “tradicionales”, las producidas por las empresas especializadas en la construcción de fuentes secundarias temáticas, hay que añadir todo el universo de las bases de datos en Internet. Quizás las más conocidas sean los *buscadores* en el Web, como *Yahoo!*, *AltaVista*, *Excite*, *HotBot*, etc.

En efecto, en muy pocos años, el Web se ha constituido en un gran depósito de información. Cada día se añaden a él millones de nuevas páginas web. Es por ello que para el usuario resulta fundamental disponer de herramientas para buscar información relevante.

Lo primero que hay que tener en cuenta es que no existe ninguna base de datos que cubra *todos* los contenidos presentes en el Web.¹⁰ Por tanto la frase “buscar en el Web” no tiene realmente sentido. En todo caso lo que se hace es buscar en alguna *base de datos* que ha cubierto una cierta parte del Web. El interés de esta base de datos se mide por dos criterios básicos: su *exhaustividad* (qué porcentaje

⁹ Una interesante recopilación puede encontrarse en <http://www.internetinvisible.com/>.

¹⁰ Se cree que los mejores buscadores cubren del 15 al 20% del total de páginas existentes. Véase Lawrence & Giles (1999), Schwartz (2000), o el proyecto *Censorware* (http://www.censorware.org/web_size/).

del Web cubre), y la *calidad* de la información (qué información se ofrece de cada recursos encontrado, cómo se ordenan los resultados, etc).

Existen distintos tipos básicos de sistemas de búsqueda, o *buscadores*:

- *Directorios* (o índices): Se trata de servicios que analizan una parte muy limitada del Web (por tanto, exhaustividad limitada), pero que utilizan un equipo de *editores* humanos que *seleccionan* los *recursos* (páginas web, generalmente) a incluir en la base de datos, de acuerdo con unos criterios bien establecidos, y los *indizan* manualmente (y, por consiguiente, la calidad es elevada). Para cada recurso se escribe un comentario sobre sus contenidos, utilidad, etc (contenidos que conforman la ficha del recurso en la base de datos). Los recursos se organizan en forma de un *árbol jerárquico*, organizado de acuerdo con unos temas/subtemas (el término utilizado en estos sistemas es *categoría/subcategoría*) Buenos ejemplos de directorios son *Yahoo!* (<http://www.yahoo.com>), *Encyclopedia Britannica Internet Guide* (<http://www.ebig.com>), o *About* (<http://www.about.com>). La organización jerárquica o arborescente de los directorios facilita el espiguelo (*browsing*) de los usuarios, es decir, la observación de los diferentes recursos propuestos con el fin de contrastarlos con la información buscada.
- *Motores*: La base de datos de estos servicios se alimenta de lo que *softwares de rastreo*, los denominados *spiders*, van encontrando en su monitorización sistemática de la Red. De manera automática, van entrando una detrás de otra en millones de páginas web, cada día, analizan el texto que contienen, lo indizan, y con ello se construye, también automáticamente, un registro que pasa a incrementar la base de datos (o sea, "se ha encontrado tal palabra en tal página"). Los *spiders* de motores importantes, como el *Scooter* de *Altavista*, visita e indiza más de seis millones de páginas web cada día. En los motores de búsqueda no hay clasificación jerárquica, sino que para realizar su búsqueda el usuario pone directamente en un formulario los términos que describen lo que está buscando, y el motor responde con los registros que responden a su petición (palabras encontradas). Aunque estos sistemas acostumbran a ser más exhaustivos, normalmente también su calidad es más baja. Son buenos ejemplos de motores *Altavista* (<http://www.altavista.com>), *Excite* (<http://www.excite.com>), o *HotBot* (<http://www.hotbot.com>).

Hay que tener en cuenta que ambos tipos de sistemas ofrecen una *herramienta de búsqueda*, o sea, un formulario en el que introducir los términos que se están buscando. En el caso de *Yahoo!*, por ejemplo, tanto se puede recorrer el árbol temático como buscar directamente mediante un formulario de búsqueda. Es por ello que hemos utilizado la denominación genérica *buscadores* tanto para directorios como para motores. La principal diferencia radica en cómo construyen su base de datos: los directorios a través de una localización e indización *humana* de los recursos, y los motores a través de softwares de rastreo o *spiders*.

Aunque obviamente útiles para orientarse a través de la inmensa cantidad de contenidos del Web, los buscadores presentan algunas limitaciones:

- Por un lado, las posibilidades de búsqueda son muy limitadas. Por lo general, el sistema sólo permite buscar si un término aparece en el registro de recursos en la base de datos. En el caso de los directorios, el sistema buscará si el indizador humano ha utilizado ese término en su descripción de un recurso. En el caso de los motores, en los que el spider indiza todas y cada una de las palabras encontradas en una página web, el sistema buscará si en alguna página de las indizadas se ha encontrado el término introducido por el usuario. En algunos buscadores se puede utilizar *combinaciones booleanas* de términos (AND, OR, NOT), siempre en un nivel muy elemental. Y muy difícilmente se encuentran buscadores que permitan buscar términos en partes concretas de los registros (o sea, que permitan la *búsqueda por campos específicos*). Las posibilidades de búsqueda de los buscadores están explicadas, normalmente, en un apartado de “Búsqueda avanzada”, situado cerca del botón general de búsqueda. Finalmente, hay que recordar que el éxito en la búsqueda casi siempre se basa en haber determinado previamente con claridad qué es lo que se está buscando. La clave de toda búsqueda es en efecto la combinación de *una buena pregunta*, el lugar adecuado donde buscar, y del tiempo de quién busca para ir *ajustando de manera iterativa* e interactiva lo buscado con lo que se va encontrando.
- Por otra parte, los recursos que el sistema presenta como resultado de una búsqueda están ordenados de acuerdo con un *criterio de relevancia* no siempre claro. En los motores, el criterio de ordenación se corresponde, por lo general, con el *lugar* de la página en la que se ha encontrado el término buscado (por ejemplo, en el título de la página web, o en los titulares de la página, ordenador quizás por el tamaño de la fuente utilizada, etc). Otro criterio muy utilizado consiste en la *frecuencia de aparición* del término; así, una página que contiene muchas veces el término buscado por el usuario aparece primero en la lista de resultados que otra página en la que aparece en menos ocasiones. Sea cual sea el criterio utilizado, normalmente el usuario no es informado sobre cuál es éste, de manera que no puede saber por qué los recursos presentados en primer lugar de la lista son “más relevantes” que los que le siguen.
- Finalmente, a pesar de que a menudo tenemos la impresión de que los buscadores “ven” todo el Web, su visión es más que limitada. Por una parte, ya se ha dicho que los buscadores sólo indizan una parte del Web, de acuerdo con algún criterio de exclusión (partes del Web a las que se decide no incluir en la base de datos). Pero, por otra, existen servicios a los que se accede *a través del Web* pero que no están *en el Web*. Esta parte *invisible*^{11 12} está constituida por aquellas bases de datos cuya información no existe en forma de página web *estática*, sino que se genera como resultado de una búsqueda en forma de página web *dinámica* (es decir, que en respuesta a una búsqueda del usuario, se generan *en ese momento*

¹¹ Véase al efecto <http://www.internetinvisible.com/>.

¹² Algunos autores denominan *Infranet* a esta parte “invisible” del Web.

unos resultados que se plasman en una plantilla web, que es lo que el usuario recibe en su pantalla). Puesto que los buscadores sólo pueden analizar las páginas estáticas, aquellas que *están en el Web*, todos estos contenidos en bases de datos son absolutamente *invisibles*. El buscador puede indizar las páginas estáticas de la base de datos, pero no las dinámicas. Aparte de este problema de invisibilidad por razón del carácter dinámico de las páginas, muchas de estas bases de datos son de acceso restringido, o directamente de pago,¹³ o sea, requieren una *contraseña* de acceso, cosa que normalmente imposibilita la entrada al buscador.

La investigación y desarrollo en el campo de los buscadores es tremendamente activa. A las capacidades ya comentadas se van añadiendo otras nuevas, por lo general en la línea de aumentar la facilidad de uso, la potencia del sistema de búsqueda, o la calidad de la relevancia ofrecida en la organización de los resultados.

En cuanto al *aumento de la facilidad de uso*, algunas aplicaciones a remarcar serían las siguientes:

- *Búsqueda en lenguaje natural*: Se trata de que el usuario pueda realizar su pregunta directamente en lenguaje normal, en forma de frase sintácticamente correcta y no como un conjunto de términos relacionados por medio de operadores booleanos. Este servicio está accesible en algunos grandes buscadores, como *Altavista* (<http://www.altavista.com>), y también en otros menos conocidos como *ElectricMonk* (<http://www.electricmonk.com>), que traduce la frase del usuario en una ecuación de búsqueda booleana que finalmente envía al mismo *Altavista*.
- *Extensión semántica*: En ocasiones, la búsqueda no se restringe a los términos introducidos por el usuario, sino que se *extiende* a otros *sinónimos*, a términos con un contenido semántico equivalente. Este principio es utilizado, entre otros, por el motor semántico (o conceptual) de *Excite* (<http://www.excite.com>). La búsqueda en lenguaje natural comentada en el punto anterior también hace uso de esta aplicación.
- *Traducción en frases frecuentes*: Algunos buscadores como *AskJeeves* (<http://www.aj.com>), interpretan semánticamente la pregunta introducida en lenguaje natural por parte del usuario, y la comparan con una base de datos

¹³ Para entrar en estos servicios se precisa un identificador y una palabra de paso (password), que hay que contratar previamente. Una vez obtenido, se puede empezar a navegar por el sistema. En cuanto a las tarifas, se siguen aquí distintos modelos. Hay sistemas en los que se paga conforme se use, o sea, de acuerdo con la información que se baje el usuario; un buen ejemplo de esto es *Dialog* (<http://www.dialog.com>). Otros requieren un pago anual único, que da derecho a una utilización ilimitada; utilizan este modelo servicios dirigidos básicamente a colectividades, como universidades y empresas, siendo *Proquest* (<http://proquest.umi.com>) un buen ejemplo. Finalmente, frente a este último modelo de audiencia selecta-precio alto, han emergido servicios basados en el modelo audiencia grande-precio bajo; un buen ejemplo es *ElectricLibrary* (<http://www.elibrary.com>), en el que el pago de una cuota anual de unos 60 dólares te da derecho a una utilización ilimitada del servicio, aunque contiene centenares de fuentes (revistas, libros, mapas, periódicos, etc).

de frases ya preparadas (*frases preguntadas con frecuencia*, o *FAQ* - frequently asked questions). El resultado de la búsqueda es una lista de las FAQ que más se acercan a la pregunta original del usuario; éste debe entonces decidir cuál es la que mejor “traduce” su pregunta. Algo parecido ocurre en *Answers* (<http://www.answers.com>).

En cuanto al aumento de la potencia de búsqueda, podemos destacar:

- Una de las líneas de desarrollo más interesante consiste en lo que podríamos denominar *facilitadores de búsqueda*. Se trata de softwares que el usuario carga en su ordenador y que le facilitan la *construcción* en local de una estrategia de búsqueda, que después será lanzada a un buscador, y que también facilitan un *análisis* de los resultados obtenidos mejor que el directamente ofrecido por el buscador utilizado. Algunos ejemplos son *Bullseyes* (<http://www.intelliseek.com>) o *MataHari* (<http://www.thewebtools.com>).
- Otra de las aplicaciones para aumentar la cantidad de resultados consiste en los denominados *metabusadores*, servicios que buscan simultáneamente en varios buscadores. Algunos ejemplos son *DogPile* (<http://www.dogpile.com>), *Mamma* (<http://www.mamma.com>), *MetaCrawler* (<http://www.metacrawler.com>), *Profusion* (<http://www.profusion.com>), *SavvySearch* (<http://www.savvysearch.com>), *Zinternet* (<http://www.zinternet.com>) o *Chubba* (<http://www.chubba.com>).

Y en cuanto a la mejora de los criterios de relevancia, ya se ha comentado que en la actualidad se basan en la “medida” del *lugar* de la página en la que se ha encontrado el término buscado o de la *frecuencia de aparición* del término. Existen aquí algunas propuestas interesantes como los siguientes:

- En el buscador *Google* (<http://www.google.com>), los resultados están ordenados de acuerdo con el número de páginas web con enlaces hacia ellos. En otras palabras, *Google* mide el índice de *popularidad*, en términos de *enlaces*, de una página en el Web. Si hay más páginas que enlazan hacia un recurso que páginas que enlazan hacia otro recurso, el primero de estos recursos aparece primero en la lista de resultados ofrecidos por *Google*.
- En el buscador *DirectHit* (<http://www.directhit.com>), los resultados están ordenados de acuerdo con el número de internautas que, habiendo hecho la misma o parecida pregunta anteriormente en el mismo buscador, los han preferido (han *clickado* en ellos) de entre toda la lista de resultados. En otras palabras, la relevancia se basa en la *popularidad entre los internautas* de una determinada página. Este programa se utiliza en el buscador *HotBot*, de manera que cuando la pregunta que hace el usuario es una de las típicas, a la lista normal de resultados se añade una propuesta con la denominación “Top 10 visited sites for [la pregunta en cuestión]”. Algo parecido ocurre en *Metabusca* (<http://www.metabusca.com>).

- En *RealNames* (<http://www.realnames.com>), los responsables de las páginas web que lo deseen pueden “comprar” determinados conceptos, de manera que cuando alguien hace una pregunta en un buscador que utiliza esta tecnología, el sistema responde con su lista normal de resultados junto con una lista en la que proponen aquellas páginas que han registrado en RealNames alguno de los términos incluidos en la pregunta realizada por el usuario. Puede verse un ejemplo de esto haciendo la pregunta “jeans” en Altavista. La idea es que si alguien ha pagado por registrar ese concepto, es que está muy interesado en que su web se asocie al mismo, es decir, que su web “va del tema en cuestión”, y por tanto, es relevante como respuesta al usuario.
- Finalmente, en *GoTo* (<http://www.goto.com>), los resultados de una búsqueda están ordenados de acuerdo con lo que las páginas están dispuestas a pagar por aparecer antes en la lista. La idea sería que una página muy interesada en salir primero en la lista de resultados a la pregunta “jeans”, “va mucho del tema jeans”, y por tanto es muy relevante para el usuario que buscaba por ese tema.

Aparte de los buscadores vistos hasta ahora, que podríamos designar como genéricos puesto que no se especializan en ningún tema en concreto, y cuyo objetivo es normalmente indizar el mayor número posible de páginas web, existen también *buscadores temáticos*. En este caso, el objetivo es indizar el mayor número de páginas web que tratan de una determinada materia. Por ejemplo, en *Brint* (<http://www.brint.com>) uno puede localizar miles de páginas relacionadas con la gestión del conocimiento. Hay miles de estos motores temáticos en el Web, muchos de los cuales están incluidos en *buscadores recopilatorios*, como *Beaucoup* (<http://www.beaucoup.com>) o *Buscopio* (<http://www.buscopio.com>), *Tematicos* (<http://www.tematicos.com>).

Finalmente, prácticamente todos los buscadores comentados limitan sus búsquedas al Web. Pero en Internet existen otros muchos recursos. Para encontrar algo en ellos existen buscadores especializados. Así, *Tile* (<http://tile.net>) es un buscador de listas o forums (e-lists) y grupos de noticias (newsgroups). Lo mismo ocurre con *Topica* (<http://www.topica.com>). Y en *FAQS* (<http://www.faqs.org>), se encuentra una lista de centenares de newsgroups especializados en FAQs, o sea, lugares en los que se responden a preguntas típicas sobre problemas que el usuario encuentra en su navegación por la Red.

El campo de los buscadores es uno de los más efervescentes en Internet, y en él tenemos que esperar una importante evolución en los próximos años.¹⁴

Suministradores de documentos

Aunque la mayoría de bases de datos actuales contienen la versión *íntegra* de los documentos, siguen existiendo algunas que contienen sólo sus *referencias*, o sea

¹⁴ Un lugar de referencia para aprender permanentemente sobre el tema es *SearchEngineWatch* (<http://www.searchenginewatch.com>).

información bibliográfica que permite localizar el documento sin ofrecer el contenido entero. En este caso, es preciso recurrir a algún depósito que pueda proporcionar copias de los documentos originales. Se podría designar a estos depósitos con el término genérico de *centros suministradores de documentos*.

Estos centros son de dos tipos principales. Por una parte, existen las grandes *bibliotecas* que mantienen enormes colecciones de revistas, estudios, informes, etc. En Europa, por ejemplo, se utilizan con frecuencia los servicios de *Document Supply Center* de la British Library, o los de la Cámara de Comercio de París. En España, la *Biblioteca del CINDOC* ofrece servicios similares.

Por otra parte, existen empresas que se dedican a proporcionar copias de documentos originales. Algunas se concentran en un área de especialidad, como las patentes, las normas, la documentación química, la legislación, etc. Otras son más generalistas.

Los centros suministradores de documentos facilitan copias de documentos originales en dos modalidades.

- En la primera, se suministran copias para el uso personal de los usuarios; se aplica en este caso la norma del «uso honrado» (*fair use*), según la cual no se permite la reproducción de fragmentos de libros pero se permite la copia de artículos de revista si su uso se limita al que hace el usuario con fines de investigación o estudio.
- En la segunda modalidad, se permite hacer reproducciones para usos no individuales si previamente se abonan los derechos de copia. En este sentido, se van generalizando en los países occidentales los centros de administración colectiva de estos derechos, como el *Copyright Clearing Center* en los Estados Unidos, o el centro *CEDRO* en España.

El hecho de que se esté avanzando rápidamente en el campo de la protección de contenidos digitales, o sea, que vayan apareciendo soluciones técnicas para impedir la copia ilegal de documentos en formato digital, hace que sea cada vez más frecuente encontrar documentos en texto completo directamente en el Web, de manera que los servicios de acceso a documento original van perdiendo, en muchos casos, su razón de ser.

Intermediarios de información

Que existan distintos agentes implicados en la cadena de valor de la información no es más que el resultado de que la oferta es muy general (hay muchos productos, tanto de información primaria como secundaria) y la demanda es siempre muy particular (cada necesidad de información suele ser diferente, propia de quien la plantea en forma de pregunta, y dependiente del momento en que lo hace). En otras palabras, entre la materia prima (la información) y el producto final (alguien que sale informado) se requieren distintos intermediarios que faciliten el proceso de localización, transmisión y aplicación de la información.

Así, por ejemplo, para responder a una pregunta como «experiencias de compensación económica de empleados en programas de calidad total», no basta con tener un ordenador conectado a la línea telefónica; debe existir alguien que haya escrito sobre el tema, alguien que haya referenciado el documento en una base de datos, alguien que distribuya esta base, y alguien que suministre copias del documento en cuestión una vez ha sido localizado. Y, quizás lo más importante, alguien que sepa dónde hay que buscar (en cuál de las miles de bases de datos disponibles) y cómo hay que plantear la pregunta.

Este último intermediario puede tener distintas caras, según el nivel de valor añadido que dé al informando. Así, una *biblioteca* es un intermediario que, por lo general, se limita a indicar en qué sección se puede encontrar un libro que «puede que» tenga la respuesta a una pregunta: el nivel de valor añadido es bajo. Y, en el otro extremo, un *consultor* puede que responda a una pregunta que se le ha planteado no solo con información, sino con un consejo concreto; el nivel de valor añadido puede ser aquí muy alto. En los estadios intermedios hay mucho espacio para otros intermediarios de localización de información: desde el *teledocumentalista* (que sabe utilizar las bases de datos en *línea*) al *investigador de mercados*.

Por lo general, se designa como *intermediario de información* (*infomediario* o *agente de información*, todas ellas traducciones aproximadas del término inglés *information broker*) a este profesional de la localización de información en las fuentes, a pesar de que, como se ha visto, todos los demás elementos de la cadena de valor de la información pueden ser considerados, en cierto modo, como intermediarios. Normalmente, el intermediario es alguien experto en navegar por las fuentes de información, generalmente dentro de un área de especialización. Puede ser que ponga sus servicios a disposición de los usuarios de un centro de documentación, sin que medie pago alguno, o puede que venda su habilidad para buscar información a cambio de unos emolumentos.

Por ejemplo, y aunque se ha dicho que los buscadores en el Web son cada vez mejores, existen en el propio Web servicios de búsqueda realizados por humanos. Podríamos denominar justamente *buscadores humanos* a estos servicios. Se trata de expertos en información que interpretan la necesidad de información del usuario, la traducen en una ecuación de búsqueda en distintos buscadores, y que analizan los resultados para ordenarlos de acuerdo con un criterio de relevancia útil. Son ejemplo de estos servicios *HumanSearch* (<http://www.humansearch.com>), un servicio gratuito, o *KimSal* (<http://www.kimsal.com/search/>).

Uno de los retos importantes de la intermediación en información consiste en entender que el proceso de «informar» es algo más complejo que simplemente «llenar la cesta vacía de quien plantea la pregunta» (idea conocida como *paradigma de la cesta vacía*). Que alguien resulte «informado» como resultado de un proceso de información depende fuertemente de factores como su motivación, capacidad y voluntad de absorber, conocimiento previo, estado de ánimo, la «química» entre el informador y el informando, etc. Ello implica que la utilidad del intermediario no depende sólo de su habilidad técnica como navegante por las fuentes, sino de su capacidad para hacer del proceso de información un proceso *constructivo* en el que participe activamente el informando. El intermediario debe

convertirse en una especie de *asesor o consultor de información* del cliente para que su función pueda desarrollarse con plenitud.

El intermediario de información es, hoy por hoy, una pieza clave de la cadena de valor de la información, posiblemente quien le da sentido y utilidad, al menos hasta que se disponga de instrumentos de búsqueda de información suficientemente sofisticados para que cualquiera pueda moverse con éxito por el universo de la información sin una preparación específica.¹⁵

Servicios de valor añadido

La intervención del intermediario de la información es aún necesaria porque no se dispone de instrumentos que determinen cuál es la fuente más adecuada para responder una determinada pregunta. Por otra parte, son escasos los lenguajes de interrogación en lenguaje natural. Pero esto cambiará conforme las tecnologías avancen. El futuro nos traerá posiblemente sistemas a los que interrogaremos preguntándoles directamente mediante nuestra voz, que de acuerdo con la petición determinarán *cuál* es la fuente más adecuada dónde buscar, y que, finalmente, de toda la información que obtengan de esa fuente *filtren* aquella que sea más relevante para nuestros objetivos.

Algunos servicios ya disponibles en la actualidad van en esta dirección:

- *Servicios de personalización* de información: permiten definir qué información se desea, de manera que cualquiera otra sea eliminada. Casi todos los grandes buscadores ofrecen esta posibilidad. Véase, por ejemplo, el caso de *MyYahoo!*. Hay sistemas que filtran la información que debe ser introducida en la intranet de una empresa, como *NewsEdge* (<http://www.newsedge-international.com/>). En España hay servicios innovadores en esta línea como *MyNewsOnline* (<http://www.mynewsonline.com/>). Un caso especialmente interesante para realizar inteligencia empresarial sobre grandes empresas de la competencia es *CompanySleuth* (<http://www.company.sleuth.com>).
- En algunos casos, se ofrece también un *servicio de alerta*, de manera que se recibe a través del correo electrónico aquellas noticias que hemos destacado como preferentes en nuestro perfil informacional. Véase por ejemplo *MyAlert* (<http://www.myalert.com/>), *Alerts.com* (<http://www.alerts.com/>), o *Latinia* (<http://www.latinia.com/>).

¹⁵ Existen en el mundo diversas asociaciones de intermediarios de información. En Estados Unidos existe la *Association of Independent Information Professionals* (<http://www.aiip.org/>), y en Europa hay los profesionales se agrupan en *Eirene* (<http://www.eirene.com/>). Uno de los objetivos principales de estas asociaciones consiste en definir *códigos éticos* (deontológicos) para sus miembros (qué métodos son lícitos para conseguir información, cómo debe ser la relación de confidencialidad y confianza con el cliente, etc). Existe un directorio mundial de intermediarios de información, el *Burwell Directory of Information Brokers*, en <http://www.burwellinc.com/>.

- *Servicios push*, como los que en su momento pudo de moda *PointCast* (hoy desaparecida), que presentaba en pantalla las noticias correspondientes a nuestro perfil, aprovechando los momentos en que no utilizábamos en ordenador para conectarse a Internet y bajar esos contenidos.
- *Servicios de visualización*, o sea, servicios que facilitan el manejo de información mediante su transformación en gráficos o imágenes. Un buen ejemplo es *NewsMaps* (<http://www.newsmaps.com>), que muestra un mapa de las noticias más importantes del día a través de un mapa donde las noticias con más titulares son los picos nevados y las menos relevantes están en los valles. En una línea parecida, *SmartMoney* (<http://www.smartmoney.com>) muestra la evolución de los valores en la bolsa a través de un mapa del mercado, en el que es muy fácil identificar la importancia de cada segmento en la capitalización total, y donde se identifican mediante colores los valores que suben (verde) y los que bajan (rojo).
- *Agente o bots*, programas diseñados para encontrar información.¹⁶ Algunos, como los *spiders* utilizados por los buscadores, están programados para ir recorriendo la Red en busca de novedades en los servidores. Otros, son *shopbots*, o robots de compra; programas de este tipo, como *Jango* (<http://www.jango.com>) o *Roboshopper* (<http://www.roboshopper.com>), permiten al usuario buscar el web donde encontrar el precio más barato para un producto. Otros son programados por un usuario y son enviados a la Red en busca de una información muy concreta (o sea, son enviados en misión de localización y captura de información). Un buen ejemplo es *DailyBriefing* (información al respecto en <http://www.autonomy.com>). Otro ejemplo es *WebWatch* (información en <http://www.ewatch.com>), que permite detectar los cambios en páginas web previamente especificadas por el usuario. O el anteriormente citado *CompanySleuth* (<http://www.company.sleuth.com>), que te envía diariamente un informe sobre todo lo que ha aparecido sobre una empresa en las oficinas de marcas, de patentes, en los newsgroups, etc (aunque actualmente sólo funciona en los Estados Unidos, y para un número limitado de empresas).

Muchos expertos creen que este es el campo de la cadena de valor de la información en el que veremos los más espectaculares, y útiles a la vez, avances en los próximos años.

¹⁶ Una recopilación de información sobre agentes y *bots* en general, está disponible en *BotSpot*, <http://www.botspot.com/>.

Bibliografía

Buchanan, S, y Gibb, F. (1998) The information audit: an integrated strategic approach. *International Journal of Information Management*, v18, n1, p 29-47.

Burk, C.F. y Horton, F.W. (1998) *Infomap : A Complete Guide to Discovering Corporate Information Sources*. EnglewoodCliffs, NJ: Prentice Hall.

Burwell, Helen et al. (1999) *Online Competitive Intelligence : Increase Your Profits Using Cyger-Intelligence*. Facts on Demand Press.

Ellis, D. Et al. (1993) Information audits, communication audits and information mapping: a review and survey. *International Journal of Information Management*, 13(134-151).

Fialka, John J. (1997) *War by Other Means*. New York, NY: W.W. Norton.

Fuld, Leonard M. (1995) *The new competitor intelligence*. New York, NY: John Wiley

Hock, R. (1999) *The extreme searcher's guide to Web Search Engines*. Medford, NJ: CyberAge Books.

Jakobiak, François (1988) *Maîtriser l'information critique*. Paris, F: Les Éditions D'Organisation.

Kahaner, L. (1996) *Competitive Intelligence*. New York, NY: Simon.

Lawrence, S. Y Giles, L. (1999) Accesibility and distribution of information on the Web. *Nature*, n400, p107-109. Resumen en <http://www.wwwmtrics.com>

Martinet, B, Marti, Y-M. (1995) *L'intelligence économique*. Paris, F: Les Editions d'Organisation.

Martinet, Bruno; Ribault, Jean-Michel (1989) *La Veille Technologique Concurrentielle et Commerciale*. Paris, F: Les Éditions D'Organisation.

Miller, Jerry P. (2000) *Millennium Intelligence: Understanding and Conducting Competitive Intelligence in the Digital Age*. Information Today.

Mutch, A. (1997) Information literacy: an exploration. *International Journal of Information Management*, v17 n5 p377-386.

Orna, E. (1990) *Practical Information Policies*. Aldershot, UK: Gower.

Palop, F, y Vicente, J.M. (1999) *Vigilancia tecnológica e inteligencia competitiva. Su potencial para la empresa española*. Madrid, E: Fundación COTEC <http://www.cotec.es>.

Rugge, Sue; Glossbrenner, Alfred (1997) *The Information Broker's Handbook*. New York, NY: McGraw-Hill.

Sammon, William L.; Kurland, Mark A.; Spitalnic, Robert (1984) *Business Competitor Intelligence : Methods for Collecting, Organizing, and Using Information*. New York, NY: Ronald Press.

Schwartz, C. (2000) Notes from de Boston 2000 Search Engine Meeting. *Bulletin of the ASIS*, aug-sep, p26). Accesible en <http://www.asis.org/Bulletin/Aug-00/schwartz.html>

Special Libraries Association (1995) *The Information Audit: An SLA Information Kit*. Washington, DC: Special Libraries Association.

Stanat, Ruth (1990) *The Intelligent Corporation*. New York, NY: Amacon.

Weitzen, H.S. (1988) *Infopreneurs : Turning Data into Dollars*. NewYork, NY: John Wiley.

Winkler, Ira (1997) *Corporate Espionage : What It Is, Why It Is Happening in Your Company, What You Must Do About It*. Rocklin, CA: Prima Publishing.

Te recordamos que el libro de Alfons Cornella *Infonomia.com: la empresa es información* (Ediciones Deusto, 2000) está enteramente dedicado a la gestión de la información en las organizaciones.