

THE DARK SIDE OF GOOGLE

*(texto original Ippolita- copyleft 2007 – Creative Commons 2.0. by-nc-sa.
www.ippolita.net; info@ippolita.net; traducción al español enero 2008)*

El lado oscuro de Google

Índice

Introducción

I. La historia de un éxito / Historias de motores

1. Motores y búsquedas
2. El nacimiento de Google: en principio fue el garaje, más bien, la universidad
3. Google.com: publicidad directa dentro de las páginas
4. El self-service publicitario: sobrevivir a la burbuja de la new economy
5. Estilo, morfología y multiplicación de los servicios
6. Google, el gigante bueno, cotiza en bolsa
7. Google, Inc.: el monopolio de la búsqueda.

II. BeGoogle

- La huida de cerebros hacia Google: la guerra para el control de la web
- Largas colas en las Redes: las entrañas de la economía de la búsqueda, Google vs.

Microsoft

- La guerra de los estándares
- Arma1: Googleplex, el capitalismo blando
- Arma2: perfeccionar la estrategia de la acumulación
- Arma3: La imagen lo es todo, pero algo de filosofía no sobra
- Arma4: Google y el Open Source

III. Google Open Source: teoría y prácticas

- Open no es free
- Los hackers de Stanford
- La era de la Open Source economy
- Seducir a los hackers: autonomía, dinero fácil, instrumentos gratuitos
- Ambientes híbridos entre universidad y empresa

IV. Algoritmos ;qué pasión!

- Algoritmos y vida real
- La estrategia de la objetividad
- Spider, base de datos y búsqueda
- De la brand-identity a la interfaz participativa
- PageRank, o la autoridad absoluta de un mundo cerrado
- PageRank, o la moneda de la ciencia

V. Además, otras funcionalidades maliciosas

- Los filtros en los algoritmos: bancos de datos preconfeccionados y control de los usuarios

- Los Cookie de Google, galletas que dejan huella
- Onanismo tecnológico: crea, recrea y consume tus contenidos
- Browser como ambientes de desarrollo
- Privacy, paranoias y poderes

VI. Calidad cantidad relación

- La emergencia de la información
- Calidad y cantidad
- El mito de la búsqueda instantánea
- Detrás del velo del mito
- Modelos de búsqueda

VII. Tecnocracia

- Tecnocracia: los expertos de la ciencia
- Milagros de la tecnología: de las opiniones subjetivas a la verdad objetiva
- Público y privado
- Líneas de fuga: p2p, medias independientes, criptografía, blog, foaf

CONCLUSIONES

APÉNDICES:

El fin del mundo en una galleta

Interzonas: influencias y dominios en los mundos digitales.

LICENCIA

Introducción

Google es el motor de búsqueda más conocido y utilizado en todo Internet, hasta tal punto que en los últimos años se ha convertido en el principal punto de acceso a la Red. De forma progresiva, los internautas se han ido adaptando a su interfaz sobria y confortante, a las inserciones publicitarias apartadas y omnipresentes; han adoptado sus confortables servicios y su uso rutinario se ha transformado en forma de comportamiento: “Si no lo sabes, pregúntale a Google”. Se recurre a Google incluso cuando se podría utilizar el *post-it* pegado a la nevera, buscar en la agenda, consultar las Páginas Amarillas, u hojear la enciclopedia que está recogiendo polvo en las estanterías, junto con las demás obras de consulta de papel, demasiado duras y difíciles de examinar.

Google ha sido capaz de explotar magistralmente nuestra necesidad de sencillez. Google anhela ser el motor de búsqueda perfecto, capaz de entender exactamente las preguntas de los usuarios y devolverles, en un santiamén, exactamente lo que desean. Las inocentes interfaces, que ya pueden ser altamente personalizadas directamente en la Web, y aun así inexorablemente identificables por su estilo minimalista, constituyen una forma de huida cotidiana desde la claustrofobia de las mesas digitales para un número impresionantes de usuarios, en constante alza. Son una bocanada de aire fresco, ventanas privilegiadas abiertas de par en par sobre el fascinante mundo de la Red. ¿Cuántas personas utilizan Google como página inicial de su browser? Sin embargo, detrás de esa sencillez y facilidad de uso se oculta un coloso, un sistema increíblemente complejo y sutil para la gestión de los conocimientos del *mare mágnum* de la Red. Google ofrece decenas de servicios gratuitos, para satisfacer todo deseo de búsqueda y comunicación: e-mail, chat, newsgroup, sistemas de indexación de los **file** en el ordenador, archivos de imágenes, videos, libros y mucho más. ¿Por qué? ¿Qué gana? Criticar a Google mediante una ponderación meticulosa de su historia y la de-construcción de los objetos matemáticos que lo componen representa una ocasión para desvelar una precisa estrategia de dominio cultural. Este estudio ofrece la oportunidad de usar un método de estudio más general y útil al descubrimiento de los tejemanejes de muchas de las aplicaciones que nos hemos acostumbrado a utilizar.

El volumen arranca con una breve panorámica sobre la historia de los motores de búsqueda, examina por tanto los momentos más significativos del avance de Google. Tras sobrevivir sin perjuicio al estallido de la burbuja de la *new economy*, Google ha entretejido relaciones sólidas con distintas multinacionales de la *Information Technology*. La continua expansión de sus actividades en cada sector de las comunicaciones digitales está difundiendo un inconfundible estilo y está modelando un completo universo cultural, el universo de la Web.

“Don’t be evil” (no seas malo) es el lema de Sergey Brin y Larry Page, los dos fundadores de Google. Los ex-estudiantes de Stanford, gracias a una meticulosa gestión de su propia imagen, han creado un “Gigante Bueno”, impaciente por archivar nuestras “intenciones de búsqueda” en sus ilimitadas bases de datos. El alter-ego digital de millones de usuarios parece estar en buenas manos, confiado al principal *datacenter* de Mountain View, California, conocido como Googleplex. Aquí, como en los demás centros de archivo de datos de Google –que están surgiendo como setas en todo el mundo- se perfeccionan verdaderas armas para combatir la guerra para el control de las redes. En primer lugar, se difunde la práctica del capitalismo blando de la abundancia: se trata de una estrategia de control biopolítico en sentido estricto, que ofrece ambientes de trabajo confortables, palmaditas en las espaldas y gratificaciones a los dependientes. Los trabajadores, satisfechos y lisonjeados, se dejan explotar alegremente y se convierten en los mayores sostenedores de la empresa, orgullosos de promocionar una imagen ganadora y “buena”.

Los objetivos y los métodos de Google son buenos para todo el mundo; de hecho, la filosofía de la empresa, basada sobre la excelencia de matriz académica, en la apuesta para la

innovación y la investigación, se encuentra expresada en las diez rápidas verdades en el mismo sitio del motor de búsqueda. Estos diez mandamientos constituyen una especie de buena nueva de la era informática, el Google-pensamiento, divulgado con la ayuda de verdaderos y auténticos “evangelizadores” (*evangelist*), es decir personalidades punteras del mundo informático. Última arma, aunque no la menos importante, es la cooptación de las metodologías de desarrollo cooperativo típicas del Open Source y el uso de software libres, no protegidos por copyright o patentes, como base para sus propios productos. De esta forma Google reduce los costes para la implementación de sus servicios, se asegura el apoyo de técnicos, manitas y hacker de todo tipo y se hace pasar por el defensor de la causa de la libre circulación de los saberes, ya que el uso del motor de búsqueda parece ofrecer la mejor forma de acceso gratuito a la Red.

Pero el sueño de Brin y Page de “Google contenedor de todo Internet”, cultivado desde los tiempos de la universidad, es nada más que una idea demagógica, útil para afirmar un culto casi positivista de la objetividad científica: en el caos de la Red solo una técnica superior puede hacerse garante de la transparencia de los procesos, de la exactitud de las respuestas, incluso de la democracia.

De hecho Google se autodefine como un instrumento “democrático”, basado en el presunto carácter “democrático” de la Web. Su algoritmo de indexación de la Red, PageRank (TM), se encarga de copiar los datos digitales en los *datacenter*, explotando las conexiones asociadas a cada página singular para determinar su valor. En la práctica, Google interpreta una conexión de la página A a la página B como un voto expresado por la primera sobre la segunda. Pero no se limita a calcular el número de votos, o conexiones, asignadas a una página. Google analiza la página que ha asignado el voto: los votos expresados por páginas “importantes” tienen más relevancia y por lo tanto contribuyen a dar “mayor valor” también a las páginas linkadas. El PageRank asigna a los sitios Web importantes y de alta calidad un voto más elevado, utilizando filtros y criterios no públicos, que Google toma en cuenta cada vez que realiza una búsqueda. La “democracia” de Google ordena, pues, la Red en base al número de votos recibidos por cada página, y por la importancia de esos votos: una democracia filtrada por la tecnología.

Existen algunos secretos alrededor del coloso de Mountain View, muchos de los que, como veréis, son secretos de Pulcinela. El halo de leyenda que envuelve la tecnología googoliana está basado en gran parte sobre la ausencia de una educación de base, de rudimentos prácticos para enfrentarse culturalmente a la onda larga de la revolución tecnológica. Por ejemplo, la extraordinaria rapidez de los resultados de búsqueda es producto de una cuidada selección para nada transparente. De hecho, ¿cómo podrían millones de usuarios hojear contemporáneamente en todo instante toda la base de datos de Google si no existieran filtros oportunos para reducir el ámbito de la búsqueda, por ejemplo limitándolo a los datos en su idioma de origen? Y si existen filtros creados para garantizar una mejor navegación lingüística ¿no es lícito suponer que existan muchos más filtros, estudiados para canalizar incluso las preferencias de los internautas? El prodigio de Google es en realidad una tecnología opaca y encubierta por el copyright y por acuerdos de no divulgación de sus recursos. La búsqueda no es ni transparente ni democrática como se nos quiere hacer creer: no podría serlo tanto por razones técnicas como por razones económicas.

El campo blanco de Google en el que se insertan las palabras claves para las búsquedas es una puerta estrecha, un filtro para nada transparente, que controla y dirige el acceso a las informaciones. En cuanto mediador de información, un simple motor de búsqueda se hace instrumento de gestión del saber y se halla pues capacitado para ejercer un poder enorme, convirtiéndose en una autoridad absoluta en un mundo cerrado. El modelo cultural de Google es por lo tanto expresión directa de un dominio tecnocrático.

Con este volumen Ippolita quiere subrayar el problema, o más bien la urgencia social de alfabetización y orientación crítica del gran público acerca del tema de la gestión de los conocimientos (*Knowledges management*). Internet ofrece extraordinarias oportunidades de

autoinformación a los usuarios, tantas como para superar incluso la formación universitaria, especialmente en ámbitos como la comunicación y la ingeniería informática. El movimiento del Software Libre, como ha mostrado Ippolita en sus precedentes trabajos, es el ejemplo más evidente de la necesidad de autoformación continua y de la posibilidad de autogestión de los instrumentos digitales.

Sin embargo, existe el reverso de esta moneda, doblemente negativo: por un lado, el menosprecio de la formación humanista, que en la Red dispone de pocos y mal organizados ámbitos de referencia; por el otro, el sustancial colapso cognitivo del usuario medio. Desorientados por la redundancia de los datos disponibles en la Red, nos entregamos a aquellos puntos de referencia de mayor visibilidad –de los que Google solo es el ejemplo más patente– sin preguntarnos qué acontece detrás de los bastidores; registramos nuestros datos personales con ligereza, seducidos por el mero uso de servicios decididamente eficaces y, como todavía acostumbra buena parte de la Red, totalmente gratuitos.

Ippolita intenta señalar el vacío, todo italiano, en la divulgación científica de los fenómenos tecnológicos que afectan el conjunto de la sociedad. Abundan los manuales técnicos, la sociología habla con desenvoltura de Sociedades en Red, la política se atreve incluso a imaginar una futurible Open Society, en la que las redes serán el sustrato tecnológico de la democracia global. Pero, ¿cuántos asiduos internautas saben qué es un algoritmo? Más bien pocos, y sin embargo, la mayoría de ellos confían en el responso de PageRank, precisamente un algoritmo, que ordena sin pausa los resultados a sus preguntas y dirige su experiencia en la Red. Hace falta el valor de centrarse de nuevo en la divulgación científica, sin encerrarse en la torre de marfil del saber académico. Hace falta hablar de macroeconomías sin ser economistas, de infomediación sin ser expertos en comunicación, de autoinformación sin ser educadores, de autogestión de los instrumentos digitales sin ser políticos. Es preciso provocar debates insistiendo sobre conceptos básicos como “algoritmo”, “datos sensibles”, “privacy”, “verdad científica”, “redes de comunicación”, demasiado a menudo debatidos entre Authority y Garantes que no pueden garantizar absolutamente nada.

La costumbre a la delega provoca un desinterés generalizado hacia los grandes cambios en curso en nuestro mundo tecnológico, que acontecen en sordina o encubiertos por el humo mediático, sin haber sido para nada asimilados por el gran público.

La postura más común oscila entre la maravilla encantada y la frustración frente a los continuos, incomprensibles “milagros de la tecnología”; se llega muy a menudo a la adoración mística, como si lo digital recubriera el mundo de una aura exotérica, comprensible exclusivamente a los pocos iniciados, conjugada con la frustración por la propia incapacidad de oficiar el culto al nuevo progreso de forma adecuada.

El grupo de investigadores de Ippolita se une precisamente alrededor de la convicción de que mediante el intercambio y el diálogo entre competencias y lenguajes diferentes sea posible convertir la llamada Revolución Digital en una materia útil para comprender la contemporaneidad, sus anomalías y probablemente incluso el tiempo a venir. La investigación científica, la tradición humanista, las pasiones políticas, el método feminista son otros lenguajes útiles en esta exploración.

La actividad de Ippolita revela que “poner en común” no es suficiente, porque el nivel de reflexión sobre las tecnologías es todavía limitado y la caja de herramientas de los usuarios es todavía demasiado rudimentaria. Es preciso asumir una actitud crítica y curiosa, desarrollar competencias sobre todo a nivel individual, comprender en qué formas se puede interactuar con los mundos digitales, poner a punto instrumentos adecuados a los objetivos personales. El desafío consiste en multiplicar los espacios y las ocasiones de autonomía sin ceder ante entusiasmos fáciles, sin sucumbir a la paranoia del control. *Just for fun*. La práctica comunitaria no es una receta capaz de convertir por arte de magia toda novedad tecnológica en un bien colectivo, no basta con conjurar el dominio tecnocrático en nombre de una gran democracia

electrónica. Se trata de una visión que confía ciegamente en el progreso, que olvida el valor de las elecciones individuales. La sinergia entre los sujetos en las redes, mundos vivos y constantemente cambiantes, no es una banal suma de las partes en juego, requiere pasión, confianza, creatividad y una continua renegociación de los instrumentos, de los métodos y de los objetivos.

Vincular los elementos más estrictamente técnicos a sus correspondientes efectos sociales es sin duda el primer y el más arduo paso a acometer. Por esta razón, el texto que tenéis entre vuestras manos se puede descargar integralmente gracias a una licencia copyleft.

I. Historia de motores

Motores y búsquedas.

Los motores de búsqueda hoy en día se presentan como sitios pertenecientes al World Wide Web (web) que permiten hallar informaciones (*information retrieval*). La mayoría de los usuarios cree erróneamente que Internet y la web coinciden, porque la web es el aspecto más sencillo e inmediato de Internet. En realidad la Red es mucho más compleja, heterogénea y variada de la web: incluye también los mundos de los chat, de los newsgroup, del correo electrónico, y en general todo lo que los individuos quieren “poner en Red”, prescindiendo de la “forma” en la que estas informaciones se presentan. Además, la Red no es estática, sino dinámica. Esto significa que las conexiones entre los recursos se modifican, están sujetos a cambios (nacimiento, transformación, muerte). Del mismo modo, los vectores de conexión física que llevan a estos recursos cambian continuamente, evolucionan desde la toma telefónica del modem hasta la “banda ancha” y la fibra óptica. Finalmente, los individuos que crean la Red proyectando su alter ego digitales en ella están cambiando continuamente, en tanto que seres vivos. La Red, pues, no es la web, sino una dinámica co-evolutiva constituida por las interacciones complejas entre tres distintos tipos de máquinas: máquinas mecánicas (ordenadores, cables, modem, etc.), máquinas biológicas (individuos humanos), máquinas significantes (recursos compartidos).

De todas formas, ya que conseguir información en el magma de la Red ha sido siempre una exigencia fundamental, por mucho que pueda parecer extraño, la historia de los motores de búsqueda es mucho más antigua que la de la Web.

La web, tal y como la conocemos hoy es un invento de Tim Bernes-lee, Robert Caillau¹ y otros científicos europeos y americanos. Entre 1989 y 1991 en el CERN de Ginebra se crean los primeros browser, el protocolo http y el lenguaje html para la composición y visualización de documentos hipertextuales, es decir documentos que contienen links (ya sean internos a cada documento ya sean entre documentos distintos). Esta nueva tecnología se presenta como una implementación de la Red de Internet, un proyecto americano resultante de la fusión de distintos proyectos universitarios y militares.

Mientras la web se difundía entre los laboratorios y las universidades de todo el mundo, los motores de búsqueda ya ofrecían, desde unos años, servicios de indexación y de búsqueda de informaciones en Internet.

Los primeros motores de búsqueda obviamente no eran consultables vía web: eran verdaderos programas instalables y configurables en el propio ordenador. Tales instrumentos indexaban los recursos de Internet mediante protocolos como ftp (para el intercambio de file), Gopher (un rival del emergente http) y otros sistemas ya caídos en desuso.

En 1994 es operativo el primer motor de búsqueda basado en la web: WebCrawler², un experimento financiado por la universidad de Washington. Las innovaciones introducidas por este motor de búsqueda eran extraordinarias: además de funcionar como un sitio Internet, y de

¹ Véase Tim Berners-Lee *Tejiendo la Red, el inventor del Worl Wide Web nos descubre su origen*. Siglo XXI,2000- y [L'architettura del nuovo Web](#)

² Todavía existe, y es un óptimo meta-motor de búsqueda que combina los resultados de los mejores motores: <http://www.webcrawler.com>

ofrecer la posibilidad de realizar búsquedas *fulltext*³, disponía de un instrumento para catalogar las páginas de forma automática, el *spider*. El *spider*, es un software que, como veremos en profundidad más adelante, desarrolla dos funciones: mientras navega en Internet, memoriza las informaciones contenidas en las páginas web y las pone a disposición de los usuarios del servicio de búsqueda.

Aunque WebCrawler fuese increíblemente revolucionario para la época, los resultados que era capaz de ofrecer eran simples elencos de direcciones web y los respectivos títulos de las páginas indexadas.

En los últimos meses de 1994 el motor de búsqueda Lycos conseguía indexar en muy pocos meses más del 90% de las páginas de aquel entonces (unos 10 millones aproximadamente). La principal innovación de Lycos era la de excluir sistemas *fulltext* para la catalogación, analizando solo las primeras 20 líneas de toda página indexada, en lugar de examinar el contenido integral del texto de cada página. Esto le permitía a Lycos mostrar en los resultados incluso una pequeña descripción extrapolada de las mismas líneas.

Con la llegada de Excite, en diciembre de 1995, los resultados empiezan a ser ordenados en base al criterio de importancia de las páginas. La introducción de un sistema de “peso” de una página posibilitaba también una forma rudimentaria de catalogación temática: por fin se apartaban los interminables listados de resultados sin ordenar. Se configuraba así, una primera consulta a la “directory” de los sitios, algo comparables a una organización estilo biblioteconómico clásico de los recursos de la web, o sea en forma de árbol, con una indexación dividida por materias, lenguas, etc.

Sin embargo, Excite ha pasado a la historia por otra razón: ha sido el primer motor de búsqueda que se ha equipado de instrumentos finalizados explícitamente a la actividad comercial. Después de haber adquirido WebCrawler, ha ofrecido a los usuarios personalizaciones y cuentas de correo gratuitas, convirtiéndose en menos de dos años en uno de los primeros portales web (1997). En poco tiempo, con el nacimiento de motores de búsqueda cada vez más evolucionados, Excite se aparta de su vocación originaria, optando por el uso de instrumentos de búsqueda de otras sociedades, entre las que hoy figura también Google.⁴

El panorama de los precursores de Google no sería completo sin una mención a lo que en 1997 era el mejor motor de búsqueda existente, y también el más popular: AltaVista. AltaVista, que quiere decir “vista desde lo alto”, se valía de los resultados de un grupo de investigación de la Digital Equipment Corporation (DEC) de Palo Alto, California, que en la primavera de 1995 había conseguido almacenar todas las palabras de una página HTML presente en Internet en un índice en el que era posible realizar búsquedas. Esto le había permitido el desarrollo de la primera base de datos abierta a la consulta directa del World Wide Web. La mente de AltaVista era Louis Monier⁵ que, usando *cluster* de ordenadores conectados entre sí, hardware a la vanguardia y lo mejor de las tecnologías entonces disponibles, había conseguido que su creación fuese el motor de búsqueda más utilizado y apreciado por los usuarios. AltaVista ha sido también

³ Gracias a la búsqueda *full-text* es posible encontrar palabras, combinaciones de palabras, frases completas, buscando en el interior de todo el documento.

⁴ La idea de usar las funcionalidades de búsqueda de otros sitios, y por lo tanto de poder ofrecer entre los resultados propios también los ajenos, no ha sido de Excite. Ya a finales de 1995, MetaCrawler ofrecía en sus páginas los servicios de todos los motores de búsqueda presentes. Hoy en día existen muchísimos meta-motores que cumplen búsquedas cruzadas sobre distintas tipologías de base de datos.

⁵ Después, Monier ha trabajado por Ebay; en 2005 ha sido contratado por Google.

el primer motor de búsqueda multilingüe en Internet y la primera tecnología de búsqueda capaz de soportar textos en chino, japonés y coreano. Finalmente, ha introducido un sistema de traducción automática Babel Fish, aún muy utilizado.

Antes de su fracaso⁶ en 1997, AltaVista daba servicio a 25 millones de búsquedas diarias y recogía 50 millones de dólares anuales de los patrocinios, ofreciendo su servicio de búsqueda a los usuarios del portal Yahoo!, que sigue siendo todavía el mayor antagonista de Google en el mundo Web.

El nacimiento de Goolge: en principio fue el garaje, más bien, la universidad.

La palabra Google procede de “Googol”, término matemático que indica un 1 seguido por 100 ceros. La leyenda quiere que éste fuese el número de páginas web que Larry Page y Sergey Brin soñaban con indexar con su nuevo instrumento de búsqueda. Se conocieron en 1995 en Stanford, cuando Larry Page (tenía entonces 24 años), licenciado en la Universidad de Michigan, visita Stanford con la idea de matricularse en el doctorado de informática. Sergey Brin (entonces 23 años) formaba parte del grupo de estudiantes que tenían que acompañar a los visitantes por el campus. Stanford era el lugar ideal para desarrollar proyectos altamente innovadores desde el punto de vista tecnológico. De hecho podía ofrecer por un lado laboratorios de investigación a la vanguardia; por otro, estrechas relaciones tanto con empresas del sector Information Technology (IT) como con personas atentas y dispuestas a invertir grandes capitales en las investigaciones académicas más prometedoras. Brin y Page están fascinados por el crecimiento vertiginoso de la web y por los problemas relativos a la búsqueda y a la organización de las informaciones; se dedican al proyecto BackRub, así llamado porque era capaz de analizar los “back links” que remitían a un sitio y de ofrecer un mapa de la situación. Después BackRub será re-bautizado como Google cuando en 1997 dispondrá de una página web.

La innovación fundamental introducida por Google es la de dar un vuelco al proceso de indexación de las páginas: ya no se enumeran los sitios en base a la “proximidad” con respecto a la pregunta realizada, sino que se muestran los resultados exactamente en el orden “correcto”, es decir los que responden a las expectativas del usuario. El primer link es el que da la respuesta “exacta” a la pregunta; los sucesivos se alejan poco a poco del centro de la pregunta⁷.

En esta óptica nace el conocido botón “hoy me siento afortunado”: clickandolo, se abre directamente el primer link de la búsqueda de Google, presentado como el link indiscutiblemente “correcto”.

El algoritmo que Larry Page “inventa” para calcular la importancia de una página web, conocido como PageRank, se basa en la matemática estadística de los primeros años del Novecientos y especialmente en las fórmulas desarrolladas por Andrej Andreevic Markov para calcular, dentro de una Red, la importancia de un nudo respecto a los demás⁸.

⁶ En 1996 el valor de esta búsqueda aún no es monetario, considerando también el estado embrionario en el que se halla la economía relacionada con Internet, pero su potencial está completamente reconocido por la DEC, sociedad que se encargaba del aspecto hardware de AltaVista. Sin embargo, el grupo DEC no será capaz de explotar la ocasión a causa de una serie de desafortunadas reestructuraciones internas que provocan una ruptura en el equipo de AltaVista, adquirido en 2003 por Overture (luego cedida a Yahoo!).

⁷ Creemos que este enfoque es absolutamente irreal, además de ser políticamente incorrecto e intrínsecamente autoritario, ya que supone la existencia de una autoridad capaz de individualizar y ofrecer una respuesta “correcta, rápida y justa” a un deseo. Pueden existir muchos recorridos individuales para construir una respuesta aceptable y satisfactoria a una necesidad/deseo (véase Cap. V).

⁸ Véase Cap. IV.

En un primer momento, Google es únicamente un proyecto académico y científico, para el que el sistema de evaluación depende en gran medida de los reseñadores (*referees*) que operan según el mecanismo de la “revisión de los pares” (*peer-review*). El método que teóricamente ofrece las mayores garantías de objetividad es el llamado “doble ciego” (*double-blind*): antes de ser publicado en una revista científica, un artículo se somete a por lo menos dos investigadores de renombre, que no deben conocer el nombre del autor (para no ser influenciados en el juicio). El segundo paso “ciego” tiene que ver con el redactor del artículo, que no debe conocer el nombre de los reseñadores. En suma, cuanto más positiva es la acogida que un artículo recibe por otros científicos (que se supone trabajan de forma independiente), tanto más el mismo artículo se considera importante y meritorio. Page traslada en su campo de búsqueda este enfoque y desarrolla la teoría según la que el número de link que conduce a una página es una forma de evaluar el valor de la página misma y, en un cierto sentido, su calidad. Más adelante veremos mejor cómo se realiza este paso de la “cantidad” de informaciones suministradas a la “calidad” del resultado que da respuestas a las expectativas del usuario⁹. Sin embargo, este criterio no era de por sí suficiente, porque no todos los links son iguales y tampoco tienen el mismo valor: más exactamente, desplazando la atención del valor estático de una conexión al valor dinámico de un recorrido, la Red es un ambiente (en términos matemáticos un *grafo*) en el que no todos los recorridos tienen el mismo valor, sino distintos “valores de rutas” según el “peso” de los nudos. Por otra parte, siguiendo con el paralelo con la práctica científica, no todas las reseñas en las revistas científicas tienen la misma importancia. La opinión positiva de reseñadores no prestigiosos, o incluso mal vistos por la comunidad científica, es nociva: demasiadas reseñas de escaso peso invalidan la autoridad de una publicación. Así, según Page, los links que proceden de sitios especialmente conectados, con un elevado número de links, son de mayor importancia. Por lo tanto, un recorrido (link) que lleva a una página desde un sitio muy popular tiene mayor relevancia que un recorrido procedente de una página poco conocida. De esta manera, un link desde una página A a una página B se interpreta como una reseña científica (aunque hay que subrayar que Brin y Page hablan explícitamente de “voto” y “democracia”), de mayor peso dependiendo de la popularidad del reseñador/votante que facilita el link. El criterio de la autoridad del reseñador científico se convierte banalmente en la medida de la popularidad de un sitio.

La evaluación de las páginas web de Google, llamada *PageRanking*, se construye pues en base a un sistema de reseña “pública” análogo al sistema de la “república científica”¹⁰, que no está limitado a los científicos sino que se extiende a todos los internautas de la Red.

Hoy en día, la organización de la república científica y la cuestión de la localización de las informaciones en general se ha convertido en un problema de importancia crucial: en este contexto de diluvio informacional¹¹, particularmente en la web, razonar no solamente sobre la

⁹ Véase Cap. VI

¹⁰ el sistema de las publicaciones científicas nace en 1665 con *The Philosophical Transactions*, la primera revista científica fundada por Henry Oldenburg de la Royal Society de Londres, una especie de público registro de la propiedad intelectual; funcionaba como una clase de oficina de patentes de las ideas científicas. La revista quería introducir claridad y transparencia en el proceso de fundación de las pretensiones innovadoras y tales pretensiones se veían consolidadas gracias al reconocimiento de un título de “nobleza intelectual”; esta a su vez garantizaba un título especial de propiedad: la intelectual. La importancia de las publicaciones contribuía a la creación de una “jerarquía entre pares”, basada en un ideal de excelencia: sin embargo, ya que los resultados para ser apreciados necesitaban publicidad, la organización de las reglas de publicación asumió un papel central, ya que para un científico ser visible se convirtió en algo fundamental para adquirir notoriedad y prestigio.

¹¹ El concepto de diluvio informacional (“*déluge informationnel*”) ha sido desarrollado sobretudo por Pierre Lévy, reelaborando la idea de “segundo diluvio” del artista Roy Ascott. De ahí procede un cuadro filosófico en el que el movimiento de “virtualización” asume un papel central, véase: Pierre Lévy, *¿Qué es lo virtual?* 1998, Ediciones Paidós Ibérica, S.A., (online la versión original francés: <http://hypermedia.univ-paris8.fr/pierre/virtuel/virt0.htm>).

importancia sino también sobre la credibilidad de la información es cada vez más complejo. Y esto lo es mucho más desde que el mecanismo de la “revisión de los pares” ha sido puesto en tela de juicio por los mismos científicos¹². Entre las propuestas alternativas más interesantes figuran mecanismos de **evaluación** distribuida, puesta a disposición gracias a licencias libres (copyleft) a través de proyectos de acceso público (Open Access), incluso por lo que respecta las obras humanistas (veáse por ejemplo HyperJournal)¹³ de las publicaciones.

Así en marzo de 1996 Page lanza su spider, un programa para explorar la web, testándolo con el algoritmo de ranking que había desarrollado.

El motor de búsqueda de los dos talentos de Stanford, construido alrededor del spider, se convierte inmediatamente en un éxito entre los estudiantes y los investigadores, adquiriendo rápidamente una extraordinaria popularidad. Sin embargo, para los administradores de la Red de Stanford el consumo de banda del motor de búsqueda empieza a convertirse en un problema significativo. Además, los propietarios de los sitios indexados se preocupaban por un lado de los derechos de copyright de sus propios materiales y por otro del hecho de que el ranking de Google no tomara en consideración alguna los premios oficiales o los demás sistemas de evaluación consolidados, salvo por la cantidad y calidad (en el sentido de popularidad) de las relaciones que una página es capaz de mover a su alrededor. A Google le interesa la economía relacional expresada en términos de link y nada más: “el spider no se interesa por los contenidos de una página”.

Una afirmación de este tipo supone que el valor de un resultado se construye mediante la evaluación de las relaciones de una página con otras, y no imponiendo una clasificación arbitraria en base a términos de búsqueda. Este forzamiento será la clave de la eficiencia de

De manera independiente, Manuel Castells habla de “informacionalismo” y de “economía informacional”. Gran parte de las lecturas económicas acerca de la época del diluvio informacional, especialmente en ámbito francés, italiano, estadounidense, tienen derivaciones marxistas y descansan sobre interpretaciones forzosas de los fenómenos de “inteligencia colectiva” de las Redes como hipótesis del *General Intellect* marxista; veáse por ejemplo: Olivier Blondeau, *Le Capitalisme Informatique* – Genève et subversión

http://www.vulgum.org/article.php?id_article=158; la revista monográfica “La società del General Intellect”, Derive approdi, Roma; Wark McKenzie, *Un manifesto hacker*, 2006, Ediciones Alpha Decay, S.A.

¹² Sitiografía: La revisión de los pares. Definiciones: http://es.wikipedia.org/wiki/Revisi%C3%B3n_por_pares; la topología de una Red de citas: Stefano Mazzocchi, “Peer Review vs. Citation Network Topology” <http://www.betaversion.org/~stefano/linotype/news/74/>; Artículos críticos recientes, The scientist: Alison McCook “Is Peer Review Broken?”. *The Scientist*. Volume 20, Sigue 2, page 26 <http://www.the-scientist.com/2006/2/1/26/1/>; Richard Gallagher. “Taking on peer review”. *The Scientist*. Volume 20, Sigue 2, page 13 <http://www.the-scientist.com/2006/2/1/13/1/>; Google y la revisión de los pares: <http://www.lib.uiowa.edu/hardin/md/notes4.html>; Il fenómeno dei Blog tra informazione e rumore di fondo: Fabio Giglietto,

<http://nextmedia.blogspot.com/2004/10/informazione-non-informazione-fra.html>; Citas y revisión de los pares, ciencias duras y ciencias mórbidas: Francesca di Donato: “Verso uno European Citation Index for the Humanities – Che cosa possono fare i ricercatori per la comunicazione scientifica”

<http://bfp.sp.unipi.it/~didonato/ovre/ECIH.html>; El exhaustivo trabajo de Jean-Claude Guédon, “In Oldenburg’s Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing” Université de Montreal, <http://www.arl.org/ar/ proceedings/138/guedon.html>.

¹³ Sitiografía: El acceso a los saberes; la libertad de la información y el acceso abierto a la literatura científica. Señalamos el volumen del Gruppo Laser, “Il sapere liberato”, Feltrinelli, Milano, 2005, online en www.ippolita.net/laser/. En campo humanista, el trabajo desarrollado por Francesca di Donato, que consigue conjugar un enfoque cualitativo y no **cientista** con la necesidad de descentrar los archivos y las informaciones, y definir un número mínimo de criterios a priori de catalogación (metadatos, por ejemplo XML), para intercambiar mediante un protocolo común. Veáse “HyperJournal. Un software open source per l’Open Publishing” (abstract y slides) <http://purl.org/hj/bfp/102/>; Seminario: “Dopo Berlin 3: politiche di accesso aperto alla letteratura scientifica” (Pisa, 16 febbraio 2006) <http://bfp.sp.unipi.it/blog/html/notizie/oapisa.html>. El sitio de HyperJournal: <http://www.hjournal.org/>.

Google en los siguientes años: los resultados no se establecen de manera absoluta una vez por todas, sino que varían dinámicamente según la posición que asumen las páginas en el contexto de la Red.

Google.com: publicidad directa dentro de las páginas

Page e Brin desarrollan y prueban Google durante todo el primer año y medio de vida usando instrumentos libres puestos a disposición por las comunidades Free Software y Open Source¹⁴, a partir del sistema operativo GNU/Linux. De esa manera implementan un instrumento fuertemente modular y escalable, que se puede expandir y modificar fácilmente incluso mientras está siendo usado. Esta estructura modular hoy en día es la base del *datacenter*¹⁵ de Google, el Googleplex, que permite sustituir o añadir hardware y software, realizar su manutenzione y actualización sin tener que bloquear nunca el servicio.

A mediados de 1998 Google da servicios a cerca de diez mil búsquedas diarias; el sistema casero apañado en la habitación alquilada por Page y Brin se halla al borde del colapso. Era preciso, pues, encontrar un sistema de financiación para el proyecto que fuese más allá de los normales fondos de ayuda para la investigación.

La salida de la universidad se atribuye al encuentro con Andy Bechtolsheim, fundador de la Sun y consumido *talent scout* de la IT, que entrega a los dos investigadores un primer cheque de cien mil dólares.

El nacimiento de la sociedad coincide con la contratación de los primeros dependientes necesarios al desarrollo y al mantenimiento del datacenter, entre los que Craig Silverstein, director de tecnología. El datacenter de Google, se configura desde el principio como un sistema altamente redundante en el que las informaciones se copian en más lugares para minimizar la posibilidad de pérdida de los datos (o sea, el papel moneda de un sistema de búsqueda) y sobre todo un sistema en el que en todo momento es posible añadir o quitar módulos existentes para conseguir una mayor eficacia. El enfoque típico de la filosofía de los dos manitas universitarios Brin y Page, que recuperan y readaptan hardware a buen mercado y utilizan software gratuito, se revela una carta ganadora. Los reducidos recursos económicos permiten luego convertir este sistema en la clave de volta de su filosofía empresarial, que se basará en la extrema escalabilidad a todos los niveles. El sistema-Google es *escalable* porque capaz de “crecer” o “decrecer” (aumentar o disminuir de *escala*) en función de las necesidades y de las disponibilidades. No es necesario replanear el sistema cuando se añaden nuevos recursos (de cualquier tipo: máquinas, servicios, dependientes, etc.): la estructura altamente modular consiente la adquisición de nuevos módulos sin solución de continuidad.

Google Inc. abre sus puertas el 7 de septiembre de 1998 en Menlo Park, California. Más exactamente, Larry abre estas puertas con un mando a distancia: las oficinas se hallan en el garaje que un viejo amigo subalquila a la sociedad. Una oficina-garaje espartana, pero dotada de

¹⁴ veáse el Cap. III, análisis más detallado de las relaciones entre estas fórmulas.

¹⁵ El datacenter de Google se compone actualmente de varios datacenter esparcidos por el mundo que elaboran las solicitudes realizadas y se encargan de propagar las modificaciones de las indexaciones. La actualización de las bases de datos, así como el cálculo del pageranking, se resiente de esta descentralización: los datos normalmente se alinean en unos días, o algo más; este desfase da lugar a la llamada “Google dance”, el baile de los ranking. Para una panorámica general, veáse el artículo <http://www.alverde.net/articoli/stampa.asp?ID=145> y <http://www.motoricerca.info/articoli/google-dance.phtml> sobre la “Google dance”.

algunos confort no tan indiferentes: una lavadora, una secadora y un hidromasaje. Desde el principio, la filosofía empresarial de Google se basa en la satisfacción de los trabajadores.

En enero de 1999 Google abandona definitivamente el campus de Stanford, declarando: “El proyecto de investigación Google se ha convertido en Google Inc. Queremos darle al mundo búsquedas de calidad superior y mejores de las que hay actualmente, y una compañía parece ser la mejor forma para llevar a cabo este objetivo. Hemos empezado a incorporar otras personas y a configurar otros servidores para hacer que el nuestro sea un sistema escalable (hemos empezado a ordenar los server en paquetes de 21). Hemos empezado también a lanzar nuestro spider con más frecuencia y nuestros resultados no solo siguen siendo tan rápidos como antes, sino que están aún más actualizados. Estamos incorporando personas de gran talento muy rápidamente y esto comporta disponer de las últimas y las mejores tecnologías por lo que atañe la Web”... Brin y Page dedican luego algunas líneas a las diez buenas razones para trabajar en Google, incluyendo tecnologías, acciones, *snacks*, bebidas gratuitas y la posibilidad de que millones de personas: “utilicen y aprecien vuestros software”.

Entre 1998 y 1999 todos los motores de búsqueda y los sitios más populares del mundo están afectados por el síndrome del portal, una verdadera manía en la construcción de sitios que quieren que los visitantes se queden a cualquier precio más tiempo en la propia página, para ello incluyen cada vez más servicios, publicidad y personalizaciones. Google es el único instrumento carente de publicidad y de servicios añadidos: sólo motor de búsqueda, entre otras cosas el mejor, el más rápido, y sin lazos comerciales o patrocinadores.

En todo caso la compañía no puede sobrevivir con la modesta suma de Bechtolsheim, sin generar provechos relevantes y siguiendo con la investigación sobre la organización y la localización de las informaciones. Mostrando una gran capacidad de interlocución con las altas esferas de la financia, e insistiendo sobre su interés en la investigación, Brin y Page consiguen poner de acuerdo dos de las principales sociedades de financiación de California, que sorpresivamente aceptan co-financiar una empresa particular: caso más único que raro, dos enormes sociedades de capital a riesgo (*venture capitalism*) se dejan convencer a compartir los riesgos, y los provechos, del eventual negocio. El 7 de junio de 1999 Google puede anunciar que la Sequoia Capital Sequoia Capital y la Kleiner Perkins Caufield & Byers le han concedido una financiación de 25 millones de dólares¹⁶.

Mientras van aumentando los dependientes de la Google Inc., los dos presidentes buscan una forma para explotar económicamente la gran cantidad de datos indexados. Inicialmente intentan vender su servicio de búsqueda como OEM (*Original Equipment Manufacturer* – Productos ofrecidos sin los soportes originales)¹⁷ a portales, pero los resultados son más bien escasos. El modelo de business que se revela más adecuado a la nueva sociedad es la publicidad directa dentro del motor de búsqueda, mediante el cálculo del número de visitantes que acceden a los sitios desde los links promocionales y de pago. Este modelo económico, CPM (Coste por Mil Impresiones –en inglés CPT, Cost per Thousand Impresión)¹⁸ se estructura de manera tal que parece el menos indiscreto posible para el usuario; de hecho, en lugar de mirar hacia la extrema

¹⁶ Ambas sociedades detienen aún hoy las cuotas mayoritarias del capital de Google.

¹⁷ Una OEM es una sociedad que construye componentes utilizados en sistemas vendidos por otra sociedad, por ejemplo, un revendedor. Sociedad que venden productos con componentes que proceden de OEM son comunes en la industria electrónica e informática. Típicamente la OEM se contrapone al RETAIL, o sea el producto se ofrece a terceros sin los soportes originales (CD de instalación, manuales, etc.), integrado en el producto por el revendedor.

¹⁸ CPM, CPC, etc.: los modelos de business publicitario online son cada vez más complejos y refinados. Para una rápida introducción véase http://earlstagevc.typepad.com/earlstagevc/2005/08/web_20_needs_bu.html.

visibilidad de los anuncios publicitarios, se muestran solo unos pocos links atentamente seleccionados en la parte de arriba de la búsqueda. Estos links tienen colores y caracteres distintos de los que no son comerciales, pero intentan no molestar de ninguna forma la actividad de búsqueda del usuario.

El self-service publicitario: más allá de la burbuja de la new-economy

Un modelo económico basado en la aparición de simples links a los que patrocinar, colocados al lado de las búsquedas, es poco ventajoso en términos de provecho; en este punto la planificación empresarial a largo plazo de Google necesita de un salto de calidad. Los presidentes miran a su alrededor en busca de soluciones más interesantes y proficuas, y tropiezan en Goto, sociedad fundada por Bill Gross¹⁹, hoy de propiedad de Overture/Yahoo!.

El negocio de Goto consistía en mezclar resultados efectivos y resultados patrocinados, facturando a los que ponían inserciones solamente los clic realizados por los usuarios en sus direcciones: un modelo conocido como CPC (*Cost Per Clic*).

Este método era especialmente innovador con respecto a los anteriores. Los links patrocinados aparecían en base a las búsquedas de los usuarios, ampliando de esa forma al máximo las posibilidades de llevar a buen término la transacción económica, contabilizada como clic hacia el sitio pagador. Google intenta arrancar una partnership con Goto, pero el decidido rechazo del presidente les obliga a implementar internamente un instrumento parecido. En esa época todos los portales (por ejemplo Excite, Lycos, Infoseek, AltaVista y Yahoo!) utilizaban el modelo CPM, y el CPC era casi un deseo prohibido. Esto demuestra como un intento fracasado de adquirir la tecnología que se cree mejor para el propio negocio conlleva la necesidad de implementarla de forma autónoma.

En marzo de 2000 estalla la burbuja de Nasdaq, arrastrando consigo todos los sueños de gloria de las dot.com. También se viene abajo el modelo CPM, es decir la ilusión de que toneladas de banner publicitarios a “miles de impresiones”, muy a menudo descontextualizados porque presentes en sitios totalmente ajenos al negocio de referencia del anunciante, fueran a ser una fuente inagotable de provechos. Google tiene entonces que enfrentarse con la necesidad de monetizar sus propias tecnología de búsqueda.

En octubre de 2000 entra en escena AdWords, una especie de self-service publicitario en el que los anunciantes pueden elegir las palabras de búsqueda a las que asociar su propio link comercial: una implementación de la “publicidad por palabras claves” (*Keyweods advertising*) ideada por Goto.

Google sobrevive por lo tanto al fracaso de Nasdaq: más bien, gracias a su naturaleza de sociedad privada aún no cotizada en bolsa, aprovecha para contratar todos los talentos en fuga de las ruinas de las demás dot.com. A mediados de 2000 Google da servicios a 18 millones de búsqueda diarias y su índice de documentos cuenta con un billón de unidades: seis meses después las búsquedas son 60 millones.

¹⁹ La historia de Goto es muy emblemática, porque muestra como incluso en el panorama de la web cada vez más se va hacia las mega fusiones típicas del capitalismo informacional. Un buen resumen en italiano hasta el 2004 se halla en <http://motoridiricerca.it/goto.htm>; en wikipedia, recursos sobre el papel comercial jugado por Goto/Overture en el business de Yahoo! http://en.wikipedia.org/wiki/Yahoo!_Search_Marketing.

Estilo, morfología y multiplicación de los servicios

A principios del 2000 Google tiene muy pocos competidores y da comienzo a una nueva fase de innovación, empezando a ofrecer una serie de nuevos servicios a los usuarios²⁰.

Cada nuevo servicio forma parte de un mosaico complejo, constantemente redefinido, con ramificaciones por todos los ámbitos del IT. Actualmente Google ofrece: diecisiete tipologías de servicios de búsqueda en base de datos de imágenes, blog, noticias, libros, mapas, videos, servicios financieros, servicios de búsqueda en los documentos presentes en el propio ordenador y mucho más. Dos servicios están relacionados con el desarrollo de aplicaciones y con las novedades sobre las que se está trabajando en los laboratorios de Google. Seis los servicios de comunicaciones: Email, VoIP, mensajería instantánea, Grupos de debate, fotos compartidas, servicios de traducción. Tres de los servicios tienen que ver con dispositivos móviles (móviles, Pda, ecc.); para terminar, un servicio propone software aconsejados por Google. No hay día que no se añadan nuevos servicios.

Incluso el usuario más desprovisto se da cuenta fácilmente del poder de estos instrumentos. Hoy es posible digitar una dirección o un número de teléfono y Google ofrece todos los datos necesarios para contactar y localizar una persona o una cosa; se pueden memorizar las preferencias personales de búsqueda de modo tal que la búsqueda sea de una facilidad desconcertante; si hay un error en la digitación de la palabra en la búsqueda, Google sugiere la búsqueda correcta con un mecanismo de corrección ortográfica muy avanzado capaz de “aprender” de cada búsqueda.

En 2001 Google lanza “Google imágenes”, un motor de búsqueda de imágenes que en pocas semanas se convierte en una de las principales fuentes de producción gráfica del do it yourself y representa una de las mayores base de datos de imágenes disponible en la web. Mientras tanto la compañía compra a Deja.com el archivo Usenet, es decir los mensajes de los newsgroups desde 1995, que con más de 650 millones de post es la “memoria histórica” de Internet antes de la web, cuando los grupos de debate eran la linfa vital de la Red. En abril de 2001 Usenet se convierte en “Google grupos”, con una interfaz bastante simple y cautivadora para seguir con facilidad complejos temas de debate. Todos estos servicios están en continua ampliación y mejora.

Desde 2001 en adelante los servicios nacen a repetición, sin ningún objetivo evidente ni contracambio económico inmediato, como si Brin y Page se divirtieran a mostrar que un datacenter ilimitado es capaz de elaborar cualquier sueño tecnológico. El ejemplo más desconcertante es el poderosísimo Google Maps (Tierra, Marte 2005, y Luna 2006), un set de software disponibles gratuitamente para visualizar con imágenes de satélite un mapa detallado o cuando menos una reconstrucción fotográfica de la superficie de los planetas. “Google Directory”, que presenta los contenidos Open Directory de Dmoz.com, organizados por agentes humanos con un sistema cooperativo y descentrado de open publishing, y desde hace tiempo consultable en la cada vez menos lacónica página de Google.

En 2005 aparecen Goolge News, la enorme base de datos de informaciones de Google que se pone así al servicio de la producción periodística, y Gmail que ofrece 1 Gigabyte de espacio mail a cada usuario. Con un mecanismo de invitaciones que crea inmediatamente una Red de relaciones toda interna al mundo de Google, se silencian las protestas acerca de la privacy de la

²⁰ Aquí aludiremos solo a los servicios más interesantes; para un listado exhaustivo de los servicios ofrecidos por Goolge, veáse <http://www.google.com/intl/en/options/> (los servicios disponibles en italiano y en otros idiomas son menos, pero en rápido aumento).

propia cuenta electrónica con una frase lapidaria “se trata de un buen producto, las ventajas son mayores de las dudas y el tiempo lo demostrará”. Nadie de la competencia tiene la capacidad de ofrecer gratuitamente tanto espacio. Con Gmail cada actor implicado está controlado desde el punto de vista del uso del servicio, ya que la enorme disponibilidad de espacio invita a dejar los propios correos en los servidores de Google. Además, puesto que la difusión del servicio se da mediante invitaciones, que cada usuario puede usar a su gusto, a Google se le facilitan una serie de informaciones esenciales sobre los conocidos y los amigos: un proyecto de invasión y enfocado al *data mining*.

Después llega el proyecto Google Scholar (Academia), un motor de búsqueda todavía experimental en fase beta (la llamada fase de *beta-testing*) que permite localizar textos de la literatura académica así como artículos a revisar, tesinas, tesis doctorales, volúmenes académicos, *preprint*, sumarios, relaciones técnicas, reseñas de todos los sectores de la investigación académica. Finalmente, Google Library, que ambiciona almacenar on line todos los libros del mundo en formato digital, sellando acuerdos con bibliotecas de todo el mundo para la digitalización de los volúmenes, pero también con las editoriales que estén interesadas. Un sueño que solamente el datacenter de Google promete poder realizar: la biblioteca digital global accesible desde las páginas de Google, en abierta ruptura con la mayoría de las editoriales estadounidenses, reunidas en la Association of American Publishers, que temen una drástica caída en los provechos; ya en 2005 la AAP había solicitado el bloqueo durante seis meses de la digitalización de las obras protegidas por copyright a la espera de explicaciones convincentes sobre la Google Library. Pero la apariencia engaña y, no obstante la oposición de los fautores del copyright, no se trata de una iniciativa para la libre circulación de los saberes, sino sólo de un cambio de monopolio de la información, que pasaría de unas grandes editoriales a Google. De hecho, como para todo sueño, la pesadilla está al acecho tras la esquina: un sujeto privado, Google, ambiciona con decidir qué va a tomar parte de la esfera de la información colectiva, poniéndolo a disposición mediante formatos propietarios. En reacción a este proyecto ha surgido la Open Content Alliance, sostenida por Internet Archive, una fundación no profit, y por Yahoo!: el objetivo es la disponibilidad total del material adquirido mediante el uso de formatos abiertos.

Paralelamente al nacimiento de nuevos servicios, Google, se revela extremadamente hábil en aprovecharse plenamente de la economía relacional, generada por la explotación en sentido comercial de los datos indexados.

En 2004 se lanza AdSense, es decir la posibilidad para cualquiera de hospedar en su propio sitio algunos links comerciales sugeridos por Google basados en los temas tratados y en algunas palabras claves. Los introitos procedentes de los clicks en estos links, se reparten entre Google y los propietarios de los sitios hospedantes. La novedad de este enfoque consiste en explotar la Red de confianza basada en los usuarios del sitio: Google ya no está en el sitio de Google, sino que se halla en cualquier sitio que quiera hospedar su “ventana”, y este pequeño espacio ofrecerá siempre datos validos e interesantes, como pasa siempre con Google, aunque ahora se tratará de sugerencias comerciales. En la práctica AdSense sustancia una “Red de Google”, una Red que cruza los datos de los usuarios y las relaciones entre ellos para las sugerencias publicitarias. En palabras de Google, es una Red que “está formada por un conjunto amplio de sitios web y otros productos, como programas de correo electrónico y blogs, que se han asociado a Google para publicar anuncios de AdWords.²¹; ¡naturalmente el sistema AdWord también forma parte de la “Red de Google”!

²¹ Definición de la Red de Google según Google:

<http://adwords.google.es/support/bin/answer.py?answer=6104&ctx=sibling>

Desde luego, una vez creada la Red, es preciso sacarle dinero. No por puro azar, siempre en el 2005, Google experimenta un “regreso” al modelo CPM a través de la oferta AdWords con la metodología “targeting por sitio”: los anunciantes vuelven a pagar por visualizaciones (*impresión*) con un mecanismo de subasta y no por número de clicks en el banner. Pueden elegir en modo extraordinariamente detallado qué características tienen que reunir los usuarios que visualizarán el anuncio: nacionalidad, área geográfica de procedencia, idioma, etc. Pero, sobre todo, las visualizaciones sólo se dan dentro de la “Red de Google”. Esta estrategia les interesa sobre todo a quienes desean visibilidad y promoción de una firma, más que llegar a usuarios a los que vender un producto, o sea privilegian una forma indirecta de marketing, frente a formas directas. Más en detalle, se trata de invertir en la llamada “concienciación de la firma” (*brand awareness*) en lugar que productos específicos indicados por palabras claves (como por ej. el *keyword advertising* del modelo CPC).

Hay una sórdida especularidad entre el modo de gestionar los propios productos inmateriales y la organización de los dependientes, del ambiente de desarrollo proyectual, perfectamente adecuado al criterio de escalabilidad modular sobre el que se fundamenta la filosofía empresarial de Google: el crecimiento infinito es preciso para que el sistema no estalle. El número de usuarios que realizan búsquedas en Google y confían a sus servidores sus propios datos tiene que crecer continuamente, para que crezca a la par que el número de nuevos servicios, nuevas máquinas para gestionarlos, nuevos dependientes para implementarlos, nuevos usuarios para utilizarlos, nuevos anunciantes para implementar los provechos.... Cada nueva “pieza” del sistema será introducida como nuevo módulo, en un ciclo de acumulación sin fin: acumulación de datos y acumulación de cerebros y acumulación de usuarios y de sus datos, bondad en el tratamiento de los datos, bondad en el tratamiento de los dependientes, bondad en el tratamiento de los usuarios y de sus datos, archivados en el *datacenter* de Google. Siempre con el imperativo de la rapidez y del desarrollo.

Brin y Page no ocultan sus ambiciones: “¿Para qué dejar que nuestros dependientes intenten fundar su propia sociedad para luego tener que comprarla si en lugar de eso podemos financiarlos para que se queden con nosotros haciendo lo que habrían hecho de todas formas?”. De hecho el GooglePlex²², la sede operativa para los trabajadores de Google en Mountain View, es una especie de campus universitario en el que las personas reciben constantes mimos y a los dependientes se les deja incluso un día a la semana para que se dediquen a trabajar en sus propios proyectos; estos proyectos se presentan luego al “Google Duo” que premia con dinero y apoyo empresarial a los talentos más prometedores.

Google, el gigante bueno, cotiza en Bolsa

“Don’t be evil”, no seas malo, puedes hacerlo todo, basta que no hagas nunca el mal. Este es lema del capitalismo humano de Google²³. A pesar de estas profesiones de buena fe, ya se van registrando muchas excepciones en esta campaña a favor del Bien: juicios de varia naturaleza, sospechas de fraude, ocultación de sitios Web, etc²⁴....

²² véase el servicio fotográfico del Time sobre Googleplex:
http://www.time.com/time/photoessays/2006/inside_google/

²³ La filosofía de Google Inc.: <http://investor.google.com/conduct.html>

²⁴ Tres ejemplos de la cantidad de informaciones al respecto: un artículo de WiRed, Google vs. Evil, www.wiRed.com/wiRed/archive/11.01/google_pr.html; un simpático cuento de Paul Ford, The banality of Google (en lugar de la notoria banalidad del mal...), <http://ftrain.com/GoogleIP.html>, y un portal entero sobre el refrán, extendido a un auspiciado retorno de la confianza en los negocios, en la política, en los medios, con tanto de manifestaciones para que Google se quede fiel a sus principios, <http://www.dontbeevil.com/>

En diciembre de 2002 Google cuenta con más de 1000 dependientes y posee más de 10.000 ordenadores. Indexa más de cuarenta mil millones de documentos y las ganancias netas (aunque reveladas con mucha cautela) rondan los 185 millones de dólares. Frente a tal tamaño los inversores empiezan a pretender mayor claridad y control, imponiendo una estructura societaria más creíble, en cuyo vértice se sitúan dos geniales y excéntricos ingenieros, pero también un manager de comprobadas capacidades empresariales. Después de una serie de contactos infructuosos y desencuentros verbales, se asigna el rol de Administrador Delegado (CEO, Chief, Executive Officer) de Google a Eric Schmidt (antes en la plantilla de Sun, luego de Novell). Los dos jóvenes fenómenos continúan tomando decisiones a todo campo, aunque la acción diplomática se revela una buena elección económica, puesto que la llegada de Schmidt coincide con el primer trimestre en activo de la compañía, la cual demuestra que por fin ha cruzado el umbral de la monetarización de sus productos.

Page y Brin habían retrasado al máximo la necesidad de cuotarse en Bolsa, temiendo que la obligación de dar cuenta pública de sus perspectivas y ganancias, haría su trabajo mucho más difícil, poniendo entre otras cosas Google en la condición de ser más transparente, y por lo tanto vulnerable a sus competidores en el mercado.

Pero ahora, después de AdSense, en 2004, a pesar de que en palabras de Page “Google no es una compañía convencional. Y no queremos serlo de ninguna de las maneras”, por razones legales el nuevo coloso se vuelve exactamente como todas las otras sociedades americanas y cotiza en Bolsa.

Precisamente tras la oferta pública de acciones, Yahoo! y otros competidores emprenden decenas y decenas de juicios por violación de patente, de copyright y mucho más contra Google, para dañar al nombre en vista de la venta pública de las acciones.

Wall Street está a punto de bajar el precio mínimo previsto para las acciones, considerada la dificultad de la operación, pero Page y Brin resuelven la contienda principal con Yahoo!, pagando al coloso de Filo y Yang una compensación en acciones y arreglando las controversias de patente. Los dos, a pesar del parecer contrario de la Bolsa, siguen con la oferta pública de las acciones, justo en la mitad de agosto de 2004, aunque con una rebaja del precio de más de 20 dólares por acción.

En un solo día de contrataciones en el mercado las acciones de Google pasan desde la base de 85 dólares a 100 dólares, con una ganancia neta de un 1000 millones de dólares. Un año después las acciones alcanzan el valor de 400 dólares, con una subida de más del 300 por ciento. Para Google Inc. todo parece marchar viento en popa, en el maravilloso mundo donde nadie es malo, todos ganan, nadie se hace daño nunca. Por supuesto, frente a tales cifras, la más mínima oscilación hacia abajo quiere decir millones de dólares en humo, tal y como se vio en marzo de 2006 con una inflexión de siete puntos. Google Inc. ya es un gigante entre los gigantes de la Bolsa mundial, y si tiembla, muchos tiemblan junto con él.

Google Inc: el monopolio de la investigación

En el otoño de 2004 Page y Brin, de viaje en su charter privado, se enteran de que el coloso AOL (America Online, el mayor provider estadounidense) ha firmado un contrato para su propio motor de búsqueda con Yahoo!. Los recién treintañeros cambian de rumbo, aterrizan en Londres y convencen AOL a romper el contrato que acaban de firmar y pasarse a Google con una oferta irresistible de 50 millones de dólares. No se trata exactamente del comportamiento abierto y amable que se podría esperar del gigante bueno intentando derrotar los malos monopolistas, pero sea como sea, los negocios son los negocios, también para los dos valiosos investigadores de Mountain View.

Mientras tanto las ganancias de Google Inc. en un periodo de cinco años han crecido de más del 400 mil por ciento, poniéndose en competencia directa con Yahoo! y Microsoft, no solo en

términos de movilización financiera, sino sobre todo por la popularidad y en consecuencia por el dominio cultural sobre el imaginario. Millones de usuarios de la Red utilizan las páginas de Google como punto de partida para sus navegaciones, confían en los resultados obtenidos con las herramientas perfeccionadas en Mountain View. Google ya es sinónimo de Web, cuando no incluso de Internet. Google puede aprovechar al máximo la Red relacional de sus usuarios y consigue exprimir cada céntimo de millones de pequeños inversionistas, al punto que en 2005 los datos hablan de cerca de seis mil millones de dólares de ingresos publicitarios (Yahoo! está parado en la cuota de cuatro mil seiscientos millones de dólares).

El primer impasse al que se enfrenta Google es la relación entre sus investigaciones y la legislación acerca de las marcas depositadas (trademark, TM). Emblemáticas las denuncias interpuestas por la *Geico* y por *American Blind & Wallpaper Factory*²⁵. En los dos casos las sociedades acusan al servicio AdWords del motor de búsqueda de vender términos bajo su trademark a su competencia, violando las leyes establecidas al propósito. La cuestión es peliaguda, pues las empresas piden a Google que no aparezcan anuncios de su competencia sino expresamente autorizadas cuando los usuarios buscan términos como “geico”, “american blind” o “american wallpaper”. Si este concepto pasara a los tribunales, una de las más importantes fuentes de ingresos de Google y de sus partners peligraría enormemente, porque permitiría a cualquiera que posea un trademark vetar su uso para AdWords o denunciar a Google por su utilización. Luis Vuitton en Francia lleva a cabo un juicio similar, y gana. La respuesta oficial de Google es que los responsables de la violación, a lo sumo, son los anunciantes, y por supuesto no Google que en este caso es un médium neutro, y que “limitar la venta de términos depositados significa una limitación de la libertad de expresión”. Una posición fuerte de parte de la sociedad y aparentemente más que compartible.

En todo caso la misma sociedad de Mountain View cae en la misma trampa tendida a las sociedades sobre la libertad de expresión, violando el vínculo de confianza que une a muchos usuarios y que constituye uno de sus recursos fundamentales. Google se ha atrincherado siempre detrás de la supuesta objetividad de los algoritmos de búsqueda y de la objetividad de las máquinas que proporcionan resultados en los que no se realiza manipulación alguna. Pero, poco antes del juicio con American Blind, Google había decidido eliminar algunas Ad Words adquiridas por los activistas de Oceana 26²⁶. La “culpa” de Oceana había sido la publicación de denuncias ecologistas de la actividad de un gran inversor de Google, la Royal Caribbean Cruise Lines, utilizando como keywords en sus anuncios palabras como “cruise vacation” o “cruise ship”, utilizadas normalmente para buscar páginas reservadas a las vacaciones en barco o similares. La motivación oficial mantiene que Google es un canal neutro, y no quiere ser vehículo de campañas de propaganda desacreditadora de otras empresas. Evidentemente en este caso la libertad de expresión no era tan importante.

Además, el mismo día de la audiencia para el caso American Blind, en el distrito de San José, donde se celebraba el juicio, los resultados de Google aparecían misteriosamente diferentes de cualquier otro lugar del mundo. Por primera vez, Google manipulaba los resultados con fines distintos de aquellos del “mejor resultado de una búsqueda”. El parecer positivo del tribunal al respecto del caso Geico (en todo parecido a aquello Con American Blind) no cancela esta realidad.

El caso más reciente y conocido es el de la entrada en China del coloso de la búsqueda. Para conquistar el creciente mercado chino, Google se ha sometido por primera vez a las peticiones de censura, haciendo inaccesibles a los usuarios que efectúan búsquedas desde el territorio chino

²⁵ De momento solo Geico ha perdido, en los demás casos Google ha sido condenado a una multa. Anotamos aquí algunas serias fuentes en italiano acerca de estas vicisitudes judiciales: el caso American Blind and Wallpaper Factory: <http://punto-informatico.it/p.asp?i=46751>; Google condenado a refundir a Luis Vuitton en Francia: <http://webnews.html.it/news/2514.htm>

²⁶ El grupo no-profit Oceana lucha para la calidad de las aguas: www.oceana.org

sitios señalados por las autoridades. Aunque ya en 2002 un estudio de Harvard había revelado que Google censuraba 113 sitios, en las versiones locales francesa y alemana (Google.fr y Google.de). Google contestó que era verdad, pero que habían sido borrados bajo indicación de entes gubernamentales, autoridades policiales y otros sujetos, y tan solo después de un atento análisis de sus contenidos: muchos sitios eran de índole racista, otros de religiones fundamentalistas. Alguien polemizó sobre el hecho que la tan publicitada transparencia de Google empezaba a desmoronarse: había que notificar a los usuarios la aplicación de una censura que podía quedarse “escondida”; otros anotaron que la culpa no era de Google sino de los ordenamientos jurídicos que contemplan la posibilidad de llevar a juicio una empresa por la presencia de tan solo un link en sus paginas²⁷. En aquel caso, es evidente que Google, para evitar posibles consecuencias legales, optó por remover las posibles consecuencia legales, decidió remover los links evaluando caso por caso. Hay que apuntar que el fin del derecho de link está destinado a pesar sobre las libertades digitales en una medida cada vez mayor: ¿Quién decidirá qué es lícito censurar? ¿Una súper-autoridad más, a lo mejor sobre-estatal? ¿El derecho del más fuerte, o sea en el sistema del mercado el derecho del que paga (o vale) más? ¿O los chantajes de los fundamentalismos locales, sobre todo de matriz religiosa, que amenazan revanchas cada vez que se condena su particular visión del mundo por algún “sitio subversivo”? El alcance del problema es tan amplio como la importancia de la libertad de expresión y seguramente no se puede solucionar en los tribunales. El ruido levantado por el caso chino sin embargo, está ligado a una censura ordenada por un gobierno: no obstante esto, Page y Brin no han cambiado de idea, conscientes de la importancia de un mercado dirigido a la cuarta parte de la población mundial. Según Google pronto el mundo será un enorme índice en el cual los recursos digitales se corresponderán del todo y perfectamente con la definición de la realidad circunstante; cada índice podrá ser recogido y ofrecido de la forma mejor a través de un algoritmo en los resultados de una búsqueda: y Google está en la posición perfecta para ser la herramienta de gestión de este índice.

Apartando la obvia y banal objeción que los mundos reales y los digitales, por cuanto entrelazados, no se corresponden para nada, desde un punto de vista técnico no existe el algoritmo perfecto cuando se habla de encontrar las informaciones en la Red, sobre todo, nada es neutral en el ámbito tecnológico, menos aún cuando se gestionan los datos de individuos reales en Internet.

Independientemente de las alianzas que se estrecharán, la convergencia tecnológica que se está realizando nos parece una dirección nueva de la visión de Google como único punto de acceso, gestión y mediación de los datos digitales. Se perfila la distopia de Google como aspirante Gran Hermano, peligrosa y fascinante como todo choque epocal: la Web es el nuevo territorio de competición salvaje hacia la afirmación de un nuevo estándar de comunicación. Un estándar paradójicamente “personalizado” con servicios y propuestas diversificadas según los gustos de los usuarios: ya desde algunos años la palabra clave es “personalización de masa”. Es cierto, un oxímoron, que, sin embargo, revela la importancia de la apuesta en juego, un verdadero deslizamiento de paradigma, desde el consumismo serial de la industria de masa al consumismo personalizado, que se vende como “libertad de elección”. Más allá de las fáciles retóricas, la vía de fuga solo puede ser una nueva dirección en nuestras elecciones: no se trata de utilizar o no Google y sus servicios, sino de elegir una forma distinta de colocar informaciones en Internet y aprender a relacionarlas de forma nueva, con recorridos nuevos y más interesantes para todos²⁸.

²⁷ La versión integral del report firmado por Jonathan Zittrain y Benjamín Edelman, se encuentra en la dirección <http://cyber.law.harvard.edu/filtering/google/>

²⁸ En este sentido, una fuente rica de referencias y propuestas, indispensable también para la Redacción de este texto, es Ferry Bite, Claudio Parrini: “I motori di ricerca nel caos della rete-kit di sopravvivenza”, shake, Milano, 2001.

Desde hace unos años Google está comprendiendo a su costa (y por supuesto sobre la piel de los usuarios) que la pureza no es una cualidad del mundo y tampoco del mundo de los negocios, que la bondad total es generalmente una tontería, más aún para una sociedad que busca ganancias, y sobre todo, que la neutralidad es un vía difícil por recorrer en la guerra de los motores de búsqueda. Además sería útil recordar que las naciones neutrales desde siempre como Suiza, también son aquellas armadas hasta los dientes. Vamos a ver cuales son las armas “buenas” de Google, las que le han permitido ser un fenómeno de relevancia mundial.

II. BeGoogle

La huida de los cerebros hacia Google: la guerra por el control de la Web

“¡Mataré a ese jodido Google! Eric Schmidt es un jodido y lo enterraré, como ya hice”²⁹: así explotó en mayo de 2005 Steve Ballmer, administrador delegado de Microsoft, a la noticia del fichaje por parte de Google de Kai-Fu-Li, dirigente puntero de la casa de Redmond y su hombre clave en China (Kai-Fu-Li había realizado el motor MSN Search para cien millones de usuarios chinos). Una invectiva dirigida con toda evidencia al actual Administrador Delegado de Google, ex top-manager de Sun Microsystems y de Novell, empresas a las que Microsoft ya se ha enfrentado en el pasado tanto en el mercado como en los tribunales. Kai-Fu-Li ha sido designado a la dirección de un laboratorio de investigación abierto en los alrededores de Shanghai. Microsoft ha interpuesto de inmediato una denuncia en contra del ex dependiente y sobre todo en contra de Google. Esto porque Kai-Fu-Li habría violado, gracias a la determinante ayuda del rey de los motores de búsqueda, unos muy reservados acuerdos contractuales estipulados con el coloso de Redmond. Siendo un director ejecutivo, Kai-Fu-Li conoce secretos comerciales e industriales y, según los abogados de Microsoft, no tendrá reparo en utilizar tecnologías y know-how socio-económico para maximizar las ganancias de la competencia. El hombre del escándalo cuesta caro, aproximadamente dos millones y medio de dólares al año de “sueldo” y 20.000 acciones de la empresa. Cifras exorbitantes que dan la medida de la altísima apuesta en juego, no solo en el mercado chino.

Después de más de un año, la contienda judicial entre los dos colosos se concluye con el pacto de un acuerdo bilateral en diciembre de 2005, justo un mes antes de acabar ante el juez: un acuerdo privado absolutamente confidencial. Puede que hayan sido pagadas grandes sumas de dinero, o quizás Microsoft ha vinculado a Kai-Fu-Li a no revelar nunca jamás ningún tipo de secreto. Esta historia solo es una de las más curiosas y emblemáticas de una tendencia en acto desde hace algún año: de hecho, Kai-Fu-Li solo es el último dirigente pasado a Google, “la sociedad que más se parece a Microsoft”, como ha tronado Bill Gates, cada día más incierto entre demonizar a los dos estudiantes-prodigio -reforzando indirectamente su notoriedad de protagonistas “buenos y generosos” del mundo del IT- o fingir que no existan, que no haya que hacerles demasiado caso.

En realidad, Bill Gates sabe de sobra como las transmigraciones de managers, particularmente en el sector tecnológico, conllevan la asunción del “core business” de la sociedad desde la cual el manager procede: Microsoft ha usado la misma estrategia con sus rivales en diversas ocasiones. La táctica comercial de robar los secretos industriales, mecanismos de producción y gestión de los recursos, mediante el fichaje de elementos punteros de las sociedades de la competencia ha sido siempre un recurso de la competición industrial; se hace cada vez más difusa y frecuente en la era de la economía de la información.

La elección del management por parte de Brin y Page es un claro indicador de los objetivos de Google: llegar a ser la plataforma de la Web más completa y personalizable, promoviendo una personalización de los servicios, sustentada por un ingente cauce de informaciones. En sustancia, Google aprieta el acelerador para catalogar todo tipo de información digital, desde los sitios Web a los foros de discusión, de las imágenes a los videos, de las e-mails a los blogs, sin que se entrevea límite alguno al horizonte; esto significa entrar en guerra con Microsoft, que de momento queda a pesar de todo el mayor competidor en el sector, con su browser Internet Explorer, su portal MSN, su servicio de correo electrónico Hotmail, etc.

²⁹ Declaración bajo juramento de Marc Lucovsky, dependiente de Microsoft, durante una sesión del tribunal del Estado de Washington, para el caso Kai-Fu-Li; Lucovsky, resaltaría también una presunta “sed de venganza” de parte de Microsoft. Las noticias proceden de muchos artículos de Punto Informatico: <http://punto-informatico.it/>

Los intereses de las dos sociedades se están solapando en medida creciente: las dos tienden a ser elementos de mediación indispensables para toda actividad informática. Microsoft ha conseguido imponer su dominación gracias a los sistemas operativos y los programas para oficina, que actualmente son un estándar en el mundo del trabajo y en el home-computing; por su parte, Google se pone hace tiempo como mediador global de los servicios web, como la búsqueda de informaciones (que es su core-business) en todas sus declinaciones, pero también servicios adjuntos como el correo electrónico. Esquematizando en una oposición neta, Microsoft está en una posición de dominio desde hace años gracias a productos que llevan a servicios; Google intenta ser dominante a través de servicios que necesitan productos de soporte. La competición por tanto se fundamenta en las elecciones de los usuarios y en los estándares futuros que Google intenta afirmar. Utilizar programas vía Web limitando al solo browser los requisitos para acceder a los servicios significa sustraer mercado a quien siempre invirtió en el producto y en la creación de nuevas formas de arquitecturas informáticas basadas en la Web. Lo mismo vale para los mercados económicos: se pasa de una economía al por mayor (Microsoft), adonde se intenta vender licencias del mismo programa/producto, a una economía totalmente personalizada para cada usuario, donde los productos se pueden encontrar en la Red.

Largas colas en las Redes: las entrañas de la economía de la búsqueda, Google vs. Microsoft

Google aprovecha, según la fundamental argumentación del extenso texto de John Battelle³⁰, el ascenso de la “economía de la búsqueda”. En su ensayo *The Search*, Battelle, periodista entre los fundadores de *WiRed*, sostiene que el futuro del comercio online estará ligado a las búsquedas personalizadas pagadas por los usuarios. Google, que gestiona el mayor data-base de las “intenciones de búsqueda” de los cibernautas, se encuentra en la posición ideal para satisfacerlas, gracias a su Red capilar, constituida por una plataforma publicitaria excepcionalmente eficaz (AdWords) y por un network de anunciantes que cuenta con millones de sitios en todo el mundo (AdSense). La apuesta de Google es la de satisfacer toda necesidad/deseo que los usuarios expresen en sus búsquedas, ofreciendo servicios para el nuevo consumismo “a la medida del individuo”: cada usuario/cliente encontrará justo lo que busca, el producto estudiado aposta para él. El ejemplo más conocido de “personalización de masa” es el de los servicios de comercio online ofrecidos por Amazon, en condición de ganar mucho más de la venta de las copias individuales de un libro o de un CD que de los picos de ventas de cientos o miles de copias de best-sellers. Los muchísimos clientes que compran libros no especialmente populares online, representan una cantidad de “eventos” que se verifican raramente, o hasta una sola vez; llegar a satisfacer estas “búsquedas personalizadas”, representa el punto de fuerza de la distribución de Amazon. De hecho para una librería tradicional, basada en almacenes, stocks, pedidos mínimos, es imposible conseguir la disponibilidad de millones de títulos como hace Amazon; las mayores ganancias proceden por tanto de las novedades y de los títulos mas vendidos. La distribución de una copia única para satisfacer la petición de un solo cliente resulta antieconómica para una librería tradicional, mientras que para Amazon es posible y rentable, aprovechando la “economía de la búsqueda” de los mercados online.

Este tipo de mercados se denomina en jerga económica de “cola larga” (Long Tail)³¹.

³⁰ John Battelle, *The Search How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Cultura*, Portfolio Handcover, 2005 (trad.it Google e gli altri. Come hanno trasformato la nostra cultura e riscritto le regole del business, Cortina, Milano, 2006).

³¹ Las colas largas: Las implicaciones comerciales de la Cola Larga (Long Tail) –en el contexto del e-commerce la definición se ha vuelto de uso común- han sido difundidas por Chris Anderson, Redactor jefe de “*WiRed*”, en un artículo de octubre de 2004 escrito para su revista. Anderson ha publicado en 2005 el volumen *The Long Tail*. Como Battelle, Anderson ha mantenido un blog sobre la génesis de su libro en www.thelongtail.com y el término se ha vuelto de uso corriente. Las colas largas no son un invento de Anderson: se trata de un concepto familiar para

La teoría estadística de las colas largas se remonta al menos a la “distribución de Pareto”³², en la que “pocos eventos se verifican con extrema frecuencia, muchos eventos se verifican con escasez”. Una distribución parecida de los eventos se describe, a nivel estadístico, con una hipérbola en la que la “cola larga” está representada por la multitud de “eventos”, singularmente casi irrelevantes, pero en su conjunto numéricamente muy elevados. Matemáticamente, una distribución a cola larga sigue la denominada “ley de potencia”³³.

La estrategia ganadora en los mercados a “cola larga” por tanto no es la bajada de los precios de los productos más populares, sino la ampliación de la oferta, que permite mayores ganancias a través de la venta de productos distintos en pocos ejemplares.

En términos comerciales, la mayor rentabilidad procede de las pequeñas ventas. En efecto, la mayor parte del business en Red se basa en la cola larga: el mismo Google produce renta vendiendo publicidad a bajo coste a millones de personas en forma de banners textuales (text ads) mucho más que vendiendo pocos espacios de elevado coste a importantes anunciantes. Battelle se interesa por la aplicación de la investigación a mercados no explotados: en el caso de Google, la enorme cantidad de datos a disposición para realizar búsquedas, ha permitido, según él, la explotación comercial de las colas largas. En el contexto del e-commerce las colas largas tienen tres implicaciones. En primer lugar, a través de Internet los productos con poca demanda pueden, colectivamente, crear un mercado superior al del pequeño grupo de los artículos muy vendidos. En segundo lugar, Internet da vida a una proliferación de vendedores, así como favorece una proliferación de mercados (es el caso del sistema de subastas online de e-Bay). Por último, gracias a la investigación, se hace posible el paso de los mercados de masa a los de nicho.

Esta última tendencia se basa sobre la formación de grupos espontáneos, un fenómeno muy difundido en las Redes. En Internet, hasta los grupos numéricamente más relevantes no son masas homogéneas de personas, sino variadas comunidades de usuarios unidos por una pasión, un interés o un fin común. La oposición entre masas y nichos no tiene por tanto mucha relevancia para la detección de segmentos de mercado útiles. Al nivel comercial esto provoca el nacimiento de sitios que proponen e-commerce de productos ligados a tipologías muy específicas de usuarios, que fuera de la distribución online no representan en ningún modo un mercado interesante: por ejemplo, camisetas con logos para cyber-maniáticos, relojes que marcan la hora en código binario, ordenadores luminosos y otros objetos dedicados a tecnófilos. La amplitud de la oferta compensará la escasez de la demanda, repartida entre un número elevadísimo de productos altamente personalizados.

En tal escenario es posible, como pone en evidencia el atinado artículo de Charles H.Ferguson³⁴, que Google y Microsoft luchen para el control de la catalogación, de la búsqueda y extracción de todas las informaciones, en todo tipo de dispositivo digital. Microsoft está invirtiendo mucho en los servicios Web. En noviembre de 2004 ha lanzado una versión beta (experimental) de un motor de búsqueda proyectado para contestar a las peticiones formuladas en lenguaje corriente y para ofrecer resultados personalizados sobre la base de la localización geográfica del usuario³⁵; en febrero de 2005 el motor MSN search, ha sido mejorado. Con MSN search es posible también

cualquiera haya sostenido un examen de estadística. Hay distribuciones estadísticas muy comunes, cuyos gráficos describen un modesto número de eventos que ocurren con frecuencia, y un gran número de eventos (la cola larga) que suceden raramente. En el conjunto, los eventos raros pueden ser más numerosos de los normales

<http://www.technologyreview.it/index.php?p=article&a=551>. La cola larga en Italia: un artículo sobre la Webmarketing: <http://marketinginformatico.com/la-teoria-della-coda-lunga.php>

³² Vilfredo Pareto (1848-1923) fue sociólogo y economista. Una buena introducción se encuentra en Wikipedia, http://es.wikipedia.org/wiki/Vilfredo_Pareto

³³ Véase: A.L. Varabais, Link, Einaudi, Torino, 2001

³⁴ Véase el artículo en italiano de Charles H.Ferguson, publicado en Technology Review: <http://www.technologyreview.it/index.php?p=article&a=340>

³⁵ El motor de búsqueda de Microsoft: <http://search.msn.com>

buscar dentro de Encarta, la enciclopedia multimedia de Microsoft; todavía al momento la consulta está limitada a dos horas y, en el curso de la navegación un contador de minutos va actualizando el tiempo que queda. En suma el gigante Microsoft ha decidido desarrollar un sistema propio para las búsquedas en la Web y en los ordenadores personales sin pasar por Google que desde varios años es el triunfador de la búsqueda (tan solo Yahoo! puede presumir de una influencia comparable).

En su conjunto, los mercados ligados a la economía de la búsqueda parecen mucho más amplios del actual mercado de los servicios de búsqueda. Microsoft anda con retraso, es cierto, pero podría poner en marcha estrategias feroces e inalcanzables por Google: por ejemplo, podría invertir en pérdida y ofrecer acceso gratuito a su motor de búsqueda, después de haberlo integrado en el browser Internet Explorer; o, quizás, vender publicidad por debajo del coste a los anunciantes para arrebatar liquidez a su competencia. Mientras tanto, el próximo sistema operativo desarrollado en Redmond, Windows Vista, debería ofrecer herramientas de búsqueda innovadoras. Por otro lado, Microsoft llevaba un fuerte retraso con respecto a Netscape (el primer browser descargable gratuitamente de la Red) en la batalla para el dominio sobre los browser, aún así Explorer lo ha conseguido, y no cierto por ser el mejor browser del mercado. Si Microsoft tiene de su parte una larga experiencia, una posición de predominio en el mercado y un enorme poder financiero, también Google tiene unas buenas cartas para jugar. Es la encarnación de la joven empresa trepante, se ha construido una fama de compromiso con la investigación y dedicación a la excelencia técnica, predica el culto de la velocidad en satisfacer las peticiones de los usuarios con interfaces sencillas y eficaces: se ha impuesto porque técnicamente mejor de los demás motores de búsqueda. En la competición para el control de la Web, Google parecería ligeramente aventajado; pero no hay que olvidar que la extensión de las actividades de Microsoft no tiene comparaciones, no se limita al Web sino que además la entera gama de la IT, desde las herramientas con Windows o MS Office, a los contenidos con Encarta, a la búsqueda de alto nivel con plataformas como dotNet, etc. Considerada la apuesta en juego - el dominio del acceso a cualquier información digital, y las ganancias derivadas- una pacífica convivencia parece cuando menos improbable. Por ahora Google ha postergado el choque, pero no podrá tardar mucho más.

La guerra de los estándares

Continuamos siguiendo Ferguson en su razonamiento: el escenario que se abre es aquello de una guerra de los estándares. Al momento hay tres actores involucrados: Google, Yahoo! y Microsoft. La industria de la investigación, como resulta también del análisis de Battelle, está en fuerte crecimiento. Las tecnologías emergentes o en vía de consolidación, como por ejemplo el *streaming audio* y video a través de la banda ancha, el VoIP (telefonía vía Internet: piénsese en Skype, y al homólogo servicio de Google, Gtalk) y la mensajería instantánea generan enormes cantidades y tipologías de datos que aún esperan ser indexados, accesibles por toda una variedad de nuevos dispositivos electrónicos (de los palmares a los celulares y los lectores de documentos audio-visuales de nueva generación, a los navegadores vía satélite, etc.), cruzados entre sí para satisfacer las peticiones de los usuarios, y en última análisis explotados por una publicidad pervasiva de todo tipo. Para poder cooperar recíprocamente y con muchos otros sistemas se necesitarán nuevos estándares, difíciles de introducir de forma indolora.

Las guerras entre los estándares se producen porqué el mercado de la tecnología de la información necesita de idiomas comunes para poder gestionar el aumento de complejidad. El valor de la información reside en su difusión; pero difundir e intercambiar textos es más simple que difundir e intercambiar audio, o incluso video: una mayor cantidad de datos requiere redes de comunicación más potentes y cambios en las modalidades de gestionar las informaciones. Históricamente, el control propietario sobre un importante estándar tecnológico ha producido

siempre enormes riquezas³⁶. A veces, como en el caso de la Web, la adopción de un estándar abierto y no propietario, como es el protocolo HTTP, lleva a una situación ventajosa para todos. Frecuentemente sin embargo, no se imponen las “mejores” tecnologías, sino que “es más importante contar con una estrategia eficaz. De todas formas los ganadores tienden a connotarse por algunas características importantes. Suelen vender plataformas universales, independientes del hardware, como los sistemas operativos Microsoft, en lugar de soluciones hardware y software fuertemente integradas, como los sistemas Apple y Sun. Las arquitecturas ganadoras son propietarias y difíciles de imitar, pero son también extremadamente “abiertas”: o sea ponen a disposición interfaces accesibles públicamente para el desarrollo de los programadores independientes y los usuarios finales. De esta forma la arquitectura puede alcanzar todos los mercados y crear un punto de enlace y de bloqueo: en otras palabras, empuja a los usuarios hacia aquella arquitectura específica, haciendo casi imposible el paso a un sistema de la competencia, sin grandes dificultades y gastos”³⁷. Se impone un estándar cerrado para explotar una situación de monopolio.

Un ejemplo que aclara la lucha para la hegemonía sobre los estándares es el desafío citado antes entre Gtalk y Skype. En estos momentos Skype ha conseguido una posición de monopolio en la gestión del VoIP para el mercado doméstico; pero es posible que no hayan sido evaluados atentamente los tiempos de asimilación de las tecnologías abiertas dentro de las comunidades de desarrollo. Hasta hace poco, Skype era la única solución realmente funcional para cualquiera que quisiera telefonar vía Internet, aún no teniendo ningún conocimiento técnico; pero paulatinamente las cosas están cambiando. Las tecnologías propietarias de Skype podrían ser superadas por Gtalk, basado totalmente en tecnologías libres (particularmente en el protocolo de comunicación “jabber”) con librerías de desarrollo concedidas bajo licencias copyleft, que favorecen la confluencia de energías en el proyecto pues estimulan los programadores a potenciar la Red VoIP de Google. En este caso la elección del Open Source podría revelarse ganadora para quitarle el predominio a Skype, la que por su parte naturalmente podría decidir hacer públicos sus códigos para mantener sin cambios la situación. El hecho de elegir adoptar tecnologías y plataformas propietarias, que cierran el acceso (concediendo a lo mejor interfaces públicas de desarrollo), o abiertas, es una parte fundamental de la estrategia para el control de la Web y de la economía de la investigación.

El acceso al mercado de la investigación ya ha sido bloqueado (en jerga económica “*lock-in*”). De hecho, es imposible que una empresa *start-up* pueda competir ahora con Google o Yahoo! para la indexación de miles de millones de páginas Web: la inversión para la maquinaria necesaria sería prohibitiva, aún cuando se dispusiera de un mejor algoritmo para su propio *spider*. De todas formas muchos aspectos colaterales, sobre todo la interfaz entre los distintos sistemas de búsqueda, dejan amplios espacios a posibles innovaciones “críticas” a un coste relativamente bajo. Es el caso de las librerías informáticas, pequeñas parcelas de software que permiten la conexión entre sistemas heterogéneos, puentes de traducción de un sistema a otro; las cuales, junto con las metodologías de integración entre distintos dispositivos y de intercambio de informaciones y resultados de búsqueda, podrían ser desarrolladas por pequeños investigadores independientes, en lugar de las grandes empresas involucradas³⁸. Más adelante profundizaremos en estos aspectos ligados a las interfaces y a las librerías.

³⁶ El predominio arquitectural es equivalente a fabricar moneda y, si gestionado de forma adecuada, puede durar para siempre. La arquitectura de los ordenadores mainframe IBM ha sido introducida en 1964; Intel ha desarrollado su primer microprocesor en 1971; el primer sistema operativo Microsoft ha sido presentado en 1981; Cisco Systems ha vendido su primer router en 1986. Ninguna de estas tecnologías está dando señales de cansancio y cada una de ellas ha generado cientos de miles de millones de dólares de facturado.

³⁷ Charles H. Ferguson, cit., Technology Review

³⁸ Véase Hipólita, Open non é Free, Eleuthera, Milano, 2005, Cáp. V

Por ahora es importante subrayar que ninguno de los actores involucrados tiene una posición de predominio absoluto, y obviamente esto es un bien para todos: intentad imaginar qué significaría un monopolio total del mercado de la investigación conseguido por un sujeto privado cualquiera a través de la imposición de hecho de un estándar. Las primeras y obvias consideraciones habría que hacerlas a propósito de la *privacy*: ¿de quien serían los datos indexados con los que se efectuarían las búsquedas y se generarían ganancias astronómicas? Además, puesto que hoy en día, simplemente digitando el nombre de una persona en Google, se accede a una impresionante cantidad de datos, en un futuro no lejano en el que la cantidad y la calidad de las informaciones, y sobre todo la posibilidad de llevar a cabo búsquedas agregadas sobre datos heterogéneos, mejorará drásticamente, el control sobre los individuos podría volverse infinitamente más agobiante y total, cruzando datos sensibles, fichas médicas, conversaciones telefónicas, informaciones del ADN. Google podría ser el punto de acceso privilegiado al panóptico³⁹ digital: veamos por tanto cuáles son sus mejores armas en esta verdadera guerra por el control de las redes.

Arma N 1: GooglePlex, el capitalismo blando

La hagiografía corriente de Google cuenta con deleite el impresionante crecimiento que ha llevado Brin y Page desde el cuchitril de Stanford al garaje de Menlo Park, California, que un amigo había alquilado a la recién nacida sociedad Google Inc., al despacho en la University Avenue de Palo Alto, hasta la mudanza al Googleplex, los actuales despachos de Mountain View, California. Entre 1998 y 2000 los dos amigos expanden la misma cultura universitaria, desarrollando una cultura de empresa hecha de innovación, creatividad y sacrificio. La dedicación universitaria para la investigación aplicada al business: esta la llave del éxito. Googleplex atrajo desde el comienzo legiones de colaboradores: fue recreada la atmósfera del campus universitario americano, donde estudio, perseverancia, deporte, ocio se complementan y fusionan: si un ambiente lleno de comodidades y relajado favorece la creatividad entre los estudiantes, por supuesto podía favorecer también la productividad de los trabajadores. El espíritu académico de la cofradía, de la élite universitaria que trabaja celeremente para conseguir siempre lo mejor, se desprende de los cuentos acerca de Googleplex: parece que secciones enteras del parking fuesen cercadas dos veces por semana para dejar espacio a los partidos de jockey sobre ruedas; hoy más que nunca en los despachos se derrocha todo tipo de gadgets, y sobre todos despuntan las lámparas-lava con sus emulsiones coloreadas; la informalidad elegida a sistema, con Larry y Sergey quienes dirigían las reuniones semanales, junto con las decenas de empleados, llamándolas TAIF (Thank God It's Friday- gracias a dios es viernes) en el *open space* entre las mesas de trabajo. Ciertamente tal atmósfera de trabajo alentó el espíritu de grupo y aceleró el intercambio de ideas. Googleplex parece casi ser el lugar para expresar las propias pasiones de investigación, en lugar de un puesto de trabajo; seguramente no es un puesto de trabajo cualquiera, a pesar de las ya enormes dimensiones. Y eso que la dimensión organizativa estilo "campus", en USA, existe desde hace al menos treinta años: Microsoft y Apple, tan solo para dar dos ejemplos, adoptaron desde siempre un perfil de este tipo. La mitología de Silicon Valley está repleta de historias acerca de la creatividad por encima de todo, de la importancia de la cooperación entre los colaboradores: en suma, de la necesidad de tener trabajadores felices de currar y considerar los objetivos de la empresa al mismo nivel de sus propios objetivos vitales, en lugar de trabajadores

³⁹ El panóptico auspiciado por Bentham y analizado por Foucault encuentra en el proyecto Echelon una de sus realizaciones más escalofrantes. Sin dejar lugar a paranoias complotistas, señalamos entre los muchísimos volúmenes y artículos un texto breve introductorio al argumento, Alex Jones y Paul Joseph Watson, Il panottico, una prigione di sorveglianza per l'umanità, 20/01/2006 http://www.ariannaeditrice.it/articolo.php?id_articolo=1878, versión original <http://www.prisonplanet.com/articles/january2006/110106thepanopticon.htm>

frustrados, aplastados por jerarquías agobiantes, atrapados por horarios de trabajo fijos, reglas rígidas, ambientes opresivos.

Quizás la novedad de Googleplex, está en haber promocionado conscientemente, desde el inicio, la imagen de sí como lugar “diferente”, “nuevo”, “para las mejores mentes”. Entrar en Googleplex es imposible si no se conoce nadie dentro. En el interior está vetado incluso hacer fotos, al menos en teoría. Como si hubiera que defenderse del mundo exterior, poblado por feroces tiburones financieros y malvados depredadores que solo quieren explotar el talento de los Google-boys.

Todo el mundo quiere trabajar en Googleplex. Una lista no oficial de las amenidades que esconde es casi obligatoria: gimnasios de empresa, piscinas, comida gratis en los cuatro restaurantes (uno de ellos vegetariano), drinks y snacks gratis por doquier (¡basta ya con la esclavitud de las maquinitas de pago Google lo paga todo!); campos de volley-ball, basket, y espacios al aire libre para hacer deporte, monopatines de motor para moverse entre los edificios del campus-empresa. Y esto no es nada, frente a la guardería infantil, a la escuela de infancia y a las escuelas primarias de empresa para los hijos de los dependientes, totalmente gratuitas, o al estudio dental, o sea un camión adaptado a estudio dental móvil: en un país como Estados Unidos, donde la instrucción y la sanidad son un lujo para pocos, se trata de oportunidades aún más increíbles.

Los puntos de trabajo son espectaculares, el sueño de todo fanático de las tecnologías digitales: monitor estándar de 21 pulgadas LCD, pantallas de plasma por doquier, juegos y juguetes (modelos del tamaño natural de los personajes de Star Wars, gadget tecnológicos de todo tipo). Omnipresentes, las lámparas-lava con sus colores fluorescentes.

Googleplex es un lugar de ensueño, sumergido en el verde, donde los horarios de trabajo son flexibles, donde todo parece posible. En resumen, Googleplex transmite la filosofía de Google, el estilo de vida de Google: por supuesto hay también una tienda donde comprar todos los gadgets corporativos que se puedan imaginar, alcanzables online en un sitio adhibido a ello; en su mayoría se trata de himnos a la inutilidad y al exceso como todo gadget que se respete, pero muy eficaces en hacer del sentido de pertenencia a la empresa una razón de orgullo. La camiseta y el gorro con el logo de la compañía para consolidar el espíritu de equipo son verdaderos anacronismos: el presente de la adhesión blanda es Googleplex.

Además, no solo Google, sino también empresas como Apple y Yahoo! ofrecen desde hace tiempo un verdadero catálogo de objetos de consumo, desde líneas completas de vestuario a todo tipo de accesorios high-tech, lectores MP3 y pen-drive USB, hasta los objetos de mobiliarios, todos oportunamente marcados con logos y lemas corporativos. El muy versátil negocio de Google, ofrece monos para los recién nacidos, y también Google Mini, el sistema de empresa para listar “como Google” todos los materiales propios.

Googleplex es la encarnación del capitalismo de la abundancia en la era del informacionalismo⁴⁰: todas las informaciones del mundo ofrecidas a todo el mundo de forma gratuita. Ya no es tiempo de escasez: la abundancia y la disponibilidad de los bienes (en este caso de las informaciones) es total. No olvidemos, sin embargo, que al fin al cabo la fuente de esta abundancia es prácticamente solo la publicidad hecha principalmente con link textuales. Todo lo demás es *free*, libre. No todo pero funciona perfectamente⁴¹.

⁴⁰ Acerca del concepto de informacionalismo vease en particular Manuel Castells, *La era de la información: economía, sociedad y cultura*, Alianza Editorial.

⁴¹ Hay muchas señalizaciones de gazapos de Google; parece obvio que se multipliquen los problemas acorde con el crecimiento del tamaño y de los servicios. Además, la práctica de publicar versiones beta que los usuarios testan extensivamente acrecienta la posibilidad de malfuncionamientos. Un ejemplo concreto: Marco Deseri en Goggle Analytics: <http://www.deseri.it/articoli/google-analytics-rallenta-il-sito>

Dos de los más influyentes personajes de este capitalismo de la abundancia son Mike Moritz⁴², galés, quien, además de Google ha financiado también Yahoo; y John Doerr⁴³, quien también tuvo importantes participaciones en Sun y en Netscape.

Arma 2: perfeccionar la estrategia de la acumulación

Después de una visita a Googleplex se aclara la razón de la huida de los cerebros hacia Google de la que hablábamos al comienzo del capítulo. Para un obrero de las IT, o mejor sería decir peón de las IT, trabajar en Googleplex es un sueño hecho realidad. En este sector cada vez son más los obreros explotados; un perfil profesional ejemplar es aquello del *coder* independiente quien desarrolla proyectos personales, a lo mejor publicándolos en Sourceforge.net o Slashdot.org, y ofrece sus conocimientos en el mercado sin tener ningún tipo de protección sindical, contrato de categoría o similares, ya restos prehistóricos en nuestra época de la flexibilidad total. En Googleplex hasta podrá dedicar el famoso 20% del tiempo de trabajo a sus proyectos, pagado y motivado a hacerlo cada vez mejor.

Entre partidos de hokey y volley-ball, perros que merodean por el campus/empresa y en los pasillos, reuniones alrededor de una mesa de ping-pong, ciertamente es difícil aburrirse. Ya que no es fácil fichar nuevos colaboradores en grado incluso de mejorar este tipo de atmósfera, Google se ha distinguido por sus inventos en materias de fichajes. La más curiosa quizás sea el test realizado en julio de 2004: en algunas estaciones del metro de la ciudadela de Cambridge, en el estado de Massachussets (USA), y en la autopista 101 en California, han aparecido grandes vallas blancas con la escrita:

{first 10 digit prime in consecutive digits of e}.com

o sea

{primer numero primo de 10 cifras consecutivas en e}.com

El logaritmo natural requerido es el numero 7427466391; en la dirección

<http://www.7427466391.com/> indicado en la valla nos encontramos en una dirección IP de

Google que nos pide de completar una secuencia de números; encontrado el numero 5966290435, se siguen las instrucciones, utilizándolo como password para entrar en una sección del sitio <http://www.linux.org> (paso para nada descontado) y desde aquí se nos rebota otra vez a Google, en el sitio <http://www.google.com/labjobs/> en donde nos piden de enviar nuestro currículum: si hemos logrado resolver los enigmas, podríamos ser un buen fichaje para Google.

Google no atrae tan solo a los mejores técnicos, hackers y manitas geniales: pronto también los managers más cotizados del mundo se dan cuenta de sus potencialidades y compiten para entrar en su plantilla.

La estrategia de la acumulación de datos para la búsqueda, y de máquinas en Red para almacenar una copia de los datos, procede al unísono con la acumulación de cerebros. Máquinas significantes, máquinas electrónicas, máquinas biológicas, se acumulan en Googleplex, para dar

⁴² Galés, periodista en el Time Magazine en los años Setenta y Ochenta, hoy se considera como un “*veteran venture capitalist*” y parece tener una receta para todos: recrear hoy fenómenos como Google? Lo importante, dice, “es crear algo que sencillamente pueda ser útil para sí mismos”. Artista, si no hubiera dado prioridad a los compromisos de trabajo, es hombre de marketing vieja escuela, pero siempre incisivo: las ideas correctas son aquellas que satisfacen a sí mismo, por tanto a los consumidores, mejora aún los “usuarios”

⁴³ John Doerr continua también después de Google, sus inversiones: últimamente ha apostado 16 millones de dólares en un proyecto de nombre “Zazzle”, cuyo *core business* es el mercado online adonde se pueden comprar y vender productos altamente personalizables. En la práctica, Doerr no quiere quedarse fuera de la “personalización de masa”, en fase de actuación avanzada en Estados Unidos. Inspiradores del proyecto son personajes como Eric Von Piel, quien, “desde su cátedra de la Sloan School of Management teoriza al MIT desde hace tiempo la idea de que el próximo paso será exactamente ese: transformar la producción de masa fordista y post-fordista en algo completamente distinto, a medida, creado y gestionado a través de la Web”

http://www.bloggers.it/Marcioweb/index.cfm?idForum=News%20da%20Mondo&blogaction=archive&file=blog_7_2005.xml

vida a un estilo de vida, a una especie de religión de la excelencia, encarnada por un evangelizador.

El manager quien mejor encarna el estilo de la empresa, el dichoso “evangelista” de Google, entre tantos jóvenes, es un gran viejo de la Red, Vinton G. Cerf, inventor junto con Robert Kahn del protocolo TCP/IP. Las circunstancias merecen un pequeño paréntesis: en febrero de 2005 Google confirmaba que el ICANN, el organismo que supervisa el sistema de los nombres y de los dominios Internet, había dado el OK para que la empresa de Mountain View pudiera entrar en el business del registro de los dominios. En septiembre de 2005 Google anuncia que Vinton Cerf ha sido nombrado “vicepresidente y Evangelizador Jefe de Internet para Google. Su misión consiste en identificar nuevas tecnologías y aplicaciones estratégicas para la empresa en Internet y en otras plataformas”⁴⁴. Hasta ese momento, entre otros cargos, Cerf estaba sentado en el Consejo de Administración de la ICANN. De forma distinta del fichaje de Eric Schmidt como Administrador Delegado, y de otros top managers quitados a la competencia, ésta es sin duda alguna una operación exclusivamente de imagen (difícil pensar que Cerf trabaje de verdad en Googleplex, por cuanto divertido pueda ser...).

Arma 3: la imagen no lo es todo, pero un poco de “filosofía” no molesta

La imagen de Google no es tan solo el sitio y el interfaz⁴⁵ que han conseguido tanto éxito por su sencillez y velocidad; no solo es Googleplex, el mundo dorado de los prodigios tecnológicos; no solo es “ser buenos” y hacer dinero, conjugando estrategias comerciales descaradas con los guiños de ojo a la cultura académica y a las comunidades Free Software y Open Source, con financiaciones e iniciativas hechas a su medida.

La imagen es también y sobre todo la “filosofía” de la empresa, la explicación clara y meridiana, en píldoras, del Google-pensamiento. No se trata precisamente de investigaciones llevadas a cabo por amor del conocimiento y de la claridad, por esto quizás el término filosofía es un poco exagerado; en cualquier caso, on-line se pueden encontrar diez puntos, los diez mandamientos del buen Google. Se inicia con una introducción, que enuncia el estilo: “Nunca conformarse con lo mejor”; de hecho, según afirma Larry Page, el objetivo último es “El motor de búsqueda perfecto”, que debería comprender exactamente la petición del usuario y devolver exactamente lo que él desea”. Por lo tanto Google no quiere ser para muchos: quiere ser para todos, satisfacer los deseos de todos, llevar la felicidad en el mundo. Para lograrlo, se dedica con tenacidad a la innovación y a la investigación, siguiendo las que define sus “Diez verdades”:

1. Atención al usuario: todo lo demás viene después. El crecimiento de Google ha sido fruto del pasa palabra, que ha atraído a usuarios encantados por sus prestaciones, y no de agresivas campañas de comunicación. La publicidad no debe ser invasiva, sino útil al usuario.

2. Es preferible dedicarse verdaderamente bien a una sola cosa.

“Google se dedica a la búsqueda. Con uno de los grupos de investigación más grandes del mundo dedicado exclusivamente a la solución de los problemas de la búsqueda, tenemos conciencia de lo que hacemos bien y de cómo podemos hacerlo mejor”. Hemos visto que Googleplex se configura exactamente como un centro de investigación universitario. Los nuevos servicios añadidos respecto al simple motor de búsqueda se ven como oportunidades ofrecidas a los usuarios para orientarse y encontrar las informaciones que desean sin perderse en masas cada vez más grandes de datos.

3. Rapidez es mejor que lentitud.

⁴⁴ Véase <http://www.google.it/intl/it/corporate/execs.html#vint> recordamos que en USA prácticamente cualquiera no sea empleado es vicepresidente de algo.

⁴⁵ Véase el Cáp. V

“Google cree en la gratificación inmediata. El usuario quiere respuestas y las quiere ya. ¿Quiénes somos nosotros para no estar de acuerdo? Google es prácticamente la única sociedad del mundo cuyo objetivo manifiesto es procurar que sus usuarios abandonen el sitio de Google lo antes posible”. Dos son las principales intuiciones y realizaciones que le han permitido a Google obtener “velocidad”: la puesta a punto y el perfeccionamiento continuo del algoritmo PageRank, que indexa sin pausa las redes; el uso de plataformas modulares fuertemente escalables ligadas entre ellas

(*cluster*). De todas formas, sería oportuno reflexionar más en profundidad sobre la velocidad como panacea: a veces, también en el entorno de las nuevas tecnologías, la lentitud es una virtud⁴⁶.

4. La democracia en la Web funciona.

“El éxito de Google está basado en millones de individuos que publican sitios Web, para determinar cuales otros sitios ofrecen contenidos validos”. Ya sabemos que Google utiliza PageRank para evaluar los sitios ligados a una página Web y asignarles un valor, basado en parte en los sitios ligados a estos. La imagen de esta democracia electrónica es idílica: el resultado del ranking de Google sería un “índice de agrado popular” basado en un algoritmo/ley electoral que permitiría a los usuarios/ciudadanos de la Red expresar sus preferencias/votos a través de los links ofrecidos y aceptados por otras páginas, y encontrarlas puntualmente expresadas en el posicionamiento de los sitios Web votados. La comparación link/voto es por lo menos forzada y simplista, pues continuamente se introducen “correcciones” para calcular el ranking, “pesando” de forma diferente los link/votos. Seguramente el link/voto de un sitio porno contará menos del link/voto de un sitio universitario. Aunque sea al menos discutible que la cultura académica sea más popular que la cultura pornográfica...Lo que es cierto es que, con el crecimiento de las informaciones, esta “democracia” está en expansión continua.

5. No hace falta estar en su propia mesa de trabajo para necesitar informaciones. “La sociedad actual es cada vez más móvil y menos dispuesta a estar vinculada a un lugar fijo. A través de los PDA, los teléfonos wireless o sus automóviles, los usuarios desean que las informaciones vayan hacia ellos”. La flexibilidad de los tiempos y de los espacios es un objetivo importante: la convergencia de los dispositivos electrónicos (TV, radio, teléfono, Internet...) en dispositivos móviles en miniatura es una oportunidad total para el mayor proveedor mundial de soluciones de búsqueda. La penetración en este mercado del futuro es estratégica, lo hemos visto hablando de la “guerra de los estándares”: más aún para Google, que no produce dispositivos electrónicos para los cuales proponer/imponer su software (Microsoft, Apple), sino interfaces de búsqueda. Cada nuevo dispositivo es un territorio por conquistar.

6. Se puede ganar sin hacerle daño a nadie.

“Google es una empresa. Los ingresos generados por la sociedad proceden de la oferta a las empresas de su tecnología de búsqueda y de la venta de publicidad colgada en Google y en otros sitios Web”. De todas formas los anuncios son textuales y poco invasivos, los links de pago son pertinentes a las búsquedas realizadas (AdWords); y los usuarios pueden transformarse en anunciantes de forma sencilla y self-service; si se trata de un gestor de sitios Web, puede hasta aprovechar la Red de Google a través de AdSense, publicando anuncios pertinentes con el contenido de sus sitios. “No seas malo”, “No hagas daño a nadie” significa también “No hagas a los demás la publicidad que no te han hecho a ti”, y por supuesto garantizar que el PageRank no es objeto de venta. La confianza de los usuarios en la precisión de los resultados es la fuerza principal de Google, y no se puede desperdiciar para ganancias a corto plazo. Se aprovecha para generar ingresos indirectos, de segundo nivel (generados por la publicidad).

7. Siempre hay más informaciones de lo que se pueda imaginar.

⁴⁶ Para una panorámica completa sobre el tema vease: Thomas Hylland Eriksen, *Tempo tiranno – Velocità e lentezza nell'era informatica*, Eleuthera, Milano, 2003

“Después de haber indexado más paginas HTML en Internet de cualquier otro servicio de búsqueda, los técnicos de Google han dirigido su atención a las informaciones no inmediatamente accesibles”. Sí, Google acumula bases de datos heterogéneas: imágenes, mensajes de newsgroups (Usenet), números de teléfono y direcciones, informes financieros. Para ser el principal mediador de informaciones del mundo, no hay que poner límites a la acumulación.

8. La necesidad de informaciones traspasa todo confin

“A pesar de que la sede de Google esté en California, tenemos oficinas en todo el mundo y nuestra misión es facilitar el acceso a las informaciones a las personas de todo el mundo”. Cultura americana, estrictamente académica, para todos. Pensar en grande, siempre y como sea: indexar todas las informaciones, ofrecer acceso a todo el mundo. Las traducciones son parte fundamental del universalismo de Google: desde el coreano a la jerga hacker al klingoniano de Star Trek, del hindy al xhosa surafricano hasta al pig latin (o sea el alfabeto “farfallino” americano), desde el zula a la interlengua y al bork bork bork (una especie de falso ruso molvano...) todos deben de poder acceder. Más de cien idiomas de interfaz. Dominios locales de más de cien naciones. Loable, pero un poco centralizador...operación quizás políticamente correcta y formalmente respetuosa de las minorías, aunque en realidad se trata de un *super-layer*, un nivel superior, una única interfaz que allana, homogeniza las diferencias y difunde el estilo de Mountain View.

9. Se puede ser serios aún sin chaqueta y corbata.

“Los fundadores de Google muchas veces han declarado que la sociedad es “seria” solo por lo que atañe a la investigación. Ellos han edificado una empresa basada en la idea de que el trabajo debe ser un desafío y que dicho desafío tiene que ser divertido”. En este mandamiento se compendia la organización de Googleplex como un campus universitario para maximizar la productividad: se afirma explícitamente que “El énfasis está en los resultados logrados por el equipo y en el orgullo de los hallazgos individuales que contribuyen al éxito global de la sociedad”, y que “Este ambiente extraordinariamente comunicativo favorece una productividad y una camaradería alimentadas

por la conciencia que millones de personas confían en los resultados de Google. Dad los instrumentos adecuados a un equipo de personas que quieren ser diferentes y lo serán”. Quizás este sea la mejor manera de aprovechar los creativos, haciéndolos al mismo tiempo entusiasmados sostenedores de la “Experiencia Google”.

10. No basta con ser excelentes.

“Dar siempre más de lo esperado. Google no acepta la idea de que el objetivo final sea el de ser los mejores, sino que tiene que ser el punto de partida. Con la innovación y la iteración, Google se dedica a mejorar algo que ya funciona bien para conseguir resultados inesperados”. Evidentemente, para satisfacer rápidamente, cada vez mejor, los deseos de todos los usuarios del mundo, hace falta retrasar un poco más allá, la satisfacción de sus propios deseos. Desear de desear de ser los mejores. En este contexto llegar segundos es todavía peor que no existir. Por lo que nos atañe, eslogan por eslogan, preferimos “Dar asco, en una sociedad que obliga a la excelencia, es un claro deber moral”⁴⁷.

Arma 4: Google y el Open Source

El arma posiblemente más compleja de Google es su estrategia de colaboración-explotación del mundo Open Source. La iniciativa Google Code (marzo de 2005) es un homenaje a la comunidad Open Source, de la cual los fundadores de Google se declaran “amigos y deudores”. El motor más usado de la Web explica que el fin de Google Code no es promocionar el

⁴⁷ SantaMariaVideo, la Tv que no transmite nada: <http://www.santamariavideo.tv/video/anonimato.mpg>

desarrollo de aplicaciones que aprovechan sus propias API (Application Programming Interface), que ya disponen de un sitio ad hoc, sino el de poner a disposición de todos los desarrolladores Open Source herramientas de utilidad pública. Los 4 proyectos inicialmente publicados en Google Code nacen alrededor de programas realizados por los mismos desarrolladores de la sociedad para agilizar la creación, la optimización o el debugging del código. Los proyectos linkados en Google Code también están hospedados en SourceForge.net y se distribuyen bajo licencia BSD 2.0 (su código puede utilizarse indistintamente dentro de otras aplicaciones Open Source o propietarias). Además, Google promete desde ya poner a disposición de la comunidad Open Source muchos otros softwares, fruto en buena medida de ese 20% de tiempo que los desarrolladores dependientes de la sociedad vienen alentados a dedicar a proyectos personales. No es casual que Google siga un proyecto dedicado al fichaje de desarrolladores Open Source, con la “Summer of Code”, una verdadera competición con un premio de 4500 dólares. Luego le toca a “Google Earth” y por fin, como toda potencia que haya llegado a representar un estilo de vida, Google presenta el sueño: “moon.google.com”, ¡Google en la Luna!

“En honor del primer desembarco en la Luna, el 20 de julio de 1969, hemos añadido algunas imágenes NASA al interfaz de Google Maps para dejar que todo el mundo pueda visitar nuestro vecino estelar. Buena navegación lunar”.

Los movimientos de Google, “pseudo-monopolistas”, por métodos y perspectivas, han recaído inmediatamente sobre la competencia: hoy Google se está transformando en un gigante que ocupa todos los espacios de mercado; cada vez nuevos servicios estrangulan las sociedades hi-tech más pequeñas, a las que les cuesta fichar ingenieros y técnicos, con el riesgo añadido de ver imitados sus nuevos productos.

El lanzamiento sin interrupción de nuevos servicios y la financiación desde dentro a potenciales *spin-off* de los dependientes, bloquea de hecho el mercado en términos de innovación tecnológica. ¿Quién financiaría hoy un proyecto Web, con la posibilidad que en pocos días sea Google a lanzarlo?

Google ha sabido representar, a los ojos de los observadores, pero también de los simples usuarios de Internet, una especie de paladín del progreso: empezando por el motor de búsqueda, concebido para ser entendido rápida y sencillamente por los usuarios, se han multiplicado las propuestas de servicios e ideas.

Con la opción Open Source la economía relacional puesta en acto por Google se hace “visión del mundo”, recibida incluso como un desarrollo sostenible, un capitalismo bueno que difunde la abundancia, una solución económicamente ética a las búsquedas de los individuos.

III. Google Open Source: teoría y prácticas.

Open no es Free

Free Software⁴⁸ (Software Libre) y Open Source⁴⁹ (Código Abierto) son sintagmas que se utilizan a menudo como sinónimos para indicar códigos o porciones de códigos informáticos; sin embargo, por mucho que a menudo describan objetos idénticos, reflejan perspectivas radicalmente distintas. Free Software es un término que nace a principios de los años Ochenta por iniciativa de Richard Stallman: se refiere a la libertad del usuario para utilizar y mejorar el software. Más en detalle, se puede resumir en cuatro libertades fundamentales:

- 1) Libertad de ejecutar el programa, para cualquier objetivo.
- 2) Libertad de modificar el programa según las propias necesidades (para que esta libertad tenga algún efecto en la práctica es preciso tener acceso al código fuente del programa, ya que aportar modificaciones a un programa sin disponer del código fuente es extremadamente difícil).
- 3) Libertad para distribuir copias del programa gratuitamente o bajo compensación.
- 4) Libertad para distribuir versiones modificadas del programa, y permitir de esta forma que la comunidad pueda disfrutar de las mejoras aportadas.

Por otra parte, la expresión Open Source nace a finales de los años Noventa, fundamentalmente por iniciativa de Bruce Perens y Eric S. Raymond, que en 1998 fundan la *Open Source Initiative*⁵⁰ (OSI); se refiere a la *Open Source Definition*, que a su vez procede de las *Debian Free Software Guidelines*, es decir una serie de 10 puntos prácticos que definen cuáles criterios legales debe reunir una licencia para que se la considere efectivamente “libre”: o, usando el nuevo término, Open Source.

Es evidente pues que por un lado el Free Software pone el acento en la libertad: como subrayado en la definición, “el Software libre es una cuestión de libertad, no de precio”⁵¹. Por otra parte, el Open Source se ocupa exclusivamente de definir, en una perspectiva totalmente interna a las lógicas del mercado, cuáles son las modalidades mejores para difundir un producto según criterios open, es decir abiertos. El Free Software tiene un sentido que va mucho más allá del mercado, aún sin excluirlo a priori: el Open Source existe, como han explicitado sus promotores, para adaptar un modelo preexistente (el “free” en el sentido de “libre”) al mercado.

⁴⁸ La filosofía empresarial del Free Software: <http://www.gnu.org/philosophy/free-sw.html>

⁴⁹ La definición de Open Source: <http://www.opensource.org/docs/definition.php>

⁵⁰ el sitio de la Open Source: <http://www.opensource.org/>

⁵¹ Incluso en artículos recientes, de gran nivel, es preciso notar esta confusión entre Free Software y Open Source, casi el segundo fuese una mejora del primero en perfecta sintonía y continuidad. El movimiento fundado por Stallman contiene, sin duda alguna, muchas limitaciones concretas, no solamente en las relaciones con el mercado sino también en la asunción de posturas demasiado ideológicas; sin embargo no nos parece correcto trazar líneas evolutivas simples cuando en realidad el panorama es muy fragmentado y complejo. Véase, por ejemplo “Economia delle Reti Open Source: storia e dinamiche del movimento del software libero”, http://www.pluto.it/files/journal/pj0601/economia_reti.html

Los hackers de Stanford

Ya desde hace algunos años el software Open Source se considera comúnmente confiable, capaz de prestaciones elevadas y con costes sostenibles; en práctica hoy en día se le considera a menudo mejor frente a los software propietarios, precisamente porque permite aumentar la confianza de un producto en virtud de una metodología distinta de desarrollo, abierta y sometida a un amplio mecanismo de revisión.

Gracias a la realización de software estables y funcionales como browser, programas para oficina, editor y completos sistemas operativos (GNU/Linux), casi todo usuario se ha dado cuenta o ha oído hablar de la existencia de los programas llamados genéricamente copyleft u Open Source.

Open Source era el término adecuado, y necesario, para sustituir el sintagma Free Software. De hecho en inglés la palabra Free tiene el doble significado de Libre y Gratuito: esta ambigüedad lingüística no despertaba demasiado interés desde el punto de vista económico. La sustitución con Open fue una fórmula estratégicamente ganadora para mantener las características de cooperación libre sin renunciar a la posibilidad de un uso más estrictamente comercial de los software.

En realidad en esa época se estaba produciendo un cambio radical en el orden de las comunidades digitales espontáneas, o sea aquellas comunidades con las que se relacionan todos los que ofrecen una definición positiva de hacking.

Estas agregaciones eran (y aún lo son) extremadamente complejas. Nos referimos a un intersticio cultural fluido en el que se forman y colaboran estudiantes, profesores, investigadores, libres profesionales, policías y criminales, programadores pagados por sociedades de desarrollo, apasionados y muchas más tipologías de hacker.

El movimiento del Free Software estaba empezando a tener una confrontación cerrada con la economía de mercado. La batalla de la *Free Software Foundation* se centraba en la difusión de la licencia GPL (*General Public License*), creada por el fundador de la FSF Richard Stallman⁵²; esta licencia vincula el artefacto en modo “viral” a las cuatro libertades más arriba nombradas. Sustancialmente cada modificación aportada al código bajo licencia GPL tiene que mantener la misma licencia de uso, imposibilitando de esta forma el cierre del código: este mecanismo se conoce como “permiso de autor” (*copyleft*, juego de palabras con copyright). Surgían y se difundían entonces las primeras distribuciones del sistema operativo GNU/Linux.

La conmixión entre el método de desarrollo libre y la net-economy iba a determinar en los años sucesivos al 2000 la explosión de los productos Open Source y el desencadenarse del debate político acerca de la patentabilidad del software, el copyright y la gestión ético-política de todo lo que actualmente se define como obra de ingenio humano.

⁵² la Free Software Foundation en español: <http://www.gnu.org/home.es.html>

La empresa Google, aunque no sea una productora directa de software, no se ha quedado al margen de la sacudida Open Source: como otras empresas dinámicas e innovadoras, Google ha cooptado sus metodologías y las ha puesto al servicio de su “misión”.

La contigüidad entre Google y Open Source es espacial y temporal: en 1998 en Stanford, precisamente mientras Brin y Page pulían su primera versión de su motor de búsqueda, emergían algunos importantes proyectos Free Software; recordemos por ejemplo SND y Protegè que, en ámbitos muy distintos, el audio y la Web semántico, iban a resultar de gran éxito en el escenario digital.

En Stanford, la cultura hacker, de la que procede en última análisis el Open Source, se respira como aire familiar: no es pues un azar que nuestro dúo que se había formado en aquellos años haya manifestado siempre una determinada predilección por el desarrollo sobre la plataforma GNU/Linux.

Si existen diferencias sustanciales entre Open Source y Free Software, hay también elementos en común y continuidad de perspectivas. En breve hablaremos pues, por sencillez y corrección, de “metodologías y prácticas abiertas”, de Open Source, para indicar el fenómeno que interseca Free Software, Open Source y competición de mercado en el mundo del ITT:

La primera característica de una comunidad Open (y en este sentido también Free) es la de poner en práctica un método de trabajo abierto a la colaboración de todo el mundo, o sea capaz de aceptar sugerencias e interacciones espontáneas procedentes de toda tipología de sujeto implicado en la construcción del artefacto informático: programador, traductor, o incluso simple usuario. Este procedimiento ha sido definido en la jerga hacker método a “bazar” y su aplicación sobre amplia escala se debe al desarrollo del Kernel Linux en los primeros años Noventa, un proyecto nacido por iniciativa de Linus Torvalds y base de todas las distribuciones GNU/Linux⁵³.

La nueva técnica cooperativa propuesta por el undergrond digital ha vuelto del revés la ley de Brooks⁵⁴ que reglamentaba las comunidades de desarrollo de los proyectos informáticos hasta aquel entonces. Según la ley de Brooks, en la medida en la que crecen las complejidades aumentan de forma exponencial los errores y por lo tanto un proyecto al que contribuyen millares de desarrolladores tendería a ser un revoltijo de código inestable y lleno de bug. En cambio, mediante la revelación de los códigos fuentes, la circulación libre en Internet de la documentación, la cooperación y el *feedback* espontáneo de un número cada vez más elevado de sujetos en juego, las comunidades libres han demostrado que es posible una enorme mejora en la construcción de artefactos digitales, tanto desde el punto de vista del resultado, como del proceso. Los software realizados de este modo se publican bajo una licencia GPL, que como hemos visto difunde de manera viral materiales copyleft.

La licencia GPL no prevé por lo tanto restricciones desde el punto de vista comercial, sin embargo, como el término Free Software resultaba demasiado radical en poner las libertades en el primer lugar, así también la GPL ha sido sustituida por licencias

⁵³ A este propósito véase: Eric S. Raymond, <http://biblioweb.sindominio.net/telematica/catedral.html>

⁵⁴ La ley de Brooks según Eric S. Raymond <http://biblioweb.sindominio.net/telematica/catedral.html>

edulcoradas con respecto al contenido ético y político que el movimiento originario quería expresar. Este es el caso de licencias como la BSD (Berkeley Software Distribution), que no comporta ninguna restricción con respecto al cierre de los códigos y por ello inhibe el efecto viral, porque partes del códigos no libre pueden ser integradas en el código libre. Entonces, una creación libre puede convertirse en propietaria. O la MPL (Mozilla Public Licence) y otras licencias desarrollada exclusivamente para los nuevos productos Open Source.

La economía de mercado se convierte así en desarrollo sostenible y la Community de los desarrolladores en el núcleo de una verdadera y auténtica Open Society⁵⁵, la quimérica Sociedad Abierta. Este imaginario está determinado no solamente por el adhesión moral que suscita la práctica de un desarrollo comunitario, sino sobre todo por la calidad superior de las aplicaciones, en aparente contradicción con la gratuidad de las competencias puestas en juego.

La era del Open Source economy: competencia y bondad de ánimo

El ingreso del Open Source en el mercado es, según algunos observadores, una de las consecuencias de la llamada <<convergencia tecnológica>>, uno de los eslóganes que se ha convertido ya casi en un paradigma de la era informacional: el acercamiento y la sinergia de variadas tecnologías, anteriormente consideradas ajenas, estudiadas y desarrolladas en ámbitos separados.

Ante estas transformaciones a menudo demasiado rápidas, la creación de estándares abiertos ha creado un umbral en la “guerra de todos contra todos” del llamado “mercado libre”: “*cooperate on standards, compete on solutions*”, es el lema de la IBM, una de las principales empresas implicadas. Si incluso Big Blue decide cooperar, significa que el juego vale la pena...

Para muchas empresas el Open Source representa de hecho una de las pocas posibilidades para contrarrestar monopolios y oligopolios ya consolidados, para rehuir las dinámicas de competición clásicas sin inversiones enormes, para limitar los costes de desarrollo y reducir pues el “precio” de sus propios servicios.

Las empresas conocen desde hace tiempo y aprecian el valor de una dinámica reticular de desarrollo y alianzas: se sabe que el valor de una Red es proporcional al cuadrado de las personas/nudos que conecta⁵⁶. Una Red cada vez más amplia comporta beneficios exponencialmente mayores.

⁵⁵ Para una profundización en este tema, remitimos a Ippolita, www.ippolita.net/onf/

⁵⁶ Se trata de una ley matemática formulada a finales de los años Setenta por Robert Metcalfe, estudiante en la Harvard University y luego fundador de la sociedad 3Com además de ser pionero en el networking (e inventor del protocolo ethernet, aún hoy fundamental para las Redes Intranet). He aquí la traducción ad hoc del fragmento: “El valor de un network crece exponencialmente al número de los ordenadores conectados al network mismo. Por lo tanto, cada ordenador que se añade al network por un lado utiliza los recursos (o las informaciones o las competencias) del network y por otro aporta nuevos recursos (o nuevas informaciones o competencias) al network, aumentando así el valor intrínseco”. De este principio general sigue que: 1) el número de relaciones posibles, o meta-relaciones, o conexiones dentro de un network crece exponencialmente al número de ordenadores que se conectan (de ahí el valor estratégico de los link dentro de una Red); 2) el valor de una comunidad crece exponencialmente al crecimiento de los usuarios registrados en esa comunidad (de ahí el valor estratégico de las comunidades virtuales). Véase http://es.wikipedia.org/wiki/Ley_de_Metcalfe

El Open Source parece ofrecer algunas garantías relevantes en el desarrollo de redes con alto valor añadido: por un lado permite que el software siga siendo de alguna forma un bien <<público>> (adopta el desarrollo abierto y se vale de comunidad de soporte); por otro mantiene muy baratos los costes de paso de un sistema a otro (los llamados *switching cost*), en especial modo el paso de los modelos propietarios a los abiertos, sobre todo en el caso de los sistemas *legacy*, es decir obsoletos. Los mayores coste en la adopción de nuevas tecnologías se deben a la formación de los usuarios, y no a la inversión en la compra de la tecnología misma; con mayor razón si se trata de excelentes software libres por un coste irrisorio.

Pero el resultado más importante, cuyo precio es difícilmente calculable, es la creación de una imagen completamente nueva: para la propia empresa y los propios productos.

El éxito de las lógicas del software libre ha producido distintos intentos de aplicación de las mismas prácticas en otros sectores. Inevitablemente se han acuñado nuevas expresiones y se han formulados predicciones a menudo exageradas: Open Laws, Open Science, incluso Open Society parecen al alcance de la mano. De hecho hoy la idea de sociedad Open Source se ha convertido casi en un paradigma de la nueva era dirigida a la búsqueda de un instrumento común para una perspectiva política “posible”. Con sociedad Open Source se hace referencia a una sociedad cuyo código sea abierto y a cuya mejora pueda participar todo el mundo libremente. Puesta en estos términos no se puede no compartirla. Pero, sorprende la ligereza con la que se intercambia un término procedente de un recorrido particular, técnico e informático, para generalizarlo, aplicándolo a teorías filosóficas, económicas, sociológicas, sin considerar las posibilidades de modificar aquel concepto o cuando menos utilizarlo con las debidas aclaraciones.

Open Source, de hecho, en el campo del software en el que ha surgido el término, significa también competencia, competiciones para hacerse con los cerebros mejores al menor precio, financiaciones de capitales de riesgo, operaciones millonarias de compra y venta, un gran business orientado a una forma más “democrática” y suave de capitalismo. Esta dinámica mira no ya a someter la fuerza de trabajo, sino a cooptarla en la realización de la misión de la empresa, que se identifica cada vez más con la realización de los propios deseos particulares⁵⁷.

Entre las muchas sociedades que hoy en día explotan esta onda para conseguir las más distintas ventajas también está Google que, por otra parte, es bueno por lema: por ello, *Don't be evil*, usa el software libre, es gratuito, es mejor que el propietario, los desarrolladores están orgullosos de colaborar. La visita a Googleplex ha mostrado claramente como en Mountain View esta estrategia de penetración en las vidas de las

⁵⁷ En un contexto atrasado como el italiano, en el que las empresas, incluso en el sector del IT, continúan aplicando una obsoletas lógicas fordistas de producción de masa, sin valorizar minimamente las potencialidades de los individuos, podría aparecer despreciativo el ataque frente al capitalismo de la abundancia de Google. Pero como no todo el mundo cree en el desarrollo sostenible y en el consumismo responsable promovidos por una hábil campaña de propaganda tercermundista, así no podemos avalar ninguna forma de explotación de los individuos, ni siquiera la de Googleplex, sutil e incluso agradable para los trabajadores convertidos ya en creativos entusiastas. No se trata de rigidez dogmática, sino de un despliegue sin fin del capital, no importa si suave o ferozmente represivo. Richard Sennett lo ha demostrado con lucidez y un amplio examen histórico en *La corrosión del carácter : las consecuencias personales del trabajo en el nuevo capitalismo*, Anagrama, Barcelona, 2000. Por otro lado, el consumismo extremo del capitalismo de la abundancia podría constituir el primer paso hacia un nuevo fascismo, como muestra la lúcida visión de J.G. Ballard en *Milenio negro*, Minotauro, Barcelona, 2004.

personas se haya pulido hasta el máximo grado: trabajadores gratificados, animados a ser creativos y felices producen mejor que los trabajadores frustrados y oprimidos.

Seducir los hackers: autonomía, dinero fácil, instrumentos gratuitos

La explotación del Open Source por parte de Google alcanza su cumbre alrededor del 2005, cuando su imagen se ve empañada por las acciones de la competencia y por vicisitudes legales no del todo claras⁵⁸.

Por mucho que el proyecto estuviese enraizado de forma notoria en la cultura informática y en la práctica de la excelencia académica, no era suficiente usar el sistema operativo GNU/Linux para que el datacenter de Google funcionara: hacía falta explicitar la confianza hacia el Open Source con una iniciativa fuerte, que reclamara la atención en el magma de las redes de producción libre.

Ya no bastaba ofrecer el soporte para el idioma h4x0r, el lenguaje de los “verdaderos” hackers, (o por el Klingoniano de Star Trek), para cautivar las simpatías de los desarrolladores. Además, la actitud elitista de cerebros universitarios empezaba a cansar a los inversores.

Arrogancia y culto meritocrático de derivación académica, aunque siempre con óptimos resultados, tienen poca respuesta de los inversores que quieren sustanciales dividendos. Era inevitable que se cerrase la fase en la que los dos podían permitirse cotizar en bolsa el título suponiendo acciones por un valor de 2.718.281.828 dólares, un número que recalca la constante matemática “e” (la base de la función logaritmo natural); o también salidas alegres como el anuncio en agosto de 2005, cuando declaran que iban a vender 14.159.265 acciones para recoger 4 billones de dólares líquidos, sin mencionar nada a los inversores del uso que iban a hacer de aquel dinero.

Para sostener concretamente el propio deseo de invertir en investigación, para demostrar cómo mediante una semejante estrategia no solo se puede competir, sino que incluso ser los mejores en el mercado, hacía falta dar un paso estratégico dirigido no tanto a los usuarios “normales”, sino más bien a los jóvenes cerebros, al futuro, a la innovación; traducido operativamente esto significa crear comunidades, ofrecer instrumentos de desarrollo, sellar acuerdos con otras sociedades del sector. O sea, cortejar el mundo del Open Source.

Google comienza a invertir en la creación de comunidades en octubre 2005, cuando financia con 350.000 dólares el Oregon State University y la Portland State University. Sucesivamente se pone en marcha en pompa magna el programa de Junio 2005: *Summer of Code*, es decir “el verano del código” promovido directamente en una página del sitio de Google, y ahora localizable en <http://code.google.com/summerofcode05.html>.

El estilo comunicativo queda bien claro: ofrecer mejores oportunidades. Todo programador que ya hubiese creado un proyecto de Open Source nuevo o hubiese aportado una mejora digna de nota a un proyecto ya existente durante el verano, iba a recibir un premio de 4500 dólares. La operación estaba evidentemente enfocada a presentarse como una explícita declaración de amor hacia el Open Source, o sea re-

⁵⁸ Veáse cap. I

marcar que el Open Source era el terreno estratégico sobre el que cultivar la innovación. En segundo lugar, miraba a atraer las simpatías de los jóvenes desarrolladores mediante una operación de soporte financiero concreto para su trabajo. Finalmente, intentaba crear una verdadera y auténtica comunidad en el estilo “abierto”, cuando menos relativamente a los patrocinadores.

Los programadores premiados, en su mayoría estudiantes, han sido más de cuatrocientos; por lo general han aportado modificaciones e introducido novedades sobre proyectos ya existentes, en lugar de dar a conocer sus nuevos software, añadiendo características a programas como Apache, Fedora, Gaim, el mismo Google, Inkscape, Jabber, KDE, Mozilla, OpenOffice.org, Pitón, Samba, Gnome, Mono, Ubuntu. Una buena ganancia para todos, especialmente para las sociedades que están detrás de estos proyectos: entre las principales, recordemos IBM, RedHat, LinSpire, Novell, Mozilla.com, Sun, HP, Ubuntu⁵⁹.

Algunos de estos proyectos, junto con los desarrollados en el notorio 20% de tiempo libre por los mismos dependientes de Google, han permitido alcanzar el segundo objetivo en el recorrido de acercamiento al Open Source: ofrecer instrumentos de desarrollo. Ya desde 2002 Google ofrecía herramientas de desarrollo que podían descargarse gratuitamente del sitio code.google.com. Ahora esta página contiene los proyectos propietarios creados por los equipos de desarrollo de Google junto con los proyectos ganadores de Summer of Code que, de alguna manera, están relacionados con sus servicios.

La sección “Code” del sitio propone algunos proyectos dedicados a los creadores de software en los más variados lenguajes de programación (Java, C++, Python, etc.). Facilitar herramientas de desarrollo es un elemento esencial para cualquiera que desee permitir la creación de software y comunidad, porque se invierte precisamente sobre los medios de trabajos necesarios para su creación. Los proyectos ideados por los programadores de Google como herramientas de desarrollo se conocen con el nombre de Google API, librerías propietarias que permiten el interfaz y la utilización de los principales servicios del coloso Mountain View.

Una librería es un conjunto de funciones compartidas: se trata de porciones de código completadas que facilitan herramientas a otros programa que necesiten de funciones simplificadas. Un ejemplo muy intuitivo lo constituyen las librerías gráficas GTK, QT y FLTK, que implementan los elementos estándares de una aplicación visual (botones, menú, iconos...), simplificando el trabajo de los programadores: estos últimos, apoyándose a la librería escogida, escribirán solamente las funciones reales del programa. De hecho, serán las propias librerías las que dibujen los botones, que gestionen los clic del ratón, que diseñen las sombras y mucha parte de lo que estamos acostumbrados a ver como usuarios. El tiempo y las capacidades de los programadores así se optimizan; dado que quién escribe código difícilmente será un entusiasta de la creación de botoncitos para las propias aplicaciones, las librerías gráficas revisten un papel esencial de ligazón entre los varios proyectos. Por un lado, las aplicaciones aparecerán relativamente homogéneas desde el punto de vista gráfico; por el otro, los programadores podrán centrarse en su trabajo, sin perder demasiado tiempo en la implementación de las interfaces.

Existen comunidades de desarrollo que se ocupan de las librerías para ofrecer herramientas genéricas y transversales para la solución de problemas complejos

⁵⁹ *Summer of code* fue reeditado en abril de 2006: <http://code.google.com/soc/>; el éxito está garantizado.

(conexiones de Red, comunicación entre las aplicaciones, gestión del texto, imágenes, compresiones). Exactamente de la misma forma en la que se escribe un software con el objetivo de alcanzar más usuarios posibles, así se confecciona una librería para que llegue al mayor número de desarrolladores.

Las librerías, por lo tanto, permiten a los programadores la creación de sus propios software partiendo de un set de elementos compartidos, verdaderos y auténticos estándar *de facto* de la programación. Apoyarse a las librerías significa implementar el trabajo valiéndose de una base de partida muy amplia y compleja, que utiliza de forma óptima el código ya disponible y estratifica las competencias. Las librerías tienen, pues, un valor estratégico tanto en la dinámica de la cooperación espontánea del FS como en la economía relacional del OS:

Las librerías de Google, o sea Google API, están publicadas bajo licencias propietarias, es decir ocultan el mecanismo de su funcionamiento al programador. Pero no es todo: incorporan también un especial dispositivo de control, el desarrollador que descarga gratuitamente las librerías tiene que autenticarse mediante un código. Este sistema le permite a Google delinear de forma pervasiva todos los movimientos y las modificaciones que proceden del uso de sus API. Los programadores que usan estas librerías tienen la oportunidad de insertar la búsqueda de Google en su propio sitio o de conocer en tiempo real el propio PageRank. Además, pueden realizar software capaces de gestionar campañas publicitarias mediante AdWords, generar mapas dinámicos de sus datos con la interfaz de Google Maps o incluso implementar client VoIP para la telefonía on-line compatibles con Gtalk. En suma, pueden desarrollar los servicios de Google como mejor creen, en el lenguaje de programación que prefieran, bajo la atenta supervisión de Mountain View.

La enorme difusión de los servicios de Google se conjuga con la posibilidad de personalización en los mínimos detalles: de hecho, mediante la escritura de apropiados documentos XML⁶⁰, es posible crear “puentes” de conexión entre los distintos servicios de Google; por ejemplo programando pieza por pieza la *homepage* de Google como si fuese una verdadera aplicación, volviéndola, así, totalmente adecuada a las propias necesidades. Algo muy parecido puede hacerse con Google Earth: es posible construir navegaciones especiales en 3D sobre las fotos de satélite, subrayando gráficamente en los ordenadores de los usuarios zonas geográficas, edificios, datos climáticos, ecc.

Todas esas herramientas predisuestas para quién sabe escribir código (al menos en un lenguaje) son esenciales para hallar nuevas combinaciones o simplemente para usar lo que Google hace público, cuando menos en parte, dentro de sus propias aplicaciones⁶¹.

⁶⁰ XML, eXtensible Markup Language, es un lenguaje extensible realizado para poder utilizar de forma simple los documentos estructurados, pensado para la Web y para superar las limitaciones de HTML (Hyper Text Markup Language), pero con posibilidad de ser utilizado también en ambientes diferentes. Desarrollado a partir de W3C, World Wide Web Consortium, XML es un subconjunto de SGML (Standard Generalized Markup Language), un estándar internacional que define las reglas para escribir markup language; intencionalmente no incluye algunas funcionalidades complejas de SGML que difícilmente pueden ser implementadas en la Web. El primer borrador de XML se remonta a noviembre de 1996, el actual es consultable en la dirección <http://www.w3.org/TR/1998/REC-xml-19980210>

⁶¹ Sorprende el hecho de que el motor de búsqueda, desde el punto de vista de la interacción con las API, esté limitado a mil solicitudes diarias. Esto significa que los software que se basan en las API, podrán realizar como mucho mil búsquedas diarias. Probablemente las limitaciones disminuirán con el tiempo, porque la capacidad de Google crece sin descanso; sin embargo, Google impone las limitaciones sin ofrecer ninguna explicación.

Incluso existe un portal, googleearthhacks.com, en el que se pueden encontrar muchas trampas y hack para utilizar en las formas más impensables este servicio, cruzando los mapas del satélite con cualquier otra base de datos.

Todas las posibilidad que las API de Google facilitan implican el respeto de dos reglas especiales: el registro y la licencia. Para activar funciones de Google API es preciso solicitar una clave, es decir un código de acceso y comunicar exactamente donde se las quiere emplear. Las API se activan solo después de haberse registrado. El segundo punto es la licencia. Ya que estas API no tienen una licencia copyleft, se pueden utilizar solo dentro de ciertas delimitaciones: por ejemplo, es preciso tener una cuenta Google, porque el acumulo de informaciones no se para nunca; además, los mapas son de propiedad exclusiva de Google (o de terceras partes) y no pueden ser modificadas bajo ningún concepto. Naturalmente, para usos comerciales es necesario estipular un contrato. El código de activación permite que Google mantenga siempre el control total sobre los nuevos programas generados sobre las API: puede bloquear a su gusto las aplicaciones o simplemente controlar tanto la manera en la que se accede a los servicios, como el uso que de ellos se hace. Todo esto es posible porque el código fuente no es público y libre y por lo tanto resulta imposible comprender el funcionamiento interno de las librerías.

Además de dejar desarrollar gratuitamente y de monitorear el desarrollo de los propios servicios, otra razón por la que Google está creando comunidad con esta extraña fórmula que podríamos definir pseudo-Open es la de conseguir una base de datos sobre la que realizar investigación y venta de estadísticas.

Hospedar gratuitamente los proyectos de los programadores individuales significa conseguir su confianza. Permitir a cualquiera de realizar búsquedas en la base de datos de los proyectos hospedados activa una sólida cadena de usuarios. Una semejante incubadora de jóvenes talentos garantiza además, la disponibilidad de material humano altamente motivado y cuya formación, es decir el coste principal en el sector de la IT, ya se ha cumplido de forma autónoma y en toda conformidad al estilo de la empresa.

La oferta de herramientas de desarrollo es un mecanismo de *talent-scouting* conocido desde hace tiempo y en particular es el caballo de batalla de algunas sólidas empresas como la Va Software Corporation, que pone a disposición gratuita para el mundo del Open Source, ordenadores extremadamente potentes y banda ancha, espacio disco y asistencia a la portada de todo el mundo. Dos paraísos digitales pueden presumir de una fama mundial y de un número de proyectos hospedados superior a toda competencia: sourceforge.net y freshmeat.net, ambos de propiedad de la Va Software. Estos dos portales tienen una resonancia tal que incluso los proyectos más pequeños, tras aparecer en las primeras páginas, llegan tranquilamente a contar con centenares de visitas diarias. Todos los proyectos en el seno de code.google.com tienen sus correspondiente página en freshmeat.net y/o en sourceforge.net.

De esta forma las aplicaciones pueden a la vez gozar de la visibilidad de Google y de todos los servicios puestos a disposición por el coloso Va Software: mailing list, ordenadores dedicados a la solución de errores de programación (*debug*), a sistemas de control de las versiones, de las revisiones y de los cambios del código (por ejemplo, *cvs*, *Concurrent Versioning System*), forum de discusión, etc.

Es fácil imaginar como, con una base de datos utilizada gratuitamente por millares de *coder*, la Va Software puede garantizar un óptimo servicio de *business to business* para las empresas relacionadas al mundo Open Source y no exclusivamente. Un *data mining* de especial interés para negocios millonarios. Entre los patrocinadores y los anunciantes de sourceforge.net encontramos Red Hat, Microsoft y muchos más.

Hay muchas formas para poner en contacto los desarrolladores con las empresas Open Source. En Italia, SUN Microsystem ofrece la posibilidad de publicar el propio currículum sobre un mapa de Google (utilizando las Google API), mediante el portal javaopenbusiness.it. Los mismos desarrolladores son los que señalan su propio perfil, creando así un mapa de las competencias Open Source en Italia a través de las herramientas que SUN y Google han puesto a disposición.

Google puede contar pues con la implementación prácticamente gratuita de los propios productos por parte de centenares de usuarios; a todo esto se añade la inversión mirada en competiciones como *Summer of Code*, festivales dedicados a la promoción y desarrollo de los propios servicios y, por último, aunque no menos importante, sistemas de reclutamiento excepcionalmente dinámicos. Entre éstos, también se sitúa el video-reclutamiento, directamente en las páginas de video.google.com, con entrevistas a dependientes entusiastas y a Sergey Brin en persona, todos concordes en ilustrar los privilegios del trabajo en Mountain View⁶².

Ambientes híbridos entre universidad y empresa

Puestas estas premisas, el acercamiento de Google al Open Source aparece bastante estratégico e interesado, por mucho que se haya originado gracias a un sentir común respecto a las dinámicas cooperativas típicas de las comunidades de desarrollo Free Software, surgida en el humus académico. La estrategia de la acumulación señalada en precedencia está manos a la obra también en este ámbito: de hecho Google se comporta como una especie de agujero negro que usa códigos abiertos, o incluso favorece la redacción de esos códigos y los atrae, para luego introducirlos en su propio circuito. Por ejemplo, ninguna de las modificaciones que los programadores de Google han aportado a las herramientas abiertas usadas se ha hecho pública jamás. En especial modo su Google Web Server (GWS) es una versión modificada de una versión de Apache, el servidor web Open Source más difundido en la Red. Esto significa sin duda explotar las potencialidades y las realizaciones del método de desarrollo abierto, sin compartir, sin embargo, las propias implementaciones y mejoras.

Un factor de importancia primaria a propósito de las relaciones con el mundo Open Source es que Google nace a Stanford, una universidad conocida por su capacidad de generar *start-up* agresivas y competitivas basándose en investigaciones de perfil elevado. Por mucho que Stanford fuera, y siga siendo, un ambiente favorable al desarrollo de proyectos Open Source, su unión de doble filo con el capital de riesgo hace que sea difícil, más bien, imposible seguir adelante en la vía de la excelencia académica una vez se salga del campus.

⁶² El video está disponible en la dirección <http://video.google.com/videoplay?docid=8618166999532839788>

Un breve esbozo sobre la investigación académica americana será necesario para comprender los orígenes de Google, entre el Open Source y la investigación orientada al provecho. A nivel más general hay que subrayar el carácter centralizador de la universidad estadounidense a propósito de creación intelectual: todos los proyectos desarrollados en campo académico tienden a ser copyright de la universidad que ha hospedado el equipo de investigación. Stanford no es una excepción: por otra parte, en los EE.UU. las academias están relacionadas históricamente con el mundo de los negocios, y, a menudo, son auténticas empresas. Las patentes universitarias sobre las invenciones de los investigadores generan royalties muy sustanciales; confieren, además, prestigio a los centros de investigación y a los estudiantes/investigadores/empresarios.

Las universidades son ambientes híbridos, entre lo público y lo privado. En los EE.UU. hasta 2002, cuando menos en teoría, los lugares de investigación públicos no podían patentar sus invenciones; lo mismo vale para los laboratorios privados pero financiados con fondos públicos (por lo tanto a menudo también las universidades). El pago de aranceles obstaculiza la libre circulación de los saberes en la investigación científica, la posibilidad de reproducir, comprobar o falsificar los resultados experimentales. Esto en base al *Experimental Use Defense*, “protección del uso experimental”, un principio que autoriza el uso gratuito de tecnología patentadas en el ámbito de la investigación, introducido en 1813, y abolido en 2002 con una sentencia a favor del investigador John Madey. Madey ha citado en juicio la Duke University, para la que trabajaba, porque usaba una maquinaria patentada por él para investigaciones laser sobre electrones libres. La Corte ha considerado que el *Experimental Use Defense* hubiese surgido para proteger el científico dedicado a la investigación libre y desinteresada, pero evidentemente en las universidades esta actividad ya no es tan inocente e, incluso en el caso en el que no sea directamente comercial, puede ser considerado como un “negocio lícito” (*legitimate business*), ya que conlleva financiaciones y precisa de fuerza de trabajo y de personal en formación (los estudiantes). Cae así toda distinción entre la investigación privada y la pública.⁶³

Naturalmente, todos los proyectos nacidos en Stanford son patentados por parte de la universidad, y esta conmixión entre incentivo a los proyectos Open Source por un lado, y patentabilidad salvaje por el otro, no beneficia el ideal, ni mucho menos la práctica de la “investigación” en sí, tan ostentada como punto de orgullo y de fuerza de Google

La cuestión de la patente se hace aún más interesante si recordamos que el éxito de Google se basa en un algoritmo ideado por Larry Page, en colaboración con Sergey Brin, cuando aún eran investigadores en la facultad de Ciencias Informáticas en Stanford. El algoritmo que ha revolucionado la indexación de la Red es por lo tanto propiedad de Stanford, sometido a patente regular. Vamos a descubrir en el detalle como funciona este prodigio, como consigue ofrecer resultados en tiempos más rápido que cualquier otro de la competencia, casi que de verdad pudiera dar a cada usuario “lo que desea”.

⁶³ Laser: <http://www.e-laser.org/htm/new.asp?idNews=223> y <http://www.e-laser.org/htm/new.asp?idNews=151>; volumen del Grupo Laser, cit, www.ippolita.net/editoria/2/laser_-_il_sapere_liberato.pdf; fuentes directas sobre el caso Madey: www.law.nyu.edu/journals/lawreview/issues/vol79/no4/NYU406.pdf.

IV. Algoritmos ¡qué pasión!

El crecimiento vertiginoso de Google no ha mermado su fama de motor rápido y eficiente, fiable y completo: todos hemos oído decir que “si no está en Google, no existe!” y que Google es más veloz”. En la base de este éxito, además de los elementos que hemos ido analizando hasta ahora, está el algoritmo de PageRank, citado al comienzo de este texto, que guía el spider de Google al descubrimiento de las redes. Veamos en detalle de qué se trata y como funciona.

Algoritmos y vida real

Un algoritmo⁶⁴ es un método resolutivo aplicado a un problema, un procedimiento compuesto por pasos simples a ejecutar en secuencia para obtener un determinado resultado. Un algoritmo que favorece la solución de un problema se dice correcto, y si la solución se obtiene en tiempos breves se dice eficiente. Existen muchos tipos distintos de algoritmos, empleados en los campos más diferentes de las ciencias; pero no se trata de procedimientos abstractos pertenecientes a una exigua minoría de estudiosos, sino de prácticas que influyen nuestras vidas cotidianas mucho más de lo que nos parece a primera vista.

Por ejemplo, las técnicas para grabar un programa televisivo utilizan algoritmos, y también los métodos para ordenar una baraja de naipes o para planificar las paradas de un viaje particularmente largo. En un tiempo relativamente previsible, realizando una serie de pasos sencillos y repetibles de manera idéntica, elegimos más o menos implícitamente los algoritmos adecuados a la solución que estamos buscando. Sencillo significa sobre todo especificado de manera no ambigua, inmediatamente evidente para quién aplicará el algoritmo, o sea su ejecutor. En este sentido, una receta es un algoritmo: “haced hervir tres litros de agua en una olla, salad y echad quinientos gramos de arroz, colad después de doce minutos, añadid especias a voluntad” es una descripción de pasos sencillos y no ambiguos, si el destinatario de la receta está en condición de discernir pasos como “salar”, o “añadid especias a voluntad”.

Los algoritmos no son necesariamente métodos para hallar una solución en el menor tiempo posible. De echo los hay que se ocupan de obtener soluciones aceptables sin preocuparse del factor tiempo; otros permiten conseguir un resultado en el menor numero de pasos, o su prioridad es el ahorro de recursos.⁶⁵

Nos importa subrayar aquí, más allá de cualquiera profundización especializada, la naturaleza práctica, aplicativa de los algoritmos. Los algoritmos nos involucran a todos pues son prácticas concretas para alcanzar un determinado objetivo. En el campo informático se utilizan para solucionar problemas recurrentes en la programación del software, en los proyectos de redes y en

⁶⁴ “Algoritmo: conjunto de reglas o directivas aptas para ofrecer una respuesta específica a uno o más datos de input”. Para una primera introducción, <http://es.wikipedia.org/wiki/Algoritmo>. El término algoritmo deriva del nombre de “al-Khwarizmi”, importante matemático árabe del siglo noveno. Muhammad ibn Musa al-Khwarizmi introdujo el uso de los números árabes en la matemática: su obra “(Libro) de al-Khwarizmi sobre los números indios” fue traducida en latín como “Algorismi de numero Indorum”. En este texto toda Europa aprendió a utilizar el sistema de notación decimal posicional hoy todavía en vigor; la numeración romana fue progresivamente abandonada. Los procedimientos que permitían efectuar cálculos en notación decimal fueron denominadas “Algorismi” o “Algoritmi” y más tarde el mismo término fue aplicado en general a los procedimientos de cálculo necesarios para obtener un determinado resultado.

⁶⁵ El mejor método para llegar a París puede ser salir con un vuelo directo desde el aeropuerto más cercano a nuestra ciudad, o coger el primer vuelo en salida del mismo aeropuerto, bajar a la estación de llegada, retomar el primer avión, bajar y nuevamente volver a partir, continuando hasta llegar a París. Es bastante cierto que los dos métodos nos permitirán llegar a destino: con el primer método llegaremos a la meta en el menor tiempo posible y probablemente con el mayor ahorro de dinero; el segundo, al revés, nos permitirá apreciar los principales aeropuertos de las capitales europeas en los distintos periodos del año. Algoritmos diferentes describen las distintas posibilidades.

la construcción de aparatos hardware. A lo largo de los últimos años, sobre todo por la creciente importancia de los modelos reticulares de análisis e interpretación de la realidad, muchos investigadores han enfocado sus estudios en las metodologías de construcción y de recorrido de las redes y de los datos que constituyen su materia viva. La economía de la búsqueda de la que habla John Battelle⁶⁶ se hace posible por el perfeccionamiento de algoritmos para la búsqueda de informaciones, estudiados para aumentar la posibilidad de encontrar y compartir datos de forma cada vez más eficiente, veloz, fiable y segura. El caso más conocido por el gran público es el fenómeno del peer-to-peer: en lugar de crear enormes bancos de datos a los que acceder para encontrar video, audio, textos, software y todo tipo de informaciones, se desarrollan en continuación algoritmos cada vez más optimizados para facilitar la creación de redes altamente descentradas, en las que cada usuario puede ponerse en contacto directamente con otros usuarios y llevar a cabo provechosos intercambios.⁶⁷

La estrategia de la objetividad

El crecimiento exponencial de la calidad y cantidad de banda de nuestros ordenadores, junto con la constante disminución de los costes, nos ha permitido navegar en Internet mejor, para más tiempo y más rápidamente. Hace nada más veinte años los módems con pocos **baudes** (número de símbolos transmitidos por segundo) eran un lujo para pocos, mientras que ahora también en Italia la fibra óptica, por la que viajan millones de byte por segundo, es una tecnología accesible. Hace diez años hacían falta notables conocimientos informáticos para crear contenidos aptos para las redes digitales; ahora al revés la mayor facilidad de publicación de contenidos en la Web la omnipresencia del correo electrónico, las mejoras de los sistemas de escritura colectiva online, como blog, wiki, portales, mailing list y en paralelo la bajada de los costes de grabación y mantenimiento de los dominios y de los espacios Internet favorecen la transformación de los usuarios: desde simples usuarios de informaciones puestas a disposición por especialistas de las IT, progresivamente se transforman en creadores de informaciones.

La mejora de la conectividad procede por tanto a la par con el crecimiento exponencial de los datos volcados en Red y, como ya hemos anotado, implica la urgente necesidad de cada vez mejores herramientas de búsqueda. La urgencia difusa a todo nivel de servicios de búsqueda atrae forzosamente el interés de sociólogos, informáticos, ergónomos, designer, estudiosos de la comunicación en general. Por otro lado, el diluvio informacional de las redes globales no es una banal “puesta en Red” de las sociedades tal y como las conocemos, sino un fenómeno extremadamente complejo, que exige no fútiles interpretaciones. Creemos por tanto que dicho compromiso teórico y práctico, no se pueda delegar a los especialistas, sino que debe ser el resultado de una elaboración colectiva.

De hecho, si por un lado la construcción de redes autogestionadas puede ser una oportunidad para ampliar y poner en contacto entre sí zonas autónomas, por el otro el control social encuentra en las tecnologías de la información un formidable instrumento de represión.

⁶⁶ Véase el capítulo II.

⁶⁷ El peer-to-peer. Generalmente con peer-to-peer (o P2P) se entiende una Red de ordenadores o cualquier Red que no posea cliente o servidor fijos, sino un número de nodos equivalentes (“peer” significa “pares”) que hacen funciones bien sea de cliente que de servidor hacia otros nodos de la Red. Este modelo de Red es la antítesis de la arquitectura cliente-servidor, donde existe una relación jerárquica entre el proveedor del servicio (servidor) y el receptor (cliente). Por medio de esta configuración distribuida, cualquier nodo está en condición de poner en marcha y completar una transacción. Los nodos equivalentes pueden diferir de la configuración local, en velocidad de elaboración, en amplitud de banda y cantidad de datos memorizados. El ejemplo clásico de P2P es la Red para la puesta en común de archivos (fuente: <http://www2.autistici.org/inventa/doku.php?id=glossario> no-copyright etc.)

La realización de este segundo escenario, del que el caso Echelon⁶⁸ es tan solo la manifestación más clamorosa, aparece ciertamente más posible a la luz del constante aumento del número de individuos quienes producen informaciones, en abierta contradicción con la disminución continua de los proveedores de herramientas de búsqueda. El acceso a las informaciones producidas por un número cada vez mayor de individuos está dirigido por un pugnado de monopolistas que reducen una delicada cuestión social y política a una competición de marketing sin ahorro de golpes, en la que la adopción de un mejor algoritmo resulta ser el elemento ganador.

Un algoritmo de búsqueda es un instrumento técnico que activa un mecanismo de marketing extremadamente sutil: el usuario se fía de que los resultados no sean filtrados y correspondan a las preferencias de navegación que genera la comunidad de usuarios. En sustancia, se difunde un mecanismo de confianza en la objetividad de la técnica (específicamente, el procedimiento algorítmico que genera el resultado de la interrogación) que se considera “buena” en cuanto que no influenciada por las idiosincrasias y preferencias de individuos humanos. Las máquinas “buenas”, hijas de una ciencia “objetiva” y de una investigación “desinteresada”, no manipularán los resultados, no nos contarán mentiras porque no pueden mentir y de todas formas no tendrían ningún interés en hacerlo. Bien distinta es la realidad y esta creencia se revela una hipótesis demagógica, detrás de la cual las máquinas del marketing y del control acumulan fabulosas ganancias.

El caso de Google es el ejemplo meridiano de esta “estrategia de la objetividad” ligada a la técnica: el motor de búsqueda “bueno por definición” explota y graba por entero y de forma continua los comportamientos de los usuarios quienes utilizan sus servicios, con el fin de identificar sus costumbres e insertar en sus actividades (navegación, correo, gestión file, etc.) anuncios personalizados, contextuales, ligeros, omnipresentes, y posiblemente en condición de generar *feedback*, de modo que los usuarios ofrezcan informaciones útiles para los vendedores y hasta lleguen a mejorar de por sí “las sugerencias publicitarias”, expresando sus preferencias. La continua petición de la opinión de los usuarios, además de gratificar a las personas que se sienten participes de una extensa “democracia electrónica”, es de hecho la forma más simple y eficaz para conseguir preciosas informaciones comerciales acerca de los gustos de los consumidores. Son las preferencias (y la inconciencia) de los usuarios las que hacen triunfar un motor de búsqueda sobre los demás, pues un sitio muy visitado puede modificar sus contenidos sobre la base de “sugerencias” comerciales y consecuentemente activar virtuosos movimientos económicos.

Desde un punto de vista exquisitamente informático, a los motores de búsqueda les pertenece la gestión de cuatro elementos: la búsqueda de datos en la Red (spider), la memorización de las informaciones en archivos (bases de datos), un válido algoritmo para ordenar los datos según las búsquedas realizadas (interrogación) y por último el desarrollo de una interfaz capaz de satisfacer al usuario; cada uno de los tres primeros aspectos está cuidado por un tipo específico de algoritmo: búsqueda, memorización/archivación, interrogación.

La potencia de Google, como la de Yahoo! y de otros gigantes de la búsqueda en Red se basa por tanto:

8. “spider”, o sea un software para extraer contenidos de las redes;
9. hard-disk de enorme capacidad para memorizar los datos en soportes fiables y capaces, para evitar toda pérdida accidental;

⁶⁸ Echelon es el nombre de un sistema de control y espionaje electrónico global puesto en práctica por Estados Unidos. Su justificación oficial es hoy la lucha en contra del terrorismo: comunicaciones telefónicas, mail y toda comunicación digital se intercepta y analiza con fines políticos y económicos. Para un análisis en profundidad vease: Duncan Campbell, *Il mondo sotto sorveglianza-Echelon e lo spionaggio elettronico globale*, Eléuthera, Milano, 2003: <http://home.hiwaay.net~pspoole/echres.html> recoge muchísimos recursos disponibles.

10. un sistema rápido para encontrar (y ordenar) los resultados de una interrogación sobre la base del valor de ranking de las páginas;
11. una interfaz usuario vía Web (y no solo: Google Desktop y Google Earth, por ejemplo, son programas que hay que instalar en la máquina del usuario) para responder a las peticiones acerca de estas informaciones.

Spider, bases de datos y búsquedas

El spider es un **aplicativo** que, en la mayoría de los casos, se desarrolla en los laboratorios de investigación de los mismos motores de búsqueda. Su finalidad es navegar saltando de un link a otro en las páginas Web recogiendo informaciones: formatos de documentos, palabras clave, autores de las páginas, links ulteriores, etc. Al final de sus exploraciones el software-spider lo entrega todo a la base de datos que archivará las informaciones. Además el spider debe ocuparse de captar las variaciones de cada sitio y entonces programar una visita sucesiva para almacenar nuevos datos. Particularmente el spider de Google gestiona dos tipologías de chequeo de los sitios: una mensual en profundidad, *Deep-crawl*, y una diaria de puesta al día, *Fresh-crawl*. De esta forma la base de datos de Google emplea algún día para poner al día las varias indexaciones y difundir los nuevos resultados en todos los *datacenter*. Este lapso de tiempo es conocido como Google dance (danza de Google): los resultados de las búsquedas difieren hasta sensiblemente, pues se refieren a índices diferentes. Desde 2003 Google modificó sus metodologías de catalogación y actualización, limitando drásticamente los efectos de la “danza” y diluyéndolos en el tiempo: de hecho ahora los resultados de las búsquedas varían de manera dinámica y continua sin ningún desbarajuste periódico. En realidad los resultados de las búsquedas difieren también sobre la base de las anteriores navegaciones de los usuarios, que se archivan y utilizan para “mejorar”, en el sentido de “simplificar”, la recogida de las informaciones.⁶⁹

La secuencia de elecciones que la aplicación lleva a cabo para indexar un sitio es la verdadera potencia del algoritmo de Google. Mientras que el algoritmo de base de PageRank está depositado bajo patente por Stanford, por tanto público, estos ulteriores pasos algorítmicos no son publicados ni por Google, ni por ninguno de los motores de búsqueda actualmente en uso; de la misma forma no son públicos los procesos de recuperación en la base de datos.

En ámbito informático, una base de datos (*database*) es, sustancialmente, un archivo digital; en su forma más sencilla -y actualmente más difusa- se puede representar bajo la forma de una o más tablas en relación entre sí que presentan valores en entrada y en salida: se habla entonces de database relacional. Como todo archivo, una base de datos está organizada según precisas reglas de almacenaje, extracción y mejora continua de los mismos datos (recuperación de datos corruptos, corrección de entradas duplicadas, constante reingenierización de los procesos de captación de los datos, etc.).

Los informáticos estudian desde decenas de años las metodologías de búsqueda, introducción y mejora de los datos en database, experimentando lenguajes de programación y diferentes acercamientos (jerárquico, reticular, relacional, a objetos, etc.). La fase del proyecto de una base de datos es un componente crucial del proceso de desarrollo de un sistema informativo complejo como Google, pues de esa depende en esencia su funcionalidad. Para conseguir una rápida extracción de los datos y, en general, una gestión eficiente, es por tanto fundamental la correcta individuación de las finalidades del database y, en el caso de los database relacionales, de las tablas, a definir a través de sus campos y de las relaciones que las unen. Naturalmente es necesario adoptar inevitables aproximaciones en los pasos entre las lenguas naturales, analógicas, y los datos introducidos, digitales, que son evidentemente discretos: cero o uno, hay información o no la hay, no existen medias medidas. El punto doliente es la confidencialidad de

⁶⁹ Véase el cap. V

estas metodologías: como sucede en todos los proyectos propietarios de desarrollo, a diferencia de aquellos libres, es harto difícil saber cuáles instrumentos y cuáles algoritmos han sido utilizados.

A través de los textos redactados por los centros de investigación y por las universidades se pueden encontrar las escasas informaciones hechas públicas por los proyectos propietarios. En estos textos, se encuentran informaciones útiles para comprender la estructura de los ordenadores y la gestión de los datos por parte de los motores de búsqueda. Para hacernos una idea de la potencia de cálculo actualmente disponible, se describen ordenadores capaces de convertir direcciones Internet en secuencias univocas de byte útiles como índices para los database en 0,5 microsegundos y capaces de ejecutar 9000 spider en contemporánea; bajando a lo concreto, se trata de sistemas en condición de analizar y almacenar cerca de 50 millones de nuevas páginas al día.⁷⁰

El último elemento algorítmico escondido detrás de la “sencilla” fachada de Google es el dispositivo de búsqueda, o sea ese sistema que, dada una interrogación, es capaz de encontrar los resultados congruos, ordenarlos por importancia y ranking, y enviarlo a la interfaz.

Algunas universidades y laboratorios han decidido publicar sus investigaciones en tal sentido, particularmente las soluciones halladas y los diferentes métodos utilizados para optimizar la velocidad de acceso a las informaciones, la complejidad de la ordenación y la selección de los parámetros de input más interesantes.

Los motores de búsqueda, de hecho, deben ser capaces de ofrecer resultados óptimos casi instantáneamente, ofreciendo contemporáneamente un abanico de posibilidades de elección de los más amplios. Google representa sin duda el estado del arte de los motores de búsqueda: tales extraordinarios resultados solo se pueden obtener gracias a la implementación de filtros oportunos, como veremos en profundidad en el próximo capítulo.

Por ahora es importante saber que el mejor éxito está asegurado por el correcto equilibrio entre potencia de cálculo y calidad del algoritmo de búsqueda. Buscar una información entre los terabyte (1 TB=1000 Gigabyte) o petabyte (1 PT = 1000 TB = 1 millón de Gygabyte) necesita el empleo de extraordinarios soportes de almacenamiento y formidables sistemas de indexación, con la tarea de identificar tanto en qué punto del enorme archivo se encuentra la información como calcular el tiempo necesario para extraerla..

La Red reboza de leyendas no siempre verificadas ni verificables a propósito de la capacidad computacional de Google, también porque la empresa revela pocos particulares de su infraestructura tecnológica. Algunas fuentes hablan de cientos de miles de ordenadores en conexión entre ellos en miles de gigantescos *clúster* que implementan apropiadas distribuciones GNU/Linux; otras de superordenadores, dispositivos cuya estética recuerda escenarios de ciencia-ficción: silos enormes super refrigerados en los que uno o más brazos mecánicos mueven a la máxima velocidad miles de discos duros. Las dos soluciones son plausibles, junto con otras aún, sin estar necesariamente en contradicción. Lo que es cierto es que la extrema escalabilidad de las máquinas de Google permite prestaciones excepcionales, puesto que el sistema es “abierto” a toda mejora.

De la brand-identity a la interfaz participativa

Búsqueda, almacenamiento y descubrimiento de los datos son procedimientos extremadamente complejos y necesitan, para ser entendidas en profundidad, conocimientos y profundizaciones que se salen de las intenciones de este texto. Más adelante veremos algunos detalles de su funcionamiento. Una atención especial hay que dedicarla a la interfaz, pues mientras que las actuaciones del algoritmo y la arquitectura de la base de datos son elementos estructurales del

⁷⁰ Vease por ejemplo la documentación publicada por IBM Almaden Research Center: <http://www.almaden.ibm.com/webfountain/publications/>

motor de búsqueda que permanecen visibles al usuario, la interfaz está proyectada y gestionada como imagen del mismo Google.

Por interfaz entendemos antes todo el “*blank box*”,⁷¹ ese espacio vacío en el que se introducen las peticiones o “intenciones de búsqueda” en el marco de la página universal de Google, estudiada para resultar acogedora, comfortable, familiar.

Se trata de un perfil dicho universal pues viene declinado en distintos idiomas (al momento, más de 104 entre idiomas y dialectos personalizables para más de 113 países) y en cada una de estas presenta un modelo de interacción que queda invariado y unifica los comportamientos de búsqueda en un esquema único y homogéneo.

En la página de Google nos encontramos delante de una interfaz lineal compuesta por elementos esenciales, cada uno con una función bien determinada y universalmente reconocida. Está en condición de aceptar indicaciones de búsqueda de distinta naturaleza y complejidad, desde la introducción de sencillas palabras clave (es. “*ippolita*”) a palabras compuestas, que hay que poner entre comillas (es. “*comunidad escribiente*”), hasta a búsquedas miradas: por ejemplo, las búsquedas pueden ser limitadas a un sitio particular, o a un idioma específico, a páginas procedentes solo de un determinado dominio, o a documentos de un cierto formato, y más, según el nivel de precisión que se desea obtener. O sea, se trata de un logrado ejemplo de interfaz que alcanza el no fácil objetivo de asociar un significado positivo al espacio en blanco de la página. La interfaz se presenta sin inútiles decoraciones, casi vacía, o mejor rellena por un único elemento “vacío”: el *blank box*, que tranquiliza al usuario y tiende a inducir comportamientos activos, en lugar de provocar despiste debido a la ausencia de puntos de referencia, o al revés por la presencia de input visuales exuberantes. Se evita así la confusión generada por páginas demasiado llenas, casi fuesen afectadas por una especie de *horror vacui*, por una ansiedad comunicativa que, en el intento de atraer al usuario con miles de banner, efectos gráficos, juegos, muchas veces consigue el efecto contrario.

No existe una navegación verdadera en la página de Google: las distintas partes de la página tienen un significado funcional, sirven para acceder a los servicios, no para conducir al usuario en un recorrido; su utilización pone en marcha comportamientos que muy rápidamente devienen parte de una rutina de búsqueda, tanto que aparecen instintivos al poco tiempo. La interfaz del motor de búsqueda está estudiada de manera que la utilización, la dinámica de funcionamiento y las expectativas del usuario, un usuario genérico, se repitan; es más, también después de haber almacenado y digerido las “personalizaciones” del mismo usuario, las prácticas de búsqueda permanecen sustancialmente idénticas, al punto que podemos hablar de un instrumento “universal”.

La disposición de textos e imágenes es lineal y aprovecha elementos gráficos recurrentes, por ejemplo el empleo de colores elementales; las imágenes utilizadas son cualitativamente homogéneas. El estilo de construcción de la interfaz es sobrio, casi esencial y, en contraste con el design de tendencia de la *brand-identity* (y de la *corporate-identity*)⁷² orientado a la búsqueda de una especificidad estética, hace referencia a calidades perceptivas, elementales pero muy eficaces en su sencillez.

De esta inmediata identificación visual procede una facilidad de uso netamente superior con respecto a los motores de búsqueda de la competencia. El nivel de ergonomía alcanzado es sorprendente: Google no necesita presentarse como un concentrado de servicios a través de su

⁷¹ En la jerga informática, “black box” se refiere a una caja negra que recibe input, los elabora de manera no transparente para el usuario y restituye un output. El concepto de “blank box” reproduce esta forma, pero de manera implícita, por esto ambigua, pues, aún tratándose de un “espacio limpio, vacío” (blank) está cargado de significados y funciones de búsqueda altamente diversificadas.

⁷² La ideación de una nueva imagen para un producto o un servicio se conoce como brand identity; cuando se refiere a una sociedad específica se habla de corporate identity. Hoy en día el concepto de brand ha superado ampliamente la idea de “sello de distinción”, llegando a configurarse como una “marca” que necesita expansión psíquica, territorial, comercial. Para una primera introducción, véase <http://es.wikipedia.org/wiki/Marca>

interfaz; en otras palabras su arquitectura visual es aquella típica de los portales multiservicio. Las interfaces de los distintos servicios son autónomas y sustancialmente independientes, todas caracterizadas por la presencia de la “*blank box*” y no linkadas entre ellas de forma directa. Por ejemplo, se necesitan muchos pasos no intuitivos para alcanzar el servicio de `code.google.com`, pensado por técnicos de variado nivel, partiendo del servicio base de búsqueda de las imágenes (`images.google.com`), dirigido a un público más genérico: es necesario bajar “en profundidad” en el sitio `google.com` y saber qué buscar. A pesar de esta fragmentación, estamos todos en condición de reconocer la Red de servicios ofrecida por Google; además los usuarios pueden utilizar de manera combinada e integrada los recursos informativos puestos a disposición, sea para aquellos que se limitan a la simple utilización del browser, sea para los Google-adictos, los Google-totally-adicted⁷³ quienes se vuelcan entusiasmados en cada nuevo servicio.

Esta des-territorialización de los servicios genera un peculiar mecanismo relacional: los usuarios no conocen las nuevas secciones directamente desde Google, sino de la Red informal de los usuarios, desde otros sitios en los cuales los visitantes exponen sus gustos y discuten de sus costumbres. El mismo usuario localiza automáticamente la amplia gama de servicios ofrecidos por Google en el momento en que se interesa por un nuevo servicio: por ejemplo, con respecto a la zona geográfica, se presenta de inmediato la interfaz lingüística apropiada al usuario. Por otro lado, es fácil enfocar las tipologías de usuarios a los que se dirige un servicio, y evaluar el nivel de preparación técnica requerido, o el nivel de afinidad con otros usuarios. El mecanismo del pasa palabra se vuelve parecido a un “PageRank relacional”.

En una primera aproximación, existen una dimensión relacional local, en la que el pasa palabra se produce entre amigos y conocidos, y una dimensión relacional tipológica, en la cual un cierto tipo de usuarios, identificables en base a parámetros estadísticos (edad, sexo, empleo, etc.) utiliza un determinado servicio y pone en marcha la economía relacional.

Los diez problemas relativos a la usabilidad de los sitios Web, discutidos por Jacob Nielsen⁷⁴, entre los más conocidos estudiosos de interfaces-usuario, parecen no afectar lo más mínimo el sitio de Google que, no obstante esté escrito en lenguaje HTML totalmente fuera de estandar⁷⁵, consigue asegurar la total visibilidad en todos los browser, sean ellos gráficos o textuales.

La limpieza gráfica de las paginas está exaltada por una optima gestión visual de los aspectos comerciales. Ningún link publicitario en la homepage o en las páginas de documentación e información: la publicidad en Google se encuentra solo entre los resultados de las búsquedas, separada adrede de los resultados propuestos, pero no ajena a los argumentos buscados. Por tanto se puede decir que Google es capaz de expresar, al menos en la disposición escénica de sus interfaces, el correcto compromiso entre respeto de los usuarios y necesidad de retorno económico. La publicidad, fuente principal de ingresos de Google, se programa y realiza para que no sea nunca invasiva y para no distraer a los usuarios de su utilización de los servicios. Los link publicitarios están patrocinados de forma dinámica para seguir el recorrido de un usuario dentro del motor de búsqueda y por tanto, en segunda instancia, en los sitios de Internet.

⁷³ La Google-manía se extiende y genera nuevos lenguajes; para una panorámica, véase:

<http://en.wikipedia.org/wiki/Google> (search engine) : una lista de servicios y herramientas relacionados con Google.

http://en.wikipedia.org/wiki/List_of_Google_services_and_tools

⁷⁴ Jacob Nielsen, cincuentón informático danés, es una de las mayores autoridades en el campo de la usabilidad de la Web. Nielsen es conocido, entre otras, para sus críticas del exceso de gráfica y animaciones (por ejemplo Flash) que aquejan muchos sitios populares con perjuicio de la usabilidad, práctica dañina sobre todo para los discapacitados. Su último volumen publicado es Jacob Nielsen, Marie Thae, Homepage Usability, Apogeo, Milano, 2002. Los diez problemas de la usabilidad: <http://www.shinynews.it/usability/1005-errori.shtml>

⁷⁵ HTML (acrónimo para Hyper Text Mark-up Language) es un lenguaje utilizado para describir los documentos hipertextuales disponibles en la Web. No es un lenguaje de programación, sino un lenguaje de mark-up, o sea describe el contenido textual y no, de una página Web. Véase: <http://es.wikipedia.org/wiki/HTML>, y sobre todo el sitio de la W3 acerca del estándar HTML: <http://www.w3.org/MarkUp/>

Los link comerciales no son estáticos, se modifican y acompañan las búsquedas de los usuarios; esto es posible también gracias a los feed RSS (acrónimo de *RDF Site Summary*, o de *Really Simple Syndication*), uno de los formatos más utilizados para la distribución de contenidos Web, y en virtud de las diferentes fuentes informativas digitales (diarios, revistas, agencias de prensa, etc.) en condición de modificar dinámicamente la homepage de Google. De hecho Google pone su homepage a disposición de los usuarios registrados, haciéndola totalmente configurable gracias a la incorporación de feed RSS: es así posible configurar las previsiones del tiempo automáticas para las ciudades que se desea monitorear, o recorrer el archivo histórico de las búsquedas realizadas. Se pueden organizar los marcadores y los últimos mensajes de correo recibidos, y también tener controlados los file del propio ordenador sin solución de continuidad con respecto a los contenidos Web, gracias al aplicativo Google desktop.

El mecanismo de promoción publicitaria, los servicios y los sofisticados mecanismos de perfilación del usuario parecen constituir un conjunto al nivel estético y de los contenidos; por otra parte, los link patrocinados no serían en este sentido nada más que simples sugerencias, gráficamente compatibles y conceptualmente alineados con la operación de búsqueda que se está llevando a cabo. La economía de Google está altamente integrada con la interfaz, al punto de poder quedar excluida a nivel visual por quien no esté interesado, y por lo contrario aprovechada por quien encuentra interesantes los link y recorridos comerciales propuestos.

También Yahoo!⁷⁶ y muchos otros motores de búsqueda y portales ponen a disposición herramientas parecidas para la personalización de la propia homepage; pero la cantidad y calidad de las ofertas de Google, de momento, queda insuperada. Se trata de configuraciones bastante sencillas, que sin embargo, requieren una cierta costumbre con las interfaces Web y un poco de tiempo para ser realizadas. En ámbito Web el umbral de atención es notoriamente muy bajo, las páginas se visualizan y abandonan en tiempos muy rápidos, del orden de pocos segundos; por tanto un usuario que invierte muchos minutos, o decenas de minutos, lleva a cabo unas elecciones que revelan mucho de sí y de sus costumbres de consumidor. Estas informaciones, cuidadosamente almacenadas por la compañía de turno (Google o Yahoo! que sea) constituyen la verdadera riqueza producida por el mismo usuario, y son fundamentales para proponer bienes y servicios mirados por parte de las empresas anunciantes.

La personalización de las páginas hace que un sitio sea más amigable; el mismo sitio se configura como una herramienta personal en la que el usuario invierte tiempo, eligiendo colores, aspecto, contenidos. Un visitante habitual capaz de configurar su página de inicio viene cooptado y hecho participe de la construcción de la interfaz Web. Concederle al usuario el poder y el control sobre algunas páginas significa promocionarlo desde simple diana de campañas publicitarias a consumidor “inteligente” y desde luego es la mejor y más sutil manera de crear fidelidad promoviendo la interacción. Se entreven ambientes dotados de interfaces participativas y no excluyentes para recibir publicidad cada vez más personalizada, para entrar todos juntos en el dorado mundo de Google.

PageRank, o la autoridad absoluta de un mundo cerrado

El algoritmo que permite a Google asignar un valor a las páginas indexadas por el spider se conoce como PageRank.

Ya sabemos que el funcionamiento de PageRank está basado en la popularidad de una página Web, calculada sobre la base del número de sitios que hospedan al menos un link apuntando a ella. A paridad de número de link, dos páginas Web tendrán PageRank diferentes según la importancia de quienes las haya linkado: con este mecanismo se evalúa el aspecto “cualitativo”

⁷⁶ Vease por ejemplo MyYahoo!, my.yahoo.com

de los sitios. Los link presentes en las páginas Web más linkadas obtendrán importancia superior frente a los presentes en las páginas menos linkadas.

Hagamos un ejemplo concreto: frecuentemente, controlando las estadísticas de acceso referidas a un sitio, se encuentran cantidades de contactos procedentes de sitios pornográficos. Esto sucede porque Google atribuye un ranking en función de los accesos, que a su vez vienen visualizados en las páginas de estadísticas públicas. Existen por tanto programas que aprovechan la pervasividad de esta lógica de conexión y evaluación de los nodos de una Red para subir el propio rank; como ocurre con frecuencia, los primeros experimentadores son los sitios pornográficos (tal como ha sido para las galerías de imágenes en la Web, o para el comercio online).

En la práctica se utilizan algunos programas que se ocupan de buscar sitios con estadísticas de acceso públicas; entonces se produce un elevado número de peticiones, simulando visitas procedentes de un link falso presente en otro sitio, que en la mayoría de los casos es un sitio pornográfico. Este mecanismo de bombardeo multiplica el número de accesos al sitio en cuestión y en consecuencia las estadísticas muestran incrementos evidentes; de esta forma subirá sensiblemente el Google-Ranking del sitio y en última instancia también aquel del sitio pornográfico del que han llegado los link: en fin, una ganancia para todos, al menos al nivel de la visibilidad.

Este tipo de operación no es ilegal: nadie veta hacer peticiones a un sitio Internet; gracias a esta práctica los sitios a estadística pública obtienen un ranking más elevado. Además, dicho mecanismo nos permite ilustrar como también la magia tecnológica del ranking de Google, considerado objetivo y veraz, esté ligada a los “barrios bajos” de la Red, para nada dignos y a prácticas de linking más bien equivocadas.

Otras prácticas no ilegales que aprovechan la indexación de Google se conocen como SEO (*Search Engine Optimizacion*); se trata de un conjunto de actividades para mejorar el posicionamiento de un sitio Web en los resultados de búsqueda. La oferta de una posición de primer plano se concreta frecuentemente a través de e-mail de spam procedentes de direcciones improbables, evidentemente traducidas con programas automáticos, que prometen resultados vislumbrantes:

“Nosotros registramos Su sitio Internet en 910 motores de búsqueda, registro y catálogo Web. Nosotros llevamos Su sitio Internet en las primeras posiciones de Google y Yahoo! Inténtenlo! No hay ningún riesgo. Tan solo 249€ (coste único, sin suscripción) en lugar de 349€”. Por supuesto Google sigue reivindicando su transparencia: “nadie puede garantizar que vuestro sitio aparezca al primer puesto en los resultados de Google”⁷⁷

Desde el punto de vista matemático, una consecuencia del PageRank basado en el análisis de los link es la integridad de la base de datos; o sea, la determinación de un espacio circunscrito, aunque amplio, en el cual realizar las búsquedas. De hecho, si las páginas se evalúan y descubren solo a través de link esto quiere decir que no existen páginas no linkadas o islas de documentos desligados del resto de la Web: en sustancia en el mundo de Google existe siempre un recorrido que lleva de una página a cualquier otra presente en la base de datos, es decir en las redes indexadas.

Por tanto las búsquedas serán tendencialmente funcionales, evitando al máximo la posibilidad de link rotos (*broken link*) o de informaciones distintas de las anteriormente archivadas, presentes en la memoria oculta (*cache memory*). El problema es que de esta forma se inducen los usuarios a creer erróneamente que Internet sea un mundo cerrado, conexo, completo, sin carreteras poco

⁷⁷ Texto de un e-mail recibido en info@ippolita.net en mayo de 2005. La posición de Google en los SEO: <http://www.google.it/intl/it/webmaster/seo.html>. Para profundizaciones técnicas, aconsejamos la guía estratégica al posicionamiento en Google, <http://www.google.com/googlerank/>.

iluminadas o recorridos preferenciales, pues parecería que, dada una interrogación, siempre se llegue al resultado “correcto”.

Esto depende del hecho que la visión googoliana de Internet se genera por entero de los recorridos que el spider cumple rebotando de un link a otro. Si una página no se cita en ningún otro sitio, entonces no aparecerá nunca en ninguna interrogación realizada por un usuario, pues el spider nunca tuvo la posibilidad de encontrarla, pesarla y evaluarla. Pero esto no significa de ninguna de las maneras que no existan islas de datos, al revés.

Un ejemplo son los sitios dinámicos, en los cuales las funcionalidades ofrecidas están basadas en su totalidad en las elecciones del usuario. Uno de estos sitios es trenitalia.com: recopilando la ficha (*form*) el sitio es capaz de ofrecer en tiempo real los horarios de los trenes, las coincidencias, los recorridos más veloces para llegar a un destino. Google no es capaz de entender las peticiones de este form y por eso no indexa los horarios y los recorridos generados dinámicamente por trenitalia.com. Solo la intervención humana puede superar este escolio. La única solución propuesta por Google es la de incorporar en su interfaz herramientas de re-direccionamiento a los sitios de compañías aéreas o ferroviarias en el momento en que el usuario busca un recorrido, destino o llegada.

La integridad referencial propuesta por la base de datos de Google debe ser revisada, porque supone un mundo único para todos, cerrado y limitado. Al contrario, trazar un recorrido en una Red compleja quiere decir llevar a cabo una exploración que siempre determina resultados relativos y parciales.

El sueño de Google contenedor de toda Internet es una idea demagógica particularmente cómoda, útil para sostener la exhaustividad y fiabilidad de las informaciones disponibles, o sea todas las características que hacen de Google un “servicio único”, un dispensador de verdad. En las búsquedas de la vida cotidiana dicha clausura absoluta es muy útil, pues conduce rápidamente a un resultado; en realidad nos hace ilusión pensar que la libertad consiste en el hallazgo de una calidad total. Sabemos por contra que en un sistema reticular complejo no existen verdades absolutas, sino solo autoridades distribuidas según el recorrido que se quiera hacer, o tan solo en función del tiempo que estemos dispuestos a invertir en la búsqueda. La calidad depende totalmente de nuestra percepción subjetiva de la aceptabilidad del resultado. Las redes que podemos analizar, apreciar y vivir, son objetos complejos cuyos nodos y enlaces están en constante mutación. Como la tarea de aceptar un elaborado de navegación referido a una búsqueda depende en definitiva del usuario, resulta esencial ejercer la capacidad crítica, la conciencia de la subjetividad de nuestro punto de vista. Para generar el recorrido que nos interesa hace falta aceptar la hipótesis de la existencia de una Red cerrada y limitada, un mundo cerrado solo por nuestras exigencias personales, reconociendo que solo se trata de una localización subjetiva, ni absoluta ni constante en el tiempo. Explorar una Red implica la capacidad de dividir las redes en subredes de análisis y se corresponde con la creación de pequeños mundos localizados y temporales⁷⁸. En la práctica de la navegación cotidiana los enlaces casuales son de primaria importancia: la creación de enlaces nuevos e insólitos no puede ser prevista de ninguna manera por el análisis de los elementos de la Red sugerida por el Ranking de Google. Estos enlaces funcionan como “puertas dimensionales” y permiten la disminución y hasta la anulación de las distancias entre dos nodos de la Red.

[PageRank, o la moneda de la ciencia](#)

Además, el algoritmo de PageRank, diferentemente de la opinión común, no es un invento original de Google, sino que procede de los descubrimientos matemático-estadísticos de Andrej Andreevic Harkov, quien en los primeros años del siglo XX analizó fenómenos estadísticos en

⁷⁸ Para profundizar en el concepto de mundos cerrado localizados (LCW, Localized Closed World) vease el panfleto sobre Redes dirigido por Andrea Marchesini: <http://www2.autistici.org/bakunin/doc/reti/index.xml>

los sistemas cerrados, es decir esos sistemas en los que cada elemento es causa o efecto de otro elementos del mismo sistema⁷⁹.

Sergey Brin y Larry Page seguramente partieron de esta base teórica, pero las mejoras conseguidas no se han divulgado del todo, aparte de la patente depositada por Stanford.

El mejor ejemplo para aclarar la morfología de este algoritmo es el pasa palabra entre amigos y conocidos. En las relaciones interpersonales cuanto más se habla de un determinado evento, tanto más esto toma importancia y al mismo tiempo deviene parte de un patrimonio común. Si se limita la difusión de ese evento a un entorno reducido su popularidad será menor. Lo mismo vale para los trabajadores del espectáculo: cuanto más consiguen que se hable de ellos, mayor será su ranking, más se conocerán y más famosos podrán ser (por esta razón existen transmisiones auto-referenciales como “La Isla de los Famosos”...). Esta misma lógica Google la aplica a los datos online.

Google anuncia este método de forma muy convincente, propagando la imagen de Internet como una gran democracia, pues el algoritmo actúa como si los link fuesen votos referidos a los sitios: poco importa si se linka para hablar mal o bien: lo que importa es que se hable. La falsa verdad de esta “democracia global” procedente de un algoritmo es evidente para cualquiera: como si la democracia dependiera de la tecnología y no de las prácticas de los individuos.

El origen cultural de esta práctica, como ya comentado⁸⁰, deriva del sistema, extremadamente de élite, de la revisión de los pares (peer-review) por parte de los *referees* de las publicaciones científicas: así la contribución individual de cada investigador se inserta en una Red de relaciones, verificaciones y evaluaciones que permite la transmisión y el control de los resultados de la búsqueda. La democracia global de Google se configura como una aplicación del método científico de las publicaciones a la Red, gracias al algoritmo de PageRank, una especie de “*referee* tecnológico” para evaluar de forma objetiva las informaciones de la Web, contando con las preferencias expresadas por el “pueblo de los navegadores” a través de los link, y proponerlas en el orden correcto.

El paralelismo es muy estrecho: de una parte las publicaciones científicas ganan peso y autoridad en base a su colocación en el marco de su específico campo de investigación; dicha colocación se obtiene a través de las citas, o sea las referencias a la literatura. De este modo la investigación científica se garantiza una continuidad, pues cada artículo no nace en el vacío, sino que se pone como el “presente” del largo recorrido de la tradición científica. De forma parecida, el spider de Google interpreta los link de las páginas Web como “citas”, que aumentan el prestigio, es decir el ranking de esas páginas.

El elitismo científico, base del sentido de temeroso respeto que transmite la “ciencia”, curiosamente se fundamenta en la práctica de la publicación: hacer “público” no implica “accesibilidad” o comprensibilidad⁸¹. De hecho, “los descubrimientos de los científicos, teóricos o experimentales que sean, ni son ni pueden considerarse conocimiento científico hasta que se registren de forma permanente”, como sostenía en los años Setenta el sociólogo Robert Merton⁸². Quizás la afirmación sea demasiado perentoria (la ciencia antigua se traspasaba todo menos que públicamente: piénsese en la escuela pitagórica en Grecia, a la distinción entre escritos esotéricos y exotéricos, etc.) pero evidencia correctamente el carácter eminentemente

⁷⁹ Para una introducción acerca de las cadenas de Markov vease: http://en.wikipedia.org/wiki/Markov_chain

⁸⁰ Véase el Cáp. III

⁸¹ La impresión de que la ciencia sea demasiado difícil de comprender para cualquiera no sea un especialista, está socialmente radicada en todos aquellos que, a partir de su vida cotidiana, se sienten ajenos a ella. Las paRedes del trabajo técnico parecen inviolables. La común expresión inglesa “no es ciencia para cohetes” (it’s not rocket science), normalmente sarcástica observación hecha a alguien que padece insólitos problemas en el cumplimiento de tareas fáciles, solo es un ejemplo de la manifestación de reverencia pública hacia la intensidad intelectual de la ciencia y de su separación de las actividades comunes de todos los días. Véase a este propósito la actividad de CAE, Critical Art Ensemble, www.critical-art.net

⁸² Robert K. Merton, *Ciencia, tecnología y sociedad en la Inglaterra del siglo XVII*, Alianza Editorial, 1984.

público del conocimiento científico moderno. La comunicación no es por lo tanto un subproducto de la investigación, sino parte integrante de una forma del saber cuyos presupuestos son el carácter cumulativo y cooperativo. La ciencia, al menos desde el siglo XVI, por una parte está orientada a la consecución de resultados novedosos, que puedan consistir en un aumento del patrimonio del conocimiento, por otra parte asume como punto de partida los frutos de las investigaciones precedentes. Podemos esbozar una historia de la comunicación científica que evoluciona junto con los media destinados a soportarla: desde la intensa correspondencia epistolar entre científicos a la impresión periódica en revistas eruditas, hasta la comunicación digital. No es casual que los primeros nodos de Internet fueron centros de investigación académica, que necesitaban comunicar y compartir sus informaciones.

De todas formas la mutación del soporte no ha producido un cambio sustancial en el método de conexión típico de esta forma comunicativa, que sigue siendo el de las citas. Descritas como “moneda de la ciencia”, las citas serían una especie de tributo de los científicos a sus maestros e inspiradores. Más concretamente, ponen en contacto la investigación presentada con las ya desarrolladas por el mismo autor o por otros. Es razonable asumir que el número de citas recibidas por un determinado trabajo pueda representar una indicación de su importancia o al menos de su impacto en la comunidad científica. A lo largo de los años este sistema se ha vuelto materia de estudio específica: el análisis bibliométrico es una disciplina que se vale de técnicas matemáticas y estadísticas para analizar los modelos de distribución de la información, y en particular de las publicaciones. Actualmente la bibliometría, y particularmente su más conocido indicador, el *impact factor*⁸³, se utiliza comúnmente como criterio “objetivo” para evaluar la calidad del trabajo científico llevado a cabo por un investigador o una institución. Un gran archivo para el análisis bibliométrico ha sido colgado online en 1993 justo en Stanford, la cuna de Google. El proyecto SPIRES (Stanford Public Información Retrieval System)⁸⁴ nació en 1974 de la colección de notas bibliográficas acerca de artículos de física de las altas energías curada por la biblioteca universitaria de Stanford. Limitadamente al restringido campo de análisis (la física de las altas energías), SPIRES es un database completo y de acceso gratuito, que permite búsquedas complejas también de las citas, una palestra que Brin y Page han sabido aprovechar de la mejor manera para poner a punto el algoritmo de PageRank. Junto con el algoritmo hay algunos refinamientos que hacen de Google una verdadera herramienta de mediación global en la Web.

⁸³ Eugene Garfield, *The Impact Factor*, en “Current Contents”, n°37(25) 1994, pp. 3-8; <http://www.isinet.com/isi/hot/essays/journalcitationreports/7.html>

⁸⁴ Vease el artículo “L’evoluzione delle abitudini di citazione nella comunicazione scientifica primaria. Il caso della fisica delle alte energie”, Marco Fabbrichesi, Barbara Montolli; http://jekyll.comm.sissa.it/notizie/lettere02_01.htm

V. Además, otras funcionalidades maliciosas

Los filtros en los algoritmos: bases de datos preconfeccionadas y control sobre los usuarios

La teoría de los grafos⁸⁵ es la base matemática para la elaboración de todos los algoritmos en las redes, entre los que figura también PageRank. Esta rama de la matemática se ocupa, entre otras cosas, de estudiar los métodos para gestionar, crear, recorrer distintas tipologías de redes descriptibles mediante grafos y sus correspondientes medidas. La teoría de los grafos ha tenido un desarrollo especial a partir de la mitad del siglo XX con la introducción de los elaboradores electrónicos. Podemos imaginar geoméricamente un grafo como un conjunto de puntos en el espacio y de curvas continuas que conectan parejas de puntos sin intersectarse entre sí. En la teoría de los grafos, se dice *grafo* (no confundir con gráfico) una figura constituida por puntos, llamados *vértices o nudos*, y por líneas que los unen, llamadas *lados o aristas o arcos*⁸⁶.

Una Red es un tipo de grafo especial, en la que es posible asignar un valor diferente a cada arco (un “peso”), que se podrá utilizar para generar valores de recorrido distintos. Internet es un grafo y podemos definir de la misma forma el conjunto de las páginas Web; el sistema de búsqueda de Google se organiza a partir de este asunto.

Uno de los elementos fundamentales de los algoritmos en las redes es el factor tiempo en relación con el número de nudos examinados. El tiempo para completar un recorrido (o una búsqueda) que une un nudo de la Red a otro depende del número de los elementos de la Red y varía siempre entre un valor mínimo y un valor máximo, que pueden ser incluso muy diferentes entre ellos, según el algoritmo que se utiliza para generar el recorrido.

En la Red de las páginas Web cada página es un nudo del grafo global y cada link es un lado/arista/arco. Si consideramos la premisa sobre el factor tiempo, es evidente que los resultados propuestos por Google como respuesta a una pregunta cualquiera (técnicamente una query en la base de datos) no pueden proceder de una consultación de “todo” su hipotético Internet.

El spider de Google es perennemente ocupado en la ardua tarea de copiar Internet en su base de datos; sin embargo, no es creíble que a cada pregunta el motor de búsqueda consulte puntualmente toda la base de datos para encontrar los resultados más congruentes. El factor clave que consiente conseguir resultados casi inmediatos depende de una prudente serie de limitaciones en la selección general, es decir, concretamente, de la aplicación de filtros especiales. Planteada una pregunta al motor de búsqueda, un filtro garantiza la velocidad del resultado final mediante una serie de desvíos y de elecciones estudiadas adrede para limitar el intervalo (*range*) de los posibles bloques de datos a examinar.

De esta manera Google puede facilitar resultados a las preguntas en tiempos excepcionalmente rápidos. Sin embargo, si la búsqueda es tan rápida, resulta a la vez poco transparente, o sea no completamente coherente con los datos presentes globalmente en la porción de Red indexada. Dicho con otras palabras, los resultados de la búsqueda se conseguirán velozmente no solo

⁸⁵ La literatura sobre los grafos es muy extensa; para un primer acercamiento, veáse la homónima categoría de wikipedia, <http://es.wikipedia.org/wiki/Grafos>

⁸⁶ Más formalmente, dados un conjunto V de nudos y un conjunto E de arcos, un grafo G es un conjunto $G = (V, E)$.

gracias a la potencia de cálculo disponible, sino y sobre todo porque los filtros permiten reducir el pozo de informaciones del que sacar los datos.

La tarea de un filtro consiste en realizar selecciones drásticas en los nudos de la Red excluyéndolos (o valorizándolos) junto con otras eventuales conexiones asociadas. La aplicación de estos dispositivos tiene el objetivo de excluir (o incluir) bloques completos entre los resultados globales.

Todo esto es posible mediante el uso de bases de datos de búsquedas preconfeccionadas, para las que a preguntas estándares se facilitan respuestas estándares, pero también mediante la personalización del perfil del usuario basada en sus búsquedas anteriores, en base a su idioma, a su posición geográfica, etc. De esta manera, un usuario que realiza habitualmente búsquedas en italiano no preguntará la base de datos global de Google, sino solamente la parte en italiano, con un evidente ahorro de tiempo.

Sobre elevadas cantidades de datos es impensable emplear algoritmos transparentes, o sea que toquen todos los nudos de la Red; es preciso introducir manipulaciones, simplificaciones o reducciones de las posibilidades de análisis. Esto por razones técnicas, de cómputo matemático en sentido estricto, y también, obviamente, por razones económicas. Además, sin pecar de malicia injustificada, en un sistema basado en aproximaciones dictadas por filtros se puede fácilmente concebir la inserción de ulteriores filtros para añadir o desplazar en posiciones visibles los resultados comerciales, de pago, o de pura propaganda.

Sin embargo, hay que observar que desde el punto de vista de Google los filtros añadidos no están relacionados directamente con un interés económico, ya que no sirven para vender un producto, sino que están conectados al usuario, a sus costumbres y a sus intereses personales. En efecto, Google vende publicidad, no productos (o solo en mínima parte, por ejemplo los productos hardware como Google Mini y otros sistemas de indexación para empresas): su interés principal es por lo tanto conseguir datos añadidos en base a parámetros sobre los que poder realizar campañas publicitarias con precisión. La personalización de los resultados en base al destinatario es posible sobre la base de informaciones que el mismo Google facilita y que se recogen en la forma menos molesta posible. Por ejemplo, mailbox, blogs, hardisks virtuales, y otros servicios semejantes representan otros tantos bancos de datos útiles a trazar el perfil de los usuarios mucho más de lo que nos imaginamos.

Los servicios “extras” puestos a disposición por Google, más allá del simple motor de búsqueda, son pues útiles a la empresa para experimentar nuevas vías, pero también y sobre todo porque desarrollan un papel fundamental de “concentradores de informaciones personales” de los usuarios.

Un ejemplo clamoroso han sido las mailbox de GMail, una especie de hard disk virtual en Google (hasta ahora, más de 2GB de espacio online, en constante aumento), ofrecidas a través de un sistema de propagación basado en el mismo PageRank. En breve, cada nudo de la Red de Google (usuario) dispone de un peso (un cierto número de invitaciones) y puede usarlo para ofrecer el mismo servicio (link) a sus conocidos. Gracias a este método todo sujeto implicado está controlado desde el punto de vista del uso del servicio; a la vez, pone a disposición de Google informaciones esenciales sobre sus conocidos y amigos.

En un segundo tiempo, el mecanismo se propaga a la Red de las personas invitadas, que pueden realizar nuevas invitaciones: de esa manera se va creando un grafo de las relaciones humanas

entre los usuarios de enorme valor estratégico para la personalización de las inserciones publicitarias.

Si consideramos todas las informaciones que se pueden extraer del envío de e-mail (a quién, por qué razón, en qué idioma, en qué formatos, con cuáles palabras claves, con cuáles datos adjuntos, etc.), podemos imaginar que exista en el banco datos de Google, además de una copia forzosamente parcial pero significativa de Internet, también una copia parcial, y otro tanto significativa, de las relaciones humanas, emotivas, profesionales de los usuarios del servicio.

En teoría los filtros sirven simplemente para que la interrogación sea más rápida y más ajustada a las exigencias de los individuos; pero además también son absolutamente necesarios desde un punto de vista técnico. Sin embargo, su uso demuestra que para un sujeto de hecho hegemónico en el ámbito de los servicios de búsqueda, es extremadamente sencillo explotar sus peculiaridades para aprovechar los datos a su disposición en sentido comercial y lesivo de la privacy.

En suma: actualmente, la base de datos de Google, mediante algunas palabras claves, es capaz de ordenar los resultados de una búsqueda (*query*) de manera diversificada según el tipo de usuario que tiene que atender, es decir en base a los datos relativos a aquel usuario que obran en su poder. Lejos de ser resultados “objetivos”, los resultados de toda búsqueda son pues calibrados; más aún, el uso de los servicios de búsqueda refina las capacidades del sistema para “reconocer” el usuario y ofrecerle resultados adecuados.

El uso de cada servicio está subordinado a la aceptación de los reglamentos y excepciones por parte de los usuarios, y Google en toda ocasión se compromete a no divulgar informaciones reservadas y personales en general; pero es fácil pensar que Google pueda comercializar o explotar para los más variados fines los datos de los usuarios. Sin considerar que de momento la oportunidad (sería mejor decir: la certeza) que militares y fuerzas de policía de distinto tipo puedan solicitar y tener acceso a informaciones semejantes en pro de presuntas razones de “seguridad nacional”. Añadir filtros de búsqueda para personalizar los resultados es en concreto el escenario más sencillo de imaginar.

Los cookies de Google, galletas que dejan su huella

Dibujar el perfil⁸⁷ de los usuarios se basa siempre en un sistema de selección y reconocimiento. En la Red de Internet distinguimos entre métodos de personalización explícita e implícita. La explícita necesita de un procedimiento ad hoc de registro, que implica el envío a través de un formulario (*form*) de datos personales por parte del usuario. Las informaciones enviadas, archivadas en la base de datos, son analizadas mediante una serie de parámetros útiles para segmentar en grupos homogéneos la totalidad de los usuarios registrados (edad, sexo, profesión, intereses, etc.). La personalización implícita se realiza a través del trazado de usuarios anónimos durante sus visitas a un sitio, mediante dirección IP o a través de *cookies*, o sea galletas. Los *cookies* son pequeños file de texto utilizados por los sitios web para almacenar algunas informaciones en el ordenador del usuario. Cada vez que el usuario vuelve a estos sitios, el browser devuelve las informaciones guardadas en la “galleta”. El objetivo es automatizar algunas

⁸⁷ Se trata de un sector clave de la economía de la búsqueda, en continua expansión, que incluye conceptos como: *Clicstream análisis*, CRM (*Customer relationship management*), *Data mart*, *Data minino*, *Data warehouse*, OLAP (*Online Analytical Processing*). Para una introducción elemental sobre el argumento, véase: http://www.diodati.org/scritti/2002/g_stat/stat05.asp Es interesante notar que las técnicas de personalización de los usuarios proceden, en última análisis de las actividades de *profiling* criminal.

operaciones de autenticación (*login*), recordar eventuales operaciones en curso, pero sobre todo asociar el visitador a todas las informaciones memorizadas en sus anteriores visitas.

La mayoría de los sitios Internet que ofrecen servicios online usan los cookies y Google no constituye en absoluto ninguna excepción⁸⁸. Gracias a la combinación de cookies y filtros añadidos sobre los algoritmos es posible mantener un trazado de la navegación de una persona y acumular informaciones sobre su “huella”.

Pongamos un ejemplo: el individuo X posee un número de teléfono móvil registrado a su nombre, con el que llama a su familia, algunos amigos y a los compañeros de trabajo. Después de un cierto tiempo decide cancelar aquel número y utilizar otro no registrado a su nombre, pensando así estar protegiendo su *privacy*. Con el nuevo teléfono restablece su círculo de conocidos, contactando la propia familia, algunos amigos y los compañeros de trabajo. Aquella secuencia de “conexiones sociales” (familia, amigos, compañeros), resulta ser una secuencia única entre todas las llamadas del mundo, relacionada indisolublemente con el protagonista de este ejemplo. No es imposible formalizar una secuencia semejante con un grafo que representa los nudos y los arcos de una Red: los valores (los respectivos “pesos” de los arcos que conectan los nudos) podrían ser asignados asociando a la “cercanía” un grado proporcional de “amistad” con el punto de partida del análisis, es decir con nuestro hipotético individuo X.

La remoción de los *cookies* es por lo tanto una óptima práctica de defensa de la *privacy*, pero un razonamiento análogo al ejemplo anterior puede fácilmente ser modulado sobre el uso de los motores de búsqueda. Mediante los *cookies*, buscando informaciones sobre algunos específicos intervalos temáticos, es de hecho posible identificar grupos de personas, e incluso directamente individuos en base a su huella univoca dejada en la Red.

La misma huella univoca que marca nuestros movimientos, nuestros contactos sociales (o telefónicos), es única respecto a las preferencias, los gustos, las idiosincrasias, las pasiones que nos distinguen de los demás. Nuestras pasiones son, en este caso, los sitios que visitamos, y sobre todo, las búsquedas que realizamos durante nuestra navegación. Esta cantidad de informaciones que ofrecemos a cualquier motor de búsqueda hace que sea posible la reconstrucción de nuestra “huella”⁸⁹.

Como toda galleta que se respeta, también las de Internet tienen una fecha de caducidad. Los sitios Internet que consignan a nuestro browser un *cookies* están obligados a impostar una fecha de caducidad, es decir una fecha tras la que el browser puede proceder a la eliminación de los datos contenidos en el mismo *cookie*. No es inútil utilizar al máximo los *cookies*; ciertamente es digno de nota el hecho de que Google haya sabido explotar para sus ventajas un artificio técnico conocido por los desarrolladores de todo el mundo, POSIX (el estándar internacional que permite la Inter-operabilidad de los sistemas operativos Unix y Unixlike, que incluyen también GNU/Linux). En efecto, la fecha de caducidad de Google es el 2038, es decir casi la misma

⁸⁸ Para más información sobre cookies volátiles, veáse particularmente: <http://www.google-watch.org/cgi-bin/cookie.html> y <http://www.sgroogle.org/>. Google Watch explica la gestión de los cookies por parte de Google y métodos para defenderse de la personalización. Scroogle se propone anonimizar las búsquedas hacia google, de forma que el cookie eterno de Google no pueda consentir un trazado histórico sobre las búsquedas realizadas por el usuario.

⁸⁹ Sobre la temática, veáse la intervención “La *privacy* y los motores de búsqueda” de Claudio Agosti “vecna” – s0ftpj.org en ocasión del congreso “e-privacy – fra Trusted Computing e Data Retention”, que ha tenido lugar en Florencia el 19 y el 20 de mayo de 2006 <http://e-privacy.firenze.linux.it/>, disponible en la dirección http://e-privacy.firenze.linux.it/atti/e-privacy_2006_Agosti_Privacy_e_motori_di_ricerca.pdf

fecha máxima programable en un *cookies*, y esto conlleva que el browser de nuestros sistemas operativos no removerá nunca aquel *cookies* y las informaciones guardadas en él.

Onanismo Tecnológico: crea, recrea y consume tus contenidos

Es imposible seguir las rapidísimas evoluciones e innovaciones “made in Google”: nuevos servicios se suceden de forma casi convulsa y se hace muy difícil entender cuales son los realmente destinados a tener un impacto en nuestras vidas y cuáles, por el contrario, serán olvidados en pocos meses o semanas. Por otra parte, tampoco es interesante o útil perderse en elaboraciones complejas, en taxonomías exhaustivas que inevitablemente dejan atrás excepciones significativas, si consideramos la elevada depauperación de las informaciones y de las innovaciones en Internet. La naturaleza fluida y dinámica de la Red desanima, si acaso alguien todavía sintiera la necesidad, cualquier enfoque totalizador y omni-comprensivo, que se revela veleidoso en el momento mismo de su exposición farragosa.

Sin embargo, es posible esbozar, desde una perspectiva subjetiva, una visión global del fenómeno Google, sin adentrarnos en las técnicas específicas y tampoco en los pronósticos futuribles de escasa plausibilidad. Es sin duda relevante extremar la figura del “*prosumer*”, mediante el acento puesto en el concepto de personalización.

Google es conocido por la costumbre consolidada de conceder versiones betas, no definitivas y todavía en fase de test, de sus servicios: esta dinámica, como hemos visto en el capítulo precedente, está directamente asistida por las comunidades de desarrollo Free Software. Los usuarios, dando su feedback, sugerencias e impresiones de uso, contribuyen de manera determinante a la afirmación o no de un nuevo servicio: a la vez son productores y usuarios de los servicios que usan, figuras híbridas llamadas por eso *prosumer*⁹¹. Colocándose como sujeto de mediación global de los contenidos de la web, Google vende sustancialmente tecnologías y resultados de búsqueda (mediante la publicidad) a los usuarios, que por un lado tienden a ser creadores de los contenidos de la Red, por el otro los consumidores de aquellos mismos contenidos que han modelado en primera persona a través de los servicios de Google.

Dos ejemplos, aparentemente desligados entre ellos, pueden esclarecer este ciclo cerrado de producción y de consumo de los contenidos: el Google Web Toolkit (GWT)⁹² y la convergencia entre GTalk y Nokia⁹³.

Google ha concedido en mayo del 2006 el Google Web Toolkit, un framework que permite de desarrollar aplicaciones AJAX mediante el lenguaje Java. AJAX (Asynchronous JavaScript and XML) es una técnica para desarrollar aplicaciones web interactivas y dinámicas usando una combinación del clásico HTML (o XHTML) con CSS para la parte visual y JavaScript para mostrar dinámicamente las informaciones e interactuar con ellas. En resumen sitios extremadamente dinámicos en los que no es necesario esperar cada vez que se carguen todas las informaciones de la página. Gmail, por ejemplo, utiliza AJAX. La novedad es significativa porque de este modo cambia el enfoque de la creación de las aplicaciones web, que se escriben en un lenguaje de objetos de alto nivel, Java, luego recogidas por GWT y compatibilizadas con todos los browser. Sin embargo, esta innovación no justifica el anuncio en pompa magna de la

⁹¹ Véase a este propósito el apéndice I: El fin del mundo en una galleta

⁹² <http://code.google.com/webtoolkit/>

⁹³ <http://googlesystem.blogspot/2006/05/nokia-770-will-include-gtalk.html>

Llegada de un fantomático “web 2.0” que revolucionará Internet adaptándola a las máquinas. En efecto, tecnologías para la creación de software multiplataforma, para compartir los bookmark, el *social networking*, la agregación automática de los datos, etc., son conocidas desde hace años. Además, la hipocresía de multinacionales como SUN que alardean del ingreso en la Era de la Participación (*Partecipation Age*), por un lado olvida la actitud a la cooperación en ámbito informático propia de la cultura hacker desde hace decenios, por otro vende como una gran revolución el uso de los estratos más elementales (estandarización mediante XML/RDF) de las innovaciones propuestas por organismos como el W3C a propósito de la web semántica⁹⁴.

Desde luego AJAX y tecnologías parecidas solucionan a la raíz el problema muy común de la portabilidad de los sitios web, que actualmente son difícilmente visualizables por todos los browser. El código del framework es disponible bajo licencia Apache, es decir sustancialmente Open Source, pero, como acontece con las iniciativas de code.google.com, algunas partes fundamentales (en este caso, el compilador Java-to-JavaScript y el *hosted web browser*) se distribuyen solo en binario y es preciso suscribir una licencia ad hoc, que prohíbe redistribuirlos, derivar de ellos ulteriores productos e incluirlos en productos comerciales. Además, cada vez que se usa el *hosted web browser*, que permite comprobar las aplicaciones en la propia máquina antes de publicarlas en internet, se contacta un servidor de Google, oficialmente para averiguar que se esté usando la versión más actualizada del producto. Es evidente, sin embargo, que se trata de un método de control extremadamente eficaz, aplicado a los desarrolladores en lugar de a los usuarios. Cierto, el código producido se puede distribuir libremente, incluso para fines comerciales.

GWT es en práctica un instrumento simple para crear sitios perfectamente compatibles con los sistemas de indexación de Google. Actualmente, es preciso conocer un lenguaje no banal como Java, pero no es impensable imaginar el desarrollo de instrumentos que permitan incluso a un usuario en sus primeras armas posicionar en la propia página web objetos como barras de los instrumentos, galerías de imágenes, menú de vario tipo y toda especie de objetos web sin escribir ni una línea de código. Naturalmente ya existen programas extremadamente sencillos para crear sitios web (WYSIWYG, *What You See Is What You Get*); pero GWT es vía web. Estos contenidos estarían pues inmediatamente listos para ser usados en dispositivos fijos o móviles de toda especie, siempre y cuando puedan acceder a la web.

Imaginemos ahora que Google firme acuerdos comerciales para la construcción de hardware a medida, ofreciendo a quienquiera que use sus servicios instrumentos sencillos para construir páginas web visibles en los PC, pda, ecc, e indexadas de la manera más fácil posible por sus spider. Google de hecho no proporciona programas bajo pago de licencias, como Microsoft: necesita, como sabemos, difundir sus propios estándares para poder gestionar más fácilmente su economía de búsqueda.

⁹⁴ Con respecto a la implementación de la web semántica, cosa distinta del “web 2.0” véase la *roadmap* de 1998 de Tim Berners-Lee, <http://www.w3.org/DesignIssues/Semantic.html>; la coordinación del W3C, <http://www.w3.org/2001/sw/>, un documento introductorio a la filosofía de la web semántica: <http://www2.autistici.org/bakunin/doc/reti/index.xml>

Y hete GTalk y Nokia. GoogleTalk es el servicio VoIP⁹⁵ de Google, recién integrado en la cuenta de correo de GMail, de modo que los que forman parte de la “Google community” puedan no solo enviarse correos, sino también chatear y hablar en tiempo real. A partir de finales de mayo 2006 está disponible para los nuevos dispositivos móviles de Nokia, llamados “internet tablet”, especie de móviles evolucionados pensados para navegar por la web. Así Google entra por la puerta principal en el mundo de la telefonía móvil, con la perspectiva de integrarse bien pronto con las redes públicas sin cable (*wireless* o *wimax*) que empiezan a ser instaladas en distintas ciudades, aeropuertos, áreas de servicios en las autopistas, ecc.. Una perspectiva de convergencia en el video es también plausible: video.google.com es una mina de video, y la televisión en los móviles es el próximo paso.

En otras palabras, Google, facilita los instrumentos para crear contenidos según sus estándares. Esta es la personalización extrema, el equivalente del mecanismo “colas largas” (que permite ofrecer exactamente el producto personalizado a cada consumidor) en el campo de la creación de contenidos web: el usuario crea “exactamente” lo que desea en el formato estándar de Google. La descentralización total a nivel de creación de contenidos es paralela a la descentralización total de la publicidad, y por lo tanto del suministro de productos “personalizados”. Un sistema penetrante para imponer un estándar, pero formalmente “democrático” porque en manos de los usuarios, a portada de click, a portada de browser. La que se vende como democracia electrónica se convierte en una estandarización que permite digerir los contenidos creados por una miríada de usuarios y asociarles la publicidad más adecuada.

Browser como ambientes de desarrollo

La explosión de nuevos y cada vez más potentes servicios web desde 2005 está convirtiendo el browser desde un instrumento de simple navegación en un instrumento de desarrollo. Un amplio abanico de tecnologías altera los estándares de programación web corrientes, ofreciendo a los desarrolladores un instrumento fácil, multiplataforma, completo y fiable: el browser mismo.

En los últimos años se ha ido moldeando lentamente un nuevo estilo en la creación de los sitios internet, también gracias al mayor cuidado hacia la portabilidad y la accesibilidad de los contenidos: claro ejemplo de ello son la difusión de las hojas de estilo (*Cascading Style Sheet*, CSS y CSS”) en lugar del HTML nudo y crudo, de los **validadores**, del mismo estándar XML. Gráficos y web designer hallan óptimos aliados en los browser, cada vez más sofisticados y respetuosos de los distintos estándares, que les consienten realizar sitios conformes a las normas de visualización sobre los dispositivos más disparatados, permitiéndoles a la vez mantener y ampliar la propia libertad de expresión.

En especial, el nacimiento de browser como Mozilla Firefox y su rápida difusión ha permitido una fuerte interacción entre los desarrolladores de sitios de internet y desarrolladores de este browser, consiguiendo en poco tiempo sanar casi todos los bugs y las incompatibilidades entre los estándares web. La incompatibilidad entre Internet Explorer, Opera y muchos más browser propietarios o no, es un problema bien conocido por quienquiera haya creado páginas web. La sinergia desarrollada por Mozilla, aparentemente simple y tal vez obvia, constituye una novedad

⁹⁵ El acrónimo VoIP, Voice over IP (Voz mediante protocolo Internet), indica una tecnología que hace posible realizar una conversación telefónica explotando una conexión internet u otra Red dedicada que utiliza el protocolo IP, en lugar de pasar a través de la normal línea de transmisión telefónica. Vease http://en.wikipedia.org/wiki/Voice_over_IP. No solo existen Skype y Google que ofrecen servicios VoIP, sino también programas free, como Asterisk, <http://www.asterisk.org/>

absoluta en la historia de los browser. Otra característica interesante de los productos Mozilla es la estructura modular desarrollada alrededor del motor de rendering Gecko, que permite añadir cualquier funcionalidad imaginable: entre los más conocidos y difundidos instrumentos para eliminar la publicidad en las páginas web, para tener siempre a la vista las previsiones del tiempo o la situación del mercado bursátil.

Los browser son, en resumen, instrumentos cada vez más confiables, que permiten la creación de complejos sitios internet y asumen las características de verdaderos y auténticos programas: tienden incluso a sustituir las aplicaciones de uso común. El ejemplo patente son los varios instrumentos para oficina propuestos por Google como alternativa web a la suite de productividad de Microsoft Office y también de OpenOffice⁹⁶. Actualmente, es posible utilizar Writely (producido por una sociedad adquirida) como elaborador de textos, Google Spreadsheets como hoja electrónica, Google Page Creator para crear páginas web rápidamente. Todos los servicios están en fase de *betatesting* con invitaciones: es preciso tener un account Google para utilizarlos, evidentemente bajo estricto control.

Los desarrolladores, por su parte, desplazan su interés cada vez más hacia la web, también gracias al uso de instrumentos como GWT. Naturalmente Microsoft no se queda mirando y tiene ya lista la versión beta (alineándose a la estrategia de “*betatesting* difuso” de Google, procedente como sabemos de las prácticas del Free Software) de Office System (o sea Office 2007), que prevé una integración cada vez más lanzada con respecto a las herramientas web, pese a seguir siendo una aplicación a instalar.

Los browser se están convirtiendo en ambientes de desarrollo completos para la creación de contenidos estándar, es decir SDK (*Standard Development Kit*). ¿Dónde está realmente la innovación que ha hecho posible transformar los browser en SDK? Sin duda, podemos hablar de un nuevo paradigma de la programación: se pueden crear ahora programas realmente multiplataforma, distribuidos, lado client, residentes en los server, fácilmente actualizables, y sin necesidad de complejos framework a instalar en los ordenadores de los usuarios. Los contenidos, incluidos los datos personales de los usuarios, son almacenados cada vez más en máquinas remotas (por ejemplo los server de Google), a los que se accede vía web, es decir mediante browser.

A menudo la elección de browser Open Source, como Mozilla, se debe a su sencillez de configuración, y a la disponibilidad de numerosas extensiones potentes y gratuitas. Los desarrolladores se basan en estos browser para implementar programas cada vez más complejos y estructurados. El nacimiento de programas basados solamente en la web provoca dos fenómenos, ambos con fuertes recaídas en el mercado y en los usuarios: los programas binarios a instalar devienen obsoletos y los browser se convierten en programas más complejos, a menudo fuertemente modulares, y más apetecibles en el mercado del IT. Por lo tanto en futuro podemos esperarnos menos programas .exe (Windows), .dmg (Macintosh), o paquetes Unix instalables, y más instrumentos integrados en los browser, desde las extensiones para leer los feed RSS al podcasting, de Gmail a software completos para oficina.

⁹⁶ también existen suite completas para oficina utilizables vía web, como gOffice, producido por la sociedad Silveroffice, Inc. de San Francisco, que sostiene: “gOffice no es afiliada de Google, Inc., pero nos gusta pensar que será una bella lucha mientras su misión seguirá siendo la de indexación de todas las informaciones del mundo, mientras que la nuestra es la de ayudar a la gente a crear esas informaciones. Necesitamos pues de Google para buscar los contenidos que nuestros usuarios crean con gOFFICE! Tal vez tendríamos que añadir una ficha a nuestra barra de herramientas llamada “búsqueda”... www.goffice.com

El control capilar que puede ser realizado por parte de los proveedores de los servicios web hace que esta dinámica sea potencialmente peligrosa para todos nosotros. Quien ofrece los servicios, en efecto, conoce exactamente la identidad digital de los propios usuarios, el tiempo de uso del software y los contenidos en elaboración, ya que controla todo acceso y detalle relativo al uso de los instrumentos que pone a disposición.

Se trata, a nivel técnico de una interacción continua entre client (browser) y el server (quien ofrece los servicios), donde este último es capaz de monitorar constantemente las solicitudes, los tiempos y los elaborados finales. Muta también el escenario de la seguridad: con la excusa de ofrecer un productos “a salvo” de las garras de cracker y hacker de toda clase, los mecanismos de autenticación no se encuentran ya en el programa compilado y concedido sin fuentes, sino que están directamente en los server de los distribuidores. Por lo tanto, para “entrar” en un software ya no basta con desensamblar el código en el propio ordenador para comprender como funciona, sino es preciso “perforar” server remotos.

Privacy, Paranoias, Poderes

La estrategia de la acumulación ha hecho de Google un sujeto capaz de poner en dificultad hasta Microsoft, empezando una guerra sin exclusión de golpes para la estandarización y el control del acceso a la web y a las redes que usamos cada día. En cuanto sujeto de mediación global de informaciones, el fenómeno Google tiene que ver directamente con todos los usuarios de informaciones digitales, es decir todos nosotros. Recorrer la historia de Google significa por esto re-recorrer nuestra historia de exploradores de internet y de la web, pues demasiado a menudo hemos delegado la gestión de nuestras informaciones, nuestros sitios, galerías de imágenes, cuentas de correo, blog, sms, conversaciones telefónicas, ecc. a empresas todo menos que desinteresadas.

La estrategia de la objetividad, que Google actúa poniendo el acento en la investigación, en la excelencia académica, en la superioridad tecnológica, en el cuidado estudio de las interfaces, es un velo que oculta la inquietante perspectiva de un único punto de acceso a los datos producidos por los ingenios usuarios.

La estrategia del Open Source, finalmente, es necesaria a Google precisamente para cooptar el método de desarrollo cooperativo de las comunidades digitales, haciendo que sea funcional a la propia “misión”. También en este caso hemos visto como Google se mueve con extrema desenvoltura, proponiendo métodos “nuevos” para explotar dinámicas notas: *Summer of Code* representa un ejemplo evidente de ello.

La actividad de Google, en resumen, constituye un peligro evidente para cualquiera que esté interesado en las temáticas de la privacy y, en un sentido más amplio, en la construcción conciente del propio alter ego digital. Se trata de la emersión de un conglomerado de poder que ya hoy en día influye pesadamente sobre la vida de demasiados individuos. Google detiene informaciones reservadas que analiza sin descanso para promover una personalización cada vez más cuidada del cáncer publicitario. Además, ya que la acumulación de poder generalmente favorece la ambición de dominio, sería oportuno indagar en profundidad este fenómeno.

No existen respuestas globales capaces de solucionar una vez por todas el problema de la privacy. El Gran Hermano no existe, o por lo menos, como toda paranoia que excluye horizontes de liberación y oscurece posibles vías de fuga, es útil y funcional al ejercicio del poder dominante.

Ocultar, criptografiar, estenografiar, son prácticas útiles, pero no son soluciones definitivas: el objetivo y el deseo quedan siendo la comunicación y el compartir, que solo la “publicación”, es decir el movimiento de “hacer público”, permite. La obsesión por la privacy cae muy rápidamente en la paranoia del complot: en este sentido, no es útil entregarse a la construcción de complicadas alternativas para crear redes absolutamente seguras e impenetrables. La tecnología es una ocasión de apertura y participación, ya que usar las máquinas significa usar criaturas híbridas, artefactos materiales (pertenecen en ese sentido al mundo natural) investidos de valores y significados culturales (pertenecen al universo de la “cultura”). Las redes son expresión de una dinámica co-evolutiva de máquinas mecánicas, biológicas y significantes: la tecnología es mestiza por nacimiento. Crear redes significa de hecho conectar máquinas de vario tipo; significa crear métodos para compartir, métodos de traducción, métodos de intercambio: no es posible quedarse encerrados en sí, es preciso ponerse en juego y cambiar.

Son necesarias investigación y análisis competentes; la denuncia de los mecanismos del dominio tecnocrático hace falta más que nunca. Abdicar al pensamiento crítico equivale a sucumbir a las paranoias del control, que sí se hace cada vez más difuso. Es posible utilizar en sentido paradigmático la historia de Google para trazar líneas de fuga no ingenuas, para imaginar prácticas de autogestión de las tecnologías. Google es de hecho el punto de encuentro entre el estilo meritocrático de la academia, el empuje a la innovación a toda costa y el capitalismo financiero más avanzado.

Es la ocasión precisa para desarrollar redes autónomas y descentradas, para contraponer el deseo de “explorar” y “recorrer” las redes a la necesidad de “acceder” a los datos, para focalizar la atención sobre el camino en lugar que en el resultado.

VI. Calidad cantidad relación

La emergencia de la información

La sociedad de la información presenta una extrema heterogeneidad de manifestaciones: utiliza al mismo tiempo y además con total desenvoltura sistemas de comunicación reticulares (telefonía), versiones digitales de media broadcast⁹⁷ tradicionales nacidos fuera de la Red (diarios, radio, TV. online, etc.), instrumentos de comunicación universales nacidos con la Red (e-mail) y sistemas distribuidos (p2p). En cualquier caso todos estos sistemas se basan en una única materia prima: la información. Dentro del específico ámbito de los motores de búsqueda, o sea del descubrimiento de informaciones (information retrieval), podemos deducir que la información es el conjunto de las páginas que se publican en la Web⁹⁸.

El crecimiento cuantitativo y cualitativo de estas páginas y de sus contenidos ha sido y sigue siendo desmedida. Esto depende de la facilidad con la que es posible generar nuevos contenidos para la Red. Los contenidos no son islas de datos sin conexión, sino que se constituyen en la variedad de las relaciones y de los link que ligan las páginas, los sitios, los documentos, los mismos contenidos entre sí.

El acceso directo, sin mediación, a esta masa de informaciones es absolutamente imposible, hasta en vía hipotética: sería como sostener que se puede hojear la Web “a mano”. Por esto existen los motores de búsqueda, herramientas estudiadas para filtrar la complejidad de la Red y actuar como interfaces entre nosotros y la información, solucionando las búsquedas.

La cantidad de páginas memorizadas llega a tener entonces un valor crucial para evaluar la potencia técnica y económica de un motor de búsqueda. Cuanto mayor es la amplitud del patrimonio de las páginas examinadas, tanto mayor será la potencial profundidad y fiabilidad de las búsquedas efectuadas, por supuesto en el límite de ese contexto específico.

Pero, aunque la extensión del patrimonio de un motor de búsqueda pueda ser enorme, nunca podrá ser completo y total, con independencia de cuanto tiempo, dinero y tecnologías se le dediquen. Es absurdo pensar poder conocer, o más banalmente copiar y catalogar, todo Internet: sería como pretender conocer la totalidad de un mundo vivo y en constante mutación.

Los dispositivos de almacenamiento de las informaciones de los motores de búsqueda como Google se parecen a unos vasos. Imaginemos tener que llenar un vaso muy muy capaz con esas diminutas gotas que son las páginas Web que constituyen las informaciones de la Red. Si consideramos que el vaso es capaz de contener la totalidad de las gotas (la totalidad de la Red) nuestra tarea consistirá en intentar capturarlas todas de manera sistemática y repetitiva.

Al revés si pensamos que las gotas son demasiadas para nuestro vaso y que su recogida no será previsible según algoritmos, ni rápida, sino que potencialmente podría hasta revelarse una tarea ilimitada, deberemos cambiar de táctica: también porque las informaciones varían con el tiempo, las páginas se modifican, los recursos cambian de dirección...

⁹⁷ Un sistema de transmisión *broadcast* envía información a muchos sistemas receptores de forma uni-direccional (comunicación de uno a muchos). En una Red de ordenadores, una transmisión *broadcast* envía informaciones a todos los ordenadores en conexión a esa Red; otros sistemas de transmisión son *unicast* (un solo receptor), *multicast* (muchos receptores), *anycast* (cualquier receptor dentro de un grupo). Los sistemas p2p, precisamente como los sistemas telefónicos, son el exacto contrario de los sistemas *broadcast*, porque determinan una comunicación del tipo muchos-muchos.

⁹⁸ No es este el lugar para un exhaustivo análisis del concepto de “información”, aunque esta definición está simplificada adRede. Por ejemplo no tiene en consideración que los motores de búsqueda como Google indexan también otros contenidos que no son Web, como el archivo de los mensajes Usenet o el contenido de los ordenadores que utilizan Google Desktop; de todas formas, puesto que esta masa de informaciones se presenta a través de una interfaz Web de mediación, consideramos aceptable la definición adoptada.

Acaso podríamos decidir elegir tan solo las gotas más grandes, o concentrar nuestros esfuerzos de recogida en puntos donde caen en mayor cantidad, o más simplemente, podríamos recoger aquellas que nos interesan más, y luego intentar ligarlas entre sí de la forma que consideremos más interesante.

No obstante los motores de búsqueda continúen poniéndose el objetivo de catalogarlo todo, es más correcto aplicar al Web una praxis localizada, o sea asumir que para cada “intención de búsqueda” existen muchas respuestas plausibles, entre ellas algunas son “mejores” pues cumplen con determinados requisitos de rapidez y profundidad. Debemos recordar siempre que la calidad de las respuestas depende de nuestra subjetiva percepción de la aceptabilidad del resultado. Para poder aceptar o rechazar un elaborado de búsqueda resulta esencial el ejercicio de la capacidad crítica, la conciencia de la subjetividad del propio punto de vista. Para generar el recorrido que de verdad nos interesa analizar es necesario asumir la existencia de una Red finita y limitada, un mundo cerrado solo por nuestras personales exigencias, sabiendo de todos modos que de una localización subjetiva se trata, ni absoluta ni constante en el tiempo. Desde el punto de vista analítico, explorar una Red implica la capacidad de dividir las redes en subredes de análisis y se corresponde con la creación de pequeños mundos localizados y temporales (LCW, Localized Closed World), en los que existe al menos una respuesta a las búsquedas efectuadas. En caso contrario, muchas búsquedas serían sin fin, entre otras cosas porque los datos que pueden ser analizados son muy superiores a las capacidades de elaboración de un sujeto humano. Dicha respuesta podría no ser satisfactoria: cambiar o especificar la búsqueda, enriquecer el propio punto de vista entonces podrá generar un recorrido más coherente con las exigencias. Considerar la Web un mundo cerrado localizado significa también reconocer que la dinámica misma con la que las informaciones nacen, germinan y se distribuyen en Red (hasta se abandonan a sí mismas, ya no se actualizan) es un fenómeno “emergente”, ni causal ni tampoco casual. La emergencia⁹⁹ es un comportamiento describable en términos matemáticos, una explosión de complejidad inesperada e imprevisible, sobre todo es un evento en condiciones de determinar situaciones no descriptibles de manera exhaustiva. Analizar y recorrer un “universo emergente” como la Web implica por tanto un continuo re-posicionamiento personal que determina un “mundo cerrado localizado” de competencias y expectativas, y también la abertura hacia nuevas posibilidades de exploración (hay otros mundos siempre posibles fuera del propio mundo cerrado) y entonces la conciencia de poder alcanzar resultados siempre y solo parciales.

Cantidad y Calidad

La indexación por acumulación de páginas es un fenómeno cuantitativo, que en sí no determina la calidad de las informaciones de la Web: el objetivo es recolectar todas las páginas sin elegir entre ellas. Las relaciones entre páginas, aún siendo generadas sobre la base de un criterio sencillo, el sistema de los link, generan fenómenos de emergencia. Entonces la calidad de las informaciones, a pesar de su tipología, está determinada por la capacidad de trazar recorridos, sin preocuparnos por cubrir la “totalidad” de las informaciones. La calidad depende sobre todo de la puesta en acción del propio punto de vista a lo largo de un recorrido de búsqueda: en suma, son los navegantes, los piratas, los usuarios de la Web quienes determinan y mejoran la calidad de las informaciones, entrelazando relaciones entre las páginas. La potencia de acumulación de los algoritmos de Google es útil, pero insuficiente.

⁹⁹ La emergencia es un principio que describe el comportamiento de los sistemas complejos. Se puede definir como el proceso de formación y emersión de comportamientos y esquemas complejos a partir de reglas más sencillas. Las ciencias del conocimiento utilizan el concepto de emergencia para describir “la explosión” de la complejidad típica de la inteligencia: partiendo de elementos sencillos inconscientes de por sí (las neuronas), se generan el pensamiento, la conciencia manifiesta de los procesos cerebrales:
http://es.wikipedia.org/wiki/Emergencia_%28filosof%C3%ADa%29

La evaluación del contenido de las páginas y de sus relaciones está delegada a algoritmos y, sobre todo, a las sociedades que los controlan. El fenómeno Google es el resultado de esta delegación: una potencia aparentemente sin límites, en condición de ofrecernos la posibilidad de encontrar “algo” útil e interesante, entre los muchos recursos contenidos en su patrimonio, que se contrabandea por la totalidad de la Web. Pero se silencian del todo las limitaciones de estas chispeantes ofertas: lo que falta del patrimonio, o lo que está solo parcialmente, y, sobre todo, todo lo que ha sido descartado.

Sigue siendo de urgente actualidad un delicado problema ético y político acerca de la gestión y el control de las informaciones: ¿qué sujeto puede garantizar la rectitud de un empresa, por “buena” que pueda ser, cuyo objetivo primario es el lucro?

Si por una parte ingentes recursos económicos y una óptima infraestructura tecnológica pueden mejorar continuamente el almacenamiento y la recuperación de los datos, las cuestiones políticas derivadas de la acumulación de informaciones no son de ninguna forma separables, pues Google representa de hecho una inédita concentración de informaciones, más o menos reservadas, fuente en todo caso de un enorme poder sin transparencia. Por supuesto ninguna ley sobre privacidad podrá mejorar esta situación, tampoco la creación de instituciones nacionales o internacionales ad hoc para el control de las informaciones personales y de los datos sensibles. Una respuesta sensata a la cuestión de la confidencialidad de las informaciones solo puede llegar de una responsabilidad y conciencia mayores por parte de los individuos quienes crean la Red, mediante elecciones de formación individual, aunque se trate de un camino difícil y trabajoso de realizar.

El mito de la búsqueda instantánea

Aclarado que el enorme patrimonio de Google nunca podrá corresponderse con la totalidad de la Red, es ingenuo, o viceversa sagaz, proponerse como una interfaz “instantánea” entre las peticiones de búsqueda de los usuarios y el resultado supuestamente “exacto”. Cada vez que se efectúa una navegación visitando páginas Web se siguen unos link que crean recorridos analizables según los modelos matemáticos de la teoría de los grafos. La Web está de hecho compuesta por nodos (las páginas y los sitios Internet) y por arcos (los link que los unen). Por cuanto puedan crecer las dimensiones de la Web, las directrices propuestas por los motores de búsqueda serán siempre más eficaces en llevarnos al objeto “justo”. Mediante la aplicación de criterios de eficacia/eficiencia, el recorrido hallado por la interrogación de un motor de búsqueda será optimizado, es decir el número de nodos visitados será muy bajo y el tiempo para hallar el resultado será casi instantáneo. Google llega a auspiciar un solo paso, hasta el punto de ofrecer el botón “me siento afortunado” en su página principal.

Esta optimización limita la búsqueda a un esquema secuencial articulado en tres puntos: *usuario-algoritmo-objetivo*. Con el tiempo esta dinámica lleva a una postura de pasividad digital, “esperamos” que se nos ofrezcan unos datos para poder “elegir” entre ellos.

Además esta eficacia/eficiencia paradójicamente no se fundamenta en la ampliación del cauce de datos en los que efectuar las búsquedas, sino en la reducción del acceso al mismo patrimonio informativo, pues cada recorrido propuesto por el motor de búsqueda no será trazado en la Red nunca “en ese momento”, sino calculado en base a su patrimonio almacenado, y personalizado sobre el usuario a base de filtros algorítmicos y *cookies*.

El acceso a las informaciones propuesto por Google es rápido, mejor, en sus intenciones tiende a ser inmediato, a anular la dimensión del tiempo, a pesar de la enorme extensión de las bases de datos consultadas. La mediación de la tecnología (interfaces, algoritmos, búsquedas preconfeccionadas, etc.) permite esta “anulación” temporal y produce un acceso prácticamente

instantáneo.¹⁰⁰ La calidad de la búsqueda padece una influencia negativa por la rapidez en alcanzar el resultado. Cualquiera haya intentado llevar a cabo una búsqueda, bien sabe que el tiempo empleado es un factor determinante en la composición de la experiencia: crear un propio recorrido, tomar diferentes decisiones según el momento genera sentido de participación y satisfacción. Google nos permite “localizar” en el espacio (en su espacio multidimensional) lo que queremos; pero por cuanto breve pueda ser el tiempo de espera del resultado, siempre estamos asumiendo una postura pasiva frente a un oráculo tecnológico.

En el marco de la exploración activa, por el contrario el objetivo no es acceder a los datos, sino realizar una navegación rica y diversificada, construir recorridos complejos mediante la búsqueda. El concepto de eficiencia decae. Cuanto mayor será el número de nodos que visitamos y la complejidad del retículo que construyamos, tanto mayor será la oportunidad de ampliar nuestros campos de interés, las posibilidades de operar elecciones significativas, de refinar nuestras búsquedas. Este acercamiento permite un enriquecimiento cognitivo más allá de la prestación inmediata. Por ejemplo, cuando visitamos los link propuestos por un sitio de nuestro interés y seguimos en los sitios señalados como amigos, creamos un recorrido único; a lo mejor aprovechando un marcador nuestro. Al nivel visual un proceso como este se opone a la linealidad de la secuencia usuario-algoritmo-objetivo y crea al revés un recorrido lleno de bifurcaciones, ramificaciones, saltos, giros retorcidos, respondiendo a un deseo cognitivo no lineal¹⁰¹.

En definitiva, los motores de búsqueda son excelentes herramientas en cuanto a la satisfacción de los aspectos cuantitativos de la búsqueda, cuando nos movemos dentro de recursos ya totalmente estructurados, como diccionarios, enciclopedias, etc. La cantidad es directamente proporcional a la potencia de cálculo y de acumulación; la extensión de Google es decididamente superior a aquella de cualquier otra competencia, pero para mantener esa distancia necesita de una expansión constante, en recursos, algoritmo, máquinas, usuarios, etc.

La máquina de la acumulación se alimenta de la búsqueda de la eficiencia y de la excelencia a nivel de prestaciones, a pesar de una evaluación de los contextos de búsqueda: el interés principal de un motor de búsqueda es el incremento de su propio patrimonio, y en consecuencia de su propia oferta. Una lógica exclusivamente cuantitativa se limita a la idea que cuantas más son las informaciones disponibles, tanto mayor será la posibilidad de acceder a informaciones coherentes con la intención de búsqueda: tesis que de nuevo confirma el esquema lineal que uniforma el business de los motores de búsqueda.

Por lo contrario, la calidad no reside necesariamente en la potencia tecno-económica. No está para nada descontado que a la amplitud de los almacenes de la información disponibles se corresponda con la pertinencia de los resultados ofrecidos: la emergencia del mejor recorrido no se prevé, no se calcula, sino que se construye un paso tras otro.

Detrás de la capa del mito

Los valores del posicionamiento del ranking de Google no responden a ningún criterio transparente de evaluación; a pesar de esto, las respuestas ofrecidas, en su mayoría son exhaustivas, mejor, es imposible para nosotros evaluar si algo se le ha escapado a su spider, a no ser que seamos expertos de un determinado sector y conozcamos un recurso que no resulta indexado.

¹⁰⁰ La satisfacción de los usuarios con los resultados obtenidos por este método “inmediato” es toda otra cosa: véase al propósito Jean Véronis, A comparative of six search engines, Université de Provence, <http://www.up.univ-mrs.fr/veronis>, <http://aixtal.blogspot.com>

¹⁰¹ El texto básico para acercarnos a las dinámicas cognitivas no lineales queda el de Humberto R. Maturana, Francisco J. Varela, Autopoiesis y Cognición- La realización de lo viviente, Marsilio, Venezia, 1985 (ed.or. Autopoiesis and cognition. The Realization of living)

La difusión capilar de los instrumentos de búsqueda de Google lo hacen un estándar *de facto*. El *blank box* (blank box) donde introducimos las palabras clave (*keywords*) de nuestras búsquedas es para el usuario una especie de *weltanschauung*, o *vision*, en suma difunde una bien precisa “visión del mundo”, puesto que encarna la idea de “servicio total” (*total service*): el motor de búsqueda es capaz de contestar a cualquier pregunta, de satisfacer todas las exigencias frente al instrumento Internet.

En términos epistemológicos, el *blank box* representa un modelo cognitivo de organización del conocimiento: pedimos al espacio blanco la respuesta para todas nuestras intenciones de búsqueda, bien sea cuando necesitamos documentos, profundizaciones, informaciones, que cuando simplemente deseamos “navegar”. Ya lo utilizamos hasta cuando podríamos consultar las paginas amarillas, recordar la nota que hemos marcado en un libro, llamar a alguien, pensarlo dos minutos, etc....pero Google es más cómodo. La actividad de búsqueda se identifica por completo con el objeto que ofrece el servicio, Google, del cual experimentamos una percepción redundante.

La costumbre a la utilización de las herramientas se transforma en comportamiento, en coacción a repetir: es muy difícil que intencionalmente los usuarios intenten satisfacer su “necesidad de input” con otras modalidades, habiéndose aficionado a la comfortable eficacia/eficiencia del *blank box*.

Estar en movimiento en la Red necesitando por tanto de interfaces de acceso, de herramientas para encontrar informaciones y construir recorridos, es una actividad diferenciada y profundamente contextual. La actividad de búsqueda no es para nada homogénea y no se puede reducir a la utilización del *blank box*. Lo que pedimos y deseamos no depende solo de una necesidad expresable en términos analíticos de información cuantitativa, sino también de cómo nos relacionamos con la búsqueda, del contexto de la misma, de nuestro bagaje cultural, y de la actitud frente a la novedad, a nuevos territorios, a la diversidad. Es imposible agotar las peticiones de información por medio de una solución unívoca.

Como la indexación de las páginas es parcial, se trata pues de una selección basada en el ranking, lo que Google nos propone es, banalmente, la posibilidad de encontrar “algo” útil e interesante para nosotros entre las muchas cosas que están almacenadas en su patrimonio de interés. Una intención de búsqueda, sin embargo supone en última instancia el deseo de encontrar, más bien descubrir, “todo lo que no se sabe pero es posible saber”. Entonces el gigante nos aparece tal y como es, enorme, extenso, ramificado, y no necesariamente apto para nuestras búsquedas.

Modelos de búsqueda

La ambigüedad sugerida por los motores de búsqueda según la cual parece que “busquemos en un ambiente infinito”, en lugar que en un mundo cerrado localizado sobre la base de nuestras intenciones de búsqueda, depende del solapamiento formal entre el nivel de la interfaz¹⁰² y aquello de la organización. En este específico contexto la interfaz es el elemento tecnológico que garantiza el acceso a las informaciones y la ejecución de la actividad de búsqueda; la organización es, al contrario, la arquitectura, el modelo tecnológico con el que las informaciones son ordenadas y almacenadas. Los dos niveles se influyen entre ellos, de hecho las elecciones en términos de organización determinan la utilización de interfaces específicas, y a su vez las informaciones visualizadas por medio de las interfaces reflejan la forma en la que se conservan. El problema de este solapamiento reside en la presentación bajo la forma de datos identificables de manera unívoca, separados entre ellos. El usuario de Google se mueve de forma lineal en la

¹⁰² La interfaz es el punto, el área, el territorio que pone en conexión elementos cualitativamente diferentes; en el caso de los motores de búsqueda, la interfaz permite la relación entre el sujeto que interroga y el cauce de informaciones.

lista de los resultados del ranking; para pasar de un resultado a otro debe volver al listado originario, sin conexiones transversales al nivel de la interfaz.

Con los motores de búsqueda se encuentran informaciones sin tener en consideración el recorrido realizado. La interfaz que determina nuestra interacción es el campo de las *keywords*, adonde se introducen las palabras o las fórmulas para dar comienzo a la búsqueda; en este primer nivel de acceso todas las informaciones se encuentran en el mismo plano y homogéneas, al mismo tiempo separadas y fragmentadas, posibilitando un listado de resultados ordenados por un algoritmo en base a su pertinencia.

Pero en las búsquedas diarias los resultados mismos pueden estar ligados entre ellos de distintas maneras, sin volver necesariamente al listado ordenado de los resultados, y no hay un solo resultado correcto; es más, una búsqueda que no sea dirigida a datos estructurados en enciclopedias, vocabularios y similares (que también mudan en el tiempo), podría no tener solución alguna, necesitando al revés de esfuerzos creativos, de mezcla y recombinación.

Cuando se produce identidad formal entre el nivel de la interfaz y aquello de la organización se obtiene un modelo reduccionista; en el caso de Google, la representación de una potencia de búsqueda infinita, o sea la actividad de búsqueda que se confunde con uno de los instrumentos para realizarla.

Por ejemplo, si buscamos “Penna”(bolígrafo) a través de la interfaz de Google.com, encontramos, entre más de seis millones y setecientos mil resultados: Sandro Penna”, el poeta; “Penna de autor”, el concurso literario; el “ayuntamiento de Penna San Giovanni”, en la provincia de Macerata, y más cosas. No encontramos entre los primeros resultados (*top ranking*) ni recursos referidos a las “Penne (plumas) de las aves”, ni a los “bolígrafos para escribir”, dos conceptos muy lejanos entre sí, de los que queríamos investigar eventuales relaciones. Una perspectiva más amplia acerca de la búsqueda de informaciones, que tenga en cuenta críticamente el potencial cognitivo comprendido en todo proceso informativo, nos fuerza a considerar la función acceso/búsqueda no como un proceso de localización, sino de exploración y creación. El acento se desplaza por lo tanto de la epistemología a la ontología: no solo conocer unas informaciones, sino tomar conciencia del rol de creadores de informaciones¹⁰³. Los motores de búsqueda intervienen en el acceso pero no ofrecen ningún soporte para la exploración, pues se limitan a intervenir en el primer nivel de presentación de las informaciones.

La navegación es el momento de dinamismo real en la puesta en relación entre los objetos digitales, que pueden así expresar al máximo nivel sus potencialidades comunicativas y heurísticas. Se aprende haciendo, y se cambia mientras se va aprendiendo, mientras se está involucrado en la exploración.

Existe una distinción profunda entre buscar y encontrar. Google nos hace encontrar las cosas, produce satisfacción, sensación de acumulación. Pero podría no ser tan interesante el “encontrar”, como el acto mismo de “buscar”; quizás sea interesante no encontrar para nada, pues si no se consigue encontrar quiere decir que estamos metidos en la búsqueda¹⁰⁴.

El motor de búsqueda es un modelo instrumental a través del cual se disponen las informaciones: en lugar de eso sería interesante imaginar modelos para combinar las informaciones, y crear conocimiento.

¹⁰³ Explorar las Redes significa crear las Redes: en este sentido, ampliando el perfil analítico de Maturana y Varela, podemos hablar de “autopoiésis de las Redes”, Redes que se construyen a través de nuestro hacer.

¹⁰⁴ :-)

VII. Tecnocracia

El análisis del fenómeno Google traza un panorama variado, en el que la economía de la búsqueda resulta ser sólo una ficha de un mosaico mucho más amplio y complejo. Por otra parte, si nos remitimos a las declaraciones de Eric Schmidt, en Mountain View están echando las bases para la construcción de una empresa global de las tecnologías de la información, una “empresa de 100 billones de dólares”: evidentemente, no de un simple motor de búsqueda. Se trata en efecto de un sistema pervasivo de gestión de los conocimientos, del que hemos ilustrado las fases y los métodos más significativos: estrategias que conjugan marketing agresivo y meditada gestión de la propia imagen; propagación de interfaces altamente configurables y a la vez siempre reconocibles; creación de contenidos estándar descentrada en usuarios y desarrolladores; cooptación de metodologías de desarrollo cooperativo típicas del Open Source y del Free Software; utilización de sistemas avanzados de recogida y almacenamiento de datos; sistemas de localización de las informaciones correlacionados a técnicas de “perfilación” explícitas e implícitas; personalización de la publicidad.

Google es el dominio tecnológico procedente de la búsqueda científica que se hace instrumento de gestión del conocimiento, expresión directa de la tecnocracia.

Tecnocracia: los expertos de la ciencia

Los expertos han hallado en el control y en la manipulación de la tecnología un arma sin rivales para imponer sus propios intereses personales, para mantener el poder o adquirir mayores privilegios. El mecanismo es simple: la tecnología se utiliza como narración garante de la objetividad de la investigación científica, pero también para avalar las decisiones de los politicuchos que gobiernan, y en general de cualquiera “autoridad” que tenga acceso al oráculo tecnológico.

La aplicación de la investigación científica bajo forma de tecnología es pervasiva y la realidad se interpreta constantemente en obsequio a este paradigma. La curiosidad y el deseo de saber que mueven la investigación científica se encuentran castrados por los criterios de provecho avieso que regulan las financiaciones públicas y privadas. Si la investigación no produce artefactos tecnológicos capaces de generar introitos inmediatos, no es interesante. El discurso del poder se convierte entonces en tecnocracia, el exacto contrario de la participación comunitaria, de la autogestión, de la discusión y de la mediación entre individuos. Hacer pasar la tecnocracia de Google por democracia directa es un juego lingüístico, que nos da la ilusión de participar de una gran democracia electrónica, carente de toda sustancia. Cierto, podemos publicar lo que queremos en Internet, y Google nos indexará. Pero no podemos permitirnos, siendo “profanos” y “heréticos”, observar que la estrategia de la acumulación de Google está en perfecta sintonía con el alocado sistema de la economía de mercado basada en el crecimiento infinito: de hecho, no somos economistas licenciados en la London School, ni empresarios de éxito, es decir no somos expertos, y por lo tanto no tenemos ninguna autoridad en merito. Sin embargo el buen sentido común de orwelliana memoria es más que suficiente para entender que este crecimiento sin fin y sin objetivo es la manifestación de una voluntad de potencia tecnológica interesada en los individuos solo en cuanto que potenciales consumidores.

He aquí porque el PageRank, que como hemos visto no es solo un algoritmo, se convierte en una vestimenta cultural con la que Google nos propone analizar cada cosa: en resumen, asistimos a una extensión forzosa del método de revisión de los pares, valido en ámbito científico, hacia la totalidad del conocimiento humano.

Las autoridades tradicionales, religiosas y políticas, han alcanzado el punto mínimo de credibilidad; pero su escasa capacidad sobre la realidad, en lugar de favorecer la proliferación de espacios de autonomía, ha llevado a una situación surrealista en la que ninguna aserción válida puede ser cumplida sin el apoyo de una autoridad tecnológica. La autoridad de las máquinas, en la mayoría de los casos, es una respuesta extraída de una base de datos, proporcionada por los sacerdotes de la tecnología, los “expertos”, al rico y acomodado pueblo de los *prosumer*. El relativismo extremo se realiza en los métodos para extraer “verdades” de los datos disponibles, casi sin límites: se corresponden con el número de algoritmos y filtros que se pueden aplicar. Una verdad oportuna por cada búsqueda se corresponde a un producto personalizado para cada consumidor evolucionado.

Frente a este cierre en la creación, gestión y aplicación (en nuestra propia piel) de los conocimientos, parece que quedan solo dos posibilidades: rechazar la cultura científica como fuese la causa de todo mal; o, por el contrario, aceptar con entusiasmo fideístico toda “innovación” tecnológica. Sin embargo, entre los dos extremos de la tecnofobia y de la tecnofilia, es posible proponer la curiosidad de la ética hacker es decir el compartir los saberes, la actitud crítica hacia las “verdades”, el análisis cuidadoso de las fuentes, insistiendo en el recorrido de la libre circulación de los saberes.

En este contexto, la formación es un nudo central, pero no hay salidas para una educación científica difusa. La estructura de la formación tanto en Europa como en Norteamérica gira alrededor de la sola producción de especialistas. Actualmente no hay ningún modelo pedagógico que responda a la exigencia de un saber científico de tipo “amateur”, ni siquiera en países de tradición no occidental como Brasil o la India que producen investigación científica de alto nivel y tecnología punta de bajo coste gracias a una competencia despiadada. Un saber científico no académico ni empresarial, difuso y autogestionado, no se admite ni siquiera en fase de discusión: y sin embargo sería indispensable para la formación de competencias de base y para la evaluación de las innovaciones tecnológicas que nos conciernen a todos. Especialmente, la noción global de “educación científica” debería de ser reconstruida para responder a la necesidad difusa de rudimentos prácticos, para enfrentarnos con la larga ola de la revolución tecnológica.

La afirmación de la informática como principal motor de la innovación tecnológica abre, en efecto, nuevos escenarios: la informática no es simplemente una técnica para gestionar la información de forma automática, sino que posee una lógica suya, es decir que trabaja y modifica continuamente sus mismos cimientos. Es física teórica y experimental a la vez: estudia la formalización del lenguaje (formaliza, pues, el conocimiento), la aplica a las componentes físicas de la electrónica, saca de ella lenguajes que a su vez influyen las teorías del conocimiento, se basa en una dimensión recursiva del todo especial.

En las ciencias clásicas se observan fenómenos estables: la física, por ejemplo, elabora datos naturales y sobre ellos crea teorías; en la informática, por el contrario, los fenómenos elaborados en la teoría son artificiales, cambian física y conceptualmente en la medida en que los progresos técnicos y experimentales permiten pulirlos: un software elaborado en un ordenador de hace diez años será estructuralmente distinto de uno más reciente. Lo que ayer era verdad, sabemos ya hoy que no será verdadero mañana, cuando tendremos máquinas más potentes para hacer cosas nuevas, porque hablamos de un mundo vivo y por ello constantemente en devenir.

Milagros de la tecnología: de las opiniones subjetivas a la verdad objetiva

El buen Google, en cuanto gigantesca base de datos, interviene en este panorama y sostiene que somos parte de una gran e inédita “democracia electrónica global”; sostiene que los resultados del PageRank son correctos porque expresión de la democracia directa de los links evaluada por sus algoritmos, nos devuelve pues, en cierto sentido, nuestro legítimo derecho a dar “nuestra palabra”.

La popularidad no puede sino ser considerada índice de “calidad objetiva” desde un punto de vista epistemológico ya que en este caso el concepto mismo de objetividad se basa en un malentendido, es decir en la convicción de que muchas ideas subjetivas (las “opiniones” expresadas bajo forma de link/voto) se convierten por hechizo en su exacto contrario, en objetividad revelada, apenas superan un determinado número convirtiéndose en mayoría. De este modo el ranking se transforma en expresión de calidad porque explícitamente fruto de una tecnología y manipulación de la información.

Pero ¿cómo es posible que la cantidad se convierta en calidad? Se asume subrepticamente que la mediación técnica del algoritmo sea garante de “objetividad”, y se asocia a la objetividad la califica de “bueno”, más aún de “mejor”, “verdadero”. Todo tiene que ser rápido, mejor inmediato, transparente, gracias a la eliminación del factor tiempo en la búsqueda y al estudio ergonómico de la interfaz.

El mecanismo de creación del consenso que Google estima ser manifestación de la “democracia directa” de los usuarios/voluntarios no convence principalmente por dos razones: por un lado asume que la mayoría tiene siempre razón; por el otro, implica que las opiniones de la mayoría, para ser disfrutadas por los usuarios, tienen necesariamente que pasar a través de una mediación técnica que nunca se manifiesta.

La dicotomía objetivo/subjetivo, sobrepuesta a la de la verdad/opinión, es del todo inadecuada en el mundo de las redes. Para ser precisos, la ciencia siempre ha creado híbridos naturales-culturales, es decir ha inventado técnicas y promovido tecnologías. Por un lado la observación y la experimentación con el método experimental tiene como campo de acción la “naturaleza”, y en este sentido es “objetiva”; por el otro, el fruto de la ciencia es altamente subjetivo. La ciencia está de hecho sometida a la voluntad y a la percepción político-social, en cuanto mediada por el lenguaje (cuando menos, el lenguaje de la comunicación científica), y en cuanto fuente de poder (desde la bomba atómica en adelante).

La tecnología que implementa las redes es la aplicación contemporánea del método científico para crear “híbridos de natura y de cultura”, enésimos objetos científicos, que luego se llevan cada vez más como testigos de la realidad en lugar de los humanos¹⁰⁵. Ya tiene más valor el resultado del algoritmo de PageRank, un híbrido tecnológico, respecto al de una persona cualquiera, y a menudo incluso respecto a la opinión de un experto del sector. Si no por otra cosa, por el hecho de que PageRank está siempre a portata de un click, un experto no.

¹⁰⁵ A propósito de la visión de la tecnología como “puente” de hibridación entre natura y cultura (extremidad teóricas de la esquizofrenia del moderno), véase la obra de Bruno Latour, y en particular *Non siamo mai stati moderni – Saggio di antropología simétrica*, Elèuthera, Milano, 1994. Un punto de vista parecido es el de Gilbert Simondon, que elabora un discurso sobre la “técnica transindividual”, un espacio mestizo constituido de individual y colectivo al mismo tiempo.

Como subrayado en apertura, desde un punto de vista materialista, después de todo, Internet es natural, en el sentido de material, es decir que está constituida por máquinas mecánicas y electrónicas; a la vez, es cultural, porque no existiría sin el sentido que la cultura le atribuye, constituida por interacciones significativas de los seres humanos, o mejor, de las máquinas biológicas entre ellas y con las máquinas electrónicas. El carácter híbrido de las redes es una consecuencia obvia del carácter híbrido de la tecnología *tout court*.

Otro punto de vista posible sobre la cuestión subjetividad/objetividad tiene que ver con el modelo de decisión: cómo se decide lo que es relevante. En un contexto relativista, es fácil concebir una información procedente de un sitio (sea un blog, Google, o un circuito de información oficial) como “objetivo” si este juicio de valor procede de una toma de postura clara, de una clave de lectura hecha pública, de la visión del propio pequeño mundo localizado. Una Red de confianza, es decir un grupo de personas que intercambia informaciones, pareceres y en general conocimiento, puede comunicar fácilmente el propio modo de trabajar, la propia eventual jerarquía interna y las modalidades para adherir a aquella específica Red o proyecto. Toda vez que se consulten los resultados de esta Red de confianza, podrán ser leídos como “objetivos”, en el sentido de verdaderos de por sí, relevantes para el propio conocimiento, precisamente porque fruto de tantas subjetividades distintas y de una fuerte interacción entre ellas. Si nos sentimos en sintonía con esta comunidad se podrá considerar interesante la información hallada o descartarla en función de otras redes de confianza.

Según este enfoque, si Google comunicara de forma pública el propio mecanismo de decisión y los usuarios de Internet fuesen en grado de comprenderlo se podría superar fácilmente el problema objetivo/subjetivo, llegando a identificarse cada vez, búsqueda tras búsqueda, con la Red que más nos satisface, e influenciarla directamente para que nos refleje, con nuestro gustos, ideas, idiosincrasias.

Público y privado

El PageRank ejemplifica también otra dicotomía: la que se da entre público y privado. En los hechos, todo lo que pasa a través de Google se publica, se hace público: ¿a quién no le ha pasado de encontrar e-mails personales en los resultados del ranking, igual porque habían sido enviados a mailing list públicas? Ya que una cantidad cada vez mayor de informaciones personales pasa a través de Google, y en general a través de las tecnologías informáticas, la posibilidad de reencontrar mañana las propias llamadas (tal vez realizada vía VoIP), archivadas y repetibles a través de un motor de búsqueda, parece todo menos que remota. En suma, las tecnologías llegan a hibridar incluso las esferas de lo público y de lo privado; por otra parte, conectarse a Internet significa abrir una puerta al mundo, y a través de esta ventana seguramente el mundo entra en nuestra vida.

Existen ya algunas redes cuyas prácticas desautorizan de hecho la quimera de una información objetiva, cumpliendo una elección precisa y deliberada, y del todo subjetiva, entre lo que quieren hacer público y lo que tiene que ver con la esfera privada. Este fenómeno toma cuerpo en su forma más extensa cuando el motor de búsqueda se revela incapaz de satisfacer una instancia de búsqueda cuya calidad (específica) va más allá de la disponibilidad cuantitativa y de la estructura técnica propuesta.

El ejemplo más conocido es la búsqueda tramite network Peer to peer (p2p)¹⁰⁶ como eMule u otros parecidos. La cantidad de datos que se pueden buscar en estas redes corresponde con los datos compartidos por los usuarios, y varía en el tiempo de forma no periódica; por esta razón se les define como redes “transientes”¹⁰⁷. En efecto, en estos casos el usuario realiza libremente una elección entre el material que desea hacer público y lo que, por otra parte, se quedará sellado en la esfera privada; también porque una regla básica del intercambio “a la par” es la oferta y el compartir los propios materiales, que se pueden buscar en el network, para conseguir a cambio otros materiales. En sí el p2p es una práctica legal; al límite, los datos intercambiados pueden estar protegidos por copyright (audio, video, textos, ecc.): sólo en ese caso se entra en la ilegalidad. La difusión planetaria de estas práctica pone de manifiesto la naturaleza liberticida de la legislación vigente en materia de propiedad intelectual y de patentabilidad del conocimiento: actualmente es de hecho reato penal difundir informaciones protegidas por el copyright, es decir un buen cuarto de la población italiana debería acabar en prisión por haber descargado al menos un file mp3 ilegalmente. Además, parece cada vez más urgente la elección individual sobre la frontera entre público y privado. Se abre un amplio abanico de posibilidades que se mueven entre la elección ponderada de favorecer e impulsar el intercambio de informaciones más allá de cualquier consideración legal y el simple aprovechamiento a la fuente de la piratería, aprovechando lo que se halla en la Red.

Piratería difusa no significa ciertamente estar a las puertas de la revolución popular, también porque a menudo se trata de piratería semi inconsciente, no de una asunción de responsabilidad por parte de los individuos que eligen oponerse al vigente sistema de protección del conocimiento¹⁰⁸. Más bien las tecnologías digitales transforman la reproducibilidad técnica en un problema obsoleto, y la cultura del consumismo empuja a desear mucho más de lo que se puede utilizar. Es casi obvio desear todo lo que conseguimos imaginar, por lo menos a nivel de “informaciones”, por cuanto sea evidente que nunca seremos capaces de escuchar ni siquiera una fracción de la música que descargamos de la Red, o de ver incluso solo una porción significativa de las películas que guardamos en nuestro hard disk. Pero la instancia del deseo, por su propia naturaleza ilimitado, conyugada con la posibilidad técnica, pone serios interrogantes sobre la distribución y el acceso al conocimiento. El carácter tendencialmente gratuito y libre de estos intercambios constituye un reto extraordinario a la primacía de la economía productiva, pero la difusión de las opiniones, por ejemplo bajo forma de blog, pone en crisis incluso la concepción tradicional de los mass media.

¹⁰⁶ Literalmente, p2p significa “de par a par”; en general, se trata de un modelo de comunicación en el que cada una de las partes tiene las mismas funcionalidades y cada una de las partes puede empezar la sesión de comunicación, frente a otros modelos como el server/client o el master/slave. A menudo la comunicación P2P se ve implementada ofreciendo a cada uno de los nudos de comunicación las funcionalidades de server y client. En el lenguaje corriente el término Peer to Peer se utiliza para describir las aplicaciones mediante las que los usuarios pueden, a través de Internet, intercambiarse directamente files con otros usuarios. En particular, por lo que respecta Internet, p2p es un tipo de network “transiente” que permite a un grupo de personas que utilizan el mismo programa, conectarse y acceder directamente a los recursos compartido. Napster hasta hace algún año, y ahora Emule, Lopster, ecc. son ejemplos de tales software.

¹⁰⁷ El término “transiente” sinónimo técnico de “transitorio”, procede de la astrofísica y de la acústica y se refiere a una fuente cuya radiación varía en el tiempo y de forma no periódica. Las Redes “transienti” son transitorias desde el punto de vista de los flujos informacionales que las componen; en el caso de las Redes p2p, la “transienza” depende de la variación de la cantidad de informaciones compartidas por cada individuo.

¹⁰⁸ Y sin embargo, violar las leyes sobre el copyright es un reato increíblemente difuso como grave: un reato penal! También en Italia se ha acogido la normativa europea conocida como EUCD, modelada sobre el americano DMCA, del 2003, veáse <http://www.softwarelibero.it/progetti/eucd/>

El p2p representa de hecho solo el nivel superficial, y más difuso, del intercambio no mediado por una autoridad genéricamente sobre-comunitaria. Hay muchas otras instancias cualitativas de búsqueda, todas satisfechas por network de confianza y capaces de recombinar la perspectiva de nuestra orientación en Internet. En algunos casos estos recorridos subjetivos convergen en sistemas de afinidad cultural o profesional, por ejemplo forum, newsgroup, blog especializados, o en el disenso explícito de las fuentes oficiales; no solo, a menudo constituyen un ejemplo suficientemente articulado de un modelo alternativo para la organización del conocimiento (*knowledge management*).

Es por lo tanto fácil hipotizar una (tal vez lenta) evolución del circuito de conocimiento relacionado con el instrumento del blog hacia instrumentos p2p. Intercambio de conocimiento a través de programas desligados de la publicación web y basados en network dinámicos en los que sea posible compartir flujos de informaciones y file entre los usuarios conectados, por lo que realmente pertenecientes al propio circuito de amigos. Esto permitiría hacer que la división entre público/privado no sea algo estático, sino dinámico y difuminado, en el que cada persona es libre de compartir informaciones escogiendo un nivel de accesibilidad con un valor intermedio entre público y privado.

Líneas de fuga: media independientes, criptografía, blog, foaf

El fenómeno de los blog¹⁰⁹, sitios web que contienen reflexiones personales y link del autor y comentarios de los lectores, ha dado lugar a la que ya se define como blogosfera: actualmente se estiman cerca de 60 millones de blog con casi 4 billones de link, y más de 1,2 millones de post diarios. La blogosfera aumenta de más de cien mil blog al día y se redobra cada seis-siete meses¹¹⁰.

Desde el punto de vista matemático los blog respetan la ley de potencia y se distribuyen según la característica “cola larga”: pocos centenares de blog tienen un número muy elevado de link (cerca de cuatro mil constituyen el gotha de los blogs), la mayor parte (millones) tiene poquísimos link. En este sentido, como hemos visto en el Capítulo 2, la blogosfera pertenece al marco de la economía de la búsqueda. El blog permite compartir el propio punto de vista, declaradamente subjetivo y personal, sin ningún filtro: el número de link a un blog representa su popularidad y por lo tanto su autoridad, que puede alcanzar y superar la de los periódicos y de otros media tradicionales, generando fenómenos de influencia sobre la opinión pública. Confiabilidad, confianza y reputación están ligadas al particular extensor del blog: en la medida en que crece la popularidad se hace difícil mentir sin ser penalizado inmediatamente, es decir desertado, y pronto terminar en el olvido¹¹¹.

La autoridad del blog de Beppe Grillo, el único italiano que aparece entre los primeros 100 blog del mundo por número de link, es mayor de la de los blog de Repubblica y Corriere della Sera. El sujeto Beppe Grillo se expone a nivel personal y no pretende vender verdades: expone su punto de vista. En este sentido, los blog crean redes compartidas autogestionadas; no solo, a

¹⁰⁹ <http://technorati.com> (Noviembre 2006)

¹¹⁰ Véase la relación La blogosfera e i media, www.casaleggioassociati.it

¹¹¹ Existen varios métodos para maquillar los datos relativos a los link a un blog, por ejemplo las link faro (grupos extensos de páginas web creadas para aumentar la popularidad de una página), o los fake blog, o sea blog creados de forma automática por programas, igual aprovechando recopiladores de noticias tipo RSS. Google intenta desde hace tiempo solucionar este problema, que genera el fenómeno de los clic fraudulentos afectando, al menos en parte, el mecanismo publicitario.

veces pueden convertirse en las únicas fuentes de información no alineadas. Es el caso del blogger iraquí Salam Pax (alias Salam al-Janabi) durante la segunda guerra del golfo¹¹².

La novedad más importante ligada a la información en los blogs es la de la agregación automática de fuentes distintas a través de los *feed* RSS (literalmente: “provisiones de RSS”), el estándar *de facto* para la exportación de los contenidos Web. En pocas palabras se trata de un método automático para deslizarse rápidamente en sitios web, encontrar y exportar las informaciones que interesan. La popularidad de los blogs es tal vez una de las razones principales del éxito de RSS: millares de weblog empiezan a producir contenidos en RSS y así proliferan sitios que recogen una selección de post de los blogs más seguidos (los llamados *blog aggregator*) y programas para usar los contenidos de un blog directamente en el propio terminal (*RSS reader*). De esta manera es posible buscar contenidos en la blogosfera sin pasar por Google.

La posibilidad de poder recibir cómodamente de forma automática los últimos artículos escritos sobre los argumentos que el usuario considera más interesantes en el propio ordenador es desde luego una innovación destinada a tener enormes repercusiones en el uso de la Web. Más en general, ya que el RSS es un primer paso para disponer de toda información en un formato que por su propia naturaleza se puede compartir fácilmente, transformar y ampliar, hace que una información sea accesible a todo soporte informático, sea un sitio, un programa instalado en el pc, un móvil o cualquier otro instrumento tecnológico.

Sin embargo, existen muchas otras posibilidades para buscar informaciones. Si Google se presenta como un medio objetivo y público, Indymedia.org¹¹³ por ejemplo se define como “Indymedia is a collective of independent media organizations and hundreds of journalists offering grassroots, non-corporate coverage. Indymedia is a democratic media outlet for the creation of radical, accurate, and passionate tellings of truth” donde sin embargo existe un grupo de personas que actúa según una policy pública muy específica: en la columna a la derecha de las páginas, llamada newswire, cualquiera puede publicar. No se censura nada, sin embargo “los mensajes patentemente racistas, sexistas, fascistas se ocultan, sin ser eliminados”. Indymedia es esencialmente un instrumento para usuarios activos en la creación de la información y para contribuir a la formación del conocimiento, y de verdades compartidas: la autoridad que “crea” las verdades es difusa y participada, es una autoridad ética (colectivo) y no matemática (algoritmo).

Poner en discusión las fuentes oficiales demostrando que es posible hacer información en modo creíble, descentrada e independiente es uno de los objetivos de Indymedia.org y de decenas de network surgidos alrededor de temas específicos¹¹⁴.

¿Publicar las informaciones o mantenerlas privadas? Una respuesta interesante procede del proyecto Freenet.org¹¹⁵, una Red p2p descentralizada, creada para resistir a la censura, que explota los recursos (banda pasante, espacio en disco) de sus usuarios para permitir la publicación y la fruición de cualquier tipo de información. Freenet.org se ha constituido pensando en el anonimato y en la seguridad, no en la velocidad de transmisión. Las naciones y autoridades de distinta naturaleza que censuran las comunicaciones todas comparten una característica en común: existen unos controladores que deciden qué cortar y qué mantener, lo

¹¹² La vida en Bagdad bajo los bombardeos en el blog de Salam Pax: http://dear_raed.blogspot.com/

¹¹³ www.indymedia.org; <http://www.indymedia.org/es/index.shtml>

¹¹⁴ Por ejemplo: <http://www.cyberjournalist.net>; <http://alternet.org/>; www.warnews.it; <http://zmag.org/>; <http://slashdot.org>; <http://wikipedia.org>; <http://digg.com>, ecc.

¹¹⁵ Para entrar a formar parte de la Red Frente: <http://freenetproject.org/>

que se considera ofensivo y lo que se considera inocuo. Freenet.org es una Red que elimina a quienquiera la posibilidad de imponer su escala de valores: nadie puede decidir lo que es aceptable, la tecnología se pone al servicio de la libertad absoluta de expresión.

En las redes no anónimas, si queremos que nuestras informaciones lleguen exclusivamente a destinatarios preestablecidos, la criptografía pesada se presta a nuestro caso – y, por otra parte, es una práctica fundamental para garantizar un mínimo de privacy en las comunicaciones digitales. Por ejemplo nuestros correos electrónicos encriptados aunque se puedan buscar mediante los instrumentos de búsqueda, no son legibles¹¹⁶. La encriptación se propone usar las redes de forma subjetiva y filtrada, creando espacios privados de intercambio entre particulares en lugar de espacios públicos indexables y aprovechables por cualquiera. No se trata de tener algo para ocultar, sino de preservar espacios privados y de poder decidir autónomamente y libremente qué, cómo y cuándo hacerlo público; nuestras comunicaciones personales, nuestros datos sensibles no serán públicos ni objetos de marketing por el sólo hecho de estar transitando en la Red.

Es posible también hacer que todo sea público, mejor: decidir publicar, que sigue rimando con hacer buscable y legible, la subjetiva visión de uno mismo. La propia identidad virtual (alter-ego digital) puede ser definida en los mínimos detalles y no dejada en mano de las técnicas de personalización. La propuesta, quizás más interesante la representa la comunidad FOAF, acrónimo de *Friend-of-a-friend*¹¹⁷, o sea “Amigo de un amigo”. El objetivo de este proyecto es crear un conjunto de páginas web legibles automáticamente por las máquinas (*machine-readable*) que describen las personas, sus intereses, las cosas que hacen, y sus conexiones recíprocas¹¹⁸. De esta forma el concepto mismo de verdad, de fácil presa pasa a un segundo plano, sustituido por el de “confianza”: tiendo a confiar en los amigos de los amigos, formando “redes de confianza” basadas en los gustos, afinidades, pasiones en común. FOAF es un instrumento útil para crear network sociales, promocionando una actitud activa: es preciso “declararse”, “describirse”, “publicarse”¹¹⁹.

Desde el punto de vista de las redes de confianza (*trusting networks*, base de la Web Trust) FOAF puede ser una garantía de transparencia sobre los recorridos que unen nuestros avatares virtuales, favoreciendo un crecimiento de credibilidad y confianza en las relaciones interpersonales desarrolladas en la Red. La Web se utiliza para crear espacios en los que las personas están interconectadas en un gran mapa relacional, que se puede recorrer a partir de cualquiera de sus nudos y sin la necesidad de un instrumento de búsqueda centralizado.

¹¹⁶ La criptografía le da miedo a los gobiernos, porque genera espacios privados no inmediatamente accesibles al control policiaco: no por nada la legislación estadounidense consideraba, al menos hasta hace pocos años, “armas” los instrumentos de encriptados pesados. Actualmente, la legislación en merito es muy confusa. Naturalmente la posibilidad de desencriptar un documento depende de la potencia de cálculo que se emplea. No se trata pues de seguridad absoluta, pero sí racionalmente proporcionada a la evolución tecnológica. Así utilizar claves de encriptación asimétrica pesada como el GPC (*Gnu Privacy Guard*) a 204 kb ofrece, en 2006, una protección muy elevada, aunque no total. The Gnu Privacy Guard: <http://www.gnupg.org/>

¹¹⁷ www.foaf-project.org

¹¹⁸ Más técnicamente, FOAF es un vocabulario RDF/XML para la web semántica con el objetivo de traducir en un lenguaje comprensible a las máquinas (por ejemplo los spider) lo que usualmente está escrito en una página personal sobre el autor de un sitio o de un proyecto.

¹¹⁹ Escribir el propio documento FOAF no es difícil utilizando instrumentos como Foaf-a-matic, <http://www.ldodds.com/foaf/foaf-a-matic.it.html>

Es posible usar instrumentos como FOAF también para fines declaradamente políticos: por ejemplo, el network indyvoter, que se define como “un sitio web de networking social para activistas políticos”, con el objetivos explícito de formar una clase de gobierno alternativa¹²⁰.

Las falsas alternativas binarias, natura/cultura, subjetivo/objetivo, público/privado, verdadero/falso se revelan en su poquedad. La informática es constitutivamente dúplice e impura: teoría y práctica a la vez, crea objetos que modifican nuestra forma de pensar e, utilizados, modifican la misma informática. Los mundos digitales son reales, y en la realidad las visiones en blanco y negro dejan sitio al gris, y a la infinita variedad de colores, a los matices de las opiniones, a las diferencias que son riqueza, a lo heterogéneo que no puede ser reconducido a lo único. Los instrumentos para modificar, expandir y ramificar estos espacios de libertad y estas líneas de confin están a nuestra disposición: las posibilidades están limitadas solo por nuestra superficialidad y falta de imaginación.

¹²⁰ www.indyvoter.org: “The League of Young Voters* is doing matinal voter-organizing out of New York City. Our mission is to engage pissed off 17-35 year olds in the democratic process to build a progressive governing majority in our lifetime”

Conclusiones

Hemos llegado al final de nuestra exploración. Desvelado muchos, más o menos notables, secretos de Google.

Como hemos visto se trata de un dispositivo que presume poder gestionar globalmente el complejo sistema de los conocimientos disponibles en Internet. En efecto vende por verdades objetivas respuestas que son sólo recorridos subjetivos filtrados por sus tecnologías de búsqueda. No tenemos que caer en la trampa de una lectura esotérica del fenómeno, estupefactos por la rapidez en atender nuestras solicitudes. Estos portentos, a menudo impregnados de misticismo son en realidad estrategias que conjugan el uso de sistemas avanzados de recogida y almacenaje de datos con sistemas de localización de las informaciones relacionados con técnicas de personalización directas e indirectas con la perfilación de la publicidad.

No solo, marketing de vanguardia y una oportuna gestión de la comunicación son la suma de la evangelización googoliana: el empleo de colores elementares del espectro visivo del logo es un ejemplo por todos. Añadid la difusión de las interfaces altamente configurables, capaces de conservar la propia *corporate-identity* en todo ámbito, y el plato está servido: la explotación a todo nivel de la economía relacional puesta en marcha para el usuario. Finalmente Google coopta sistemáticamente metodologías de desarrollo cooperativo típicas del Open Source y del Free Software, abatiendo los costes para la implementación de sus propios servicios y haciéndose pasar por el que apoya la causa de la libre circulación de los saberes.

Rechazar la hagiografía corriente del “motor de búsqueda perfecto” no significa empezar una campaña de boicot; por otra parte, el grupo de Ippolita ha usado a menudo Google, incluso para escribir este libro.

Del mismo modo, apelar a menudo a recursos difusos y accesibles como Wikipedia se ha revelado muy útil a nivel enciclopédico. En efecto, si ya se conoce el argumento en discusión, se puede comprobar la exactitud de las informaciones de forma independiente, recombinando las partes de la Red de modo laico y desencantado, para hablar de la misma Red. El uso crítico de las fuentes depende de la capacidad de los sujetos para aceptar y evaluar la seguridad de las informaciones, no de la bondad intrínseca de las tecnologías digitales.

La informática no es simplemente una técnica para gestionar la información de forma automática, sino que posee una lógica intrínseca, es decir que trabaja y modifica continuamente sus mismos cimientos. Es física, teórica y experimental a la vez: estudia la formalización del lenguaje (formaliza el conocimiento pues), la aplica a las componentes físicas de la electrónica, consigue lenguajes que a su vez influyen las teorías del conocimiento, se fundamenta en una dimensión recursiva del todo particular.

Google desarrolla a fondo esta lógica recursiva: es una máquina extraordinaria que se constituye mediante su mismo uso por parte de los usuarios. En este sentido, es una máquina “autopoiética”, que acumula todas las informaciones de base volcadas cada día en la Red por millones de usuarios (nombres, fotografías, mail, preferencias de búsqueda, participación a forum, uso de blog, compilación de cuestionarios y módulos, recorridos de navegación, etc.) para vender publicidad de modo capilar. Los datos de los usuarios se han convertido en un enorme patrimonio económico, social y humano. Un patrimonio a proteger, desde luego, pero también un territorio en el que curiosear, indagar, experimentar.

Google facilita una respuesta fluida a las intenciones de búsqueda de los usuarios, una respuesta articulada en servicios cada vez más sofisticados y personalizables. Pero se trata de una multiplicidad solo aparente, que mira simplemente a difundir una forma de consumismo adecuada a la economía informacional: la personalización de masa de las publicidades y de los productos. El capitalismo de la abundancia de Google practica un cuidadoso fichaje del imaginario de los productores-consumidores (*prosumer*), a todos los niveles. De hecho, los

usuarios facilitan gratuitamente sus datos personales, pero también sugerencias e impresiones de uso de los servicios; los desarrolladores colaboran a la afirmación de los instrumentos “abiertos” puestos a disposición para difundir los estándares de Google, que quedan bajo el atento control de Mountain View; los dependientes de Googleplex y de los demás *datacenter* se reconocen plenamente en la filosofía empresarial de la excelencia.

La personalización del imaginario no es sino la última etapa del proceso de colonización capitalista de las redes que hemos llamado onanismo tecnológico. La mentalidad del provecho se oculta bajo declaraciones a favor de la “libre expresión de los individuos”, para luego aprovecharse de aquellas “expresiones” para vender brillantes e inútiles productos personalizados.

Google promueve su “economía de la búsqueda” como si fuera una nueva ciberdemocracia en la que centenares de millones de personas pueden comunicar directamente y **autorganizarse**, sustrayéndose al control estatal e institucional en general, a través de sus instrumentos tecnológicos. Este mensaje simplista halla el apoyo convencido de los media y de los intelectuales “democráticos” de todo el globo. Según esta visión, la naturaleza de Internet sería intrínsecamente democrática: en Red, no solo los individuos serían empujados a tomar el lugar de las instituciones, sino que incluso las mismas instituciones serían mejores. El entusiasmo tecnocrático llega hasta el punto de presentar la informatización de la administración pública, un proceso conocido como *e-government*, como una especie de gobierno sin ideologías, que llama en causa la responsabilidad del “ciberciudadano”. Esta identidad nueva por fin participa en primera persona (digitalmente) de la emersión de una opinión pública difusa. Como si fuese posible hacer frente a la crisis de las formas clásicas de representación mediante una democracia local conectada sobre escala global. Hemos constatado las principales lagunas de este tema, que se pueden reconducir a un enfoque ideológico. La idea es que las tecnología son “neutras” por definición, y que esta presunta neutralidad es sinónimo de bondad, en cuanto fruto de una práctica de búsqueda científica objetiva, capaz de ofrecer a cada individuo lo que desea, sin esfuerzo y rápidamente.

La compleja mediación informacional puesta en acto por Google aparece como una membrana transparente de alta tecnología, garante de las elecciones libres de los usuarios/ciudadanos/consumidores, que usan/votan/compran sumergidos en la libre Red gestionada por Google por su propio bien.

A despecho de todo sueño participativo, deslumbrado por el cuento de la ciberdemocracia, pero falto de sustancia concreta, no se pueden realizar formas realmente autónomas de democracia directa mediante la centralización de las informaciones, de los saberes, de los poderes, en las manos de una empresa privada (ej. Google), y menos aún en las de un sujeto público (ej. el Authority para las telecomunicaciones).

Ni siquiera las puntas más avanzadas de los movimientos alter-mundialistas escapan de la trampa de la identidad. Desean de hecho una recomposición de la identidad de clase a través de una centralidad renovada del trabajo, en este caso telemático. Pero se quedan lejos de la esfera del deseo individual incluso cuando promueven el *social networking* como solución taumatúrgica para todas las frustraciones personales, en un ritual de autoayuda tecnológica global.

Solo una elección de autoformación puede constituir una vía de fuga a recorrer para substraerse al dominio tecnocrático. Hay que dar muchos pasos antes de poder “poner en común” algo de sí para generar sinergias. Sin una preparación técnica adecuada, el sedicente maná comunitario se revela pronto un ejercicio solipsista.

Los individuos que crean las redes tienen que aprender a gestionar sus propios datos sensibles y empezar a establecer qué hacer público y qué mantener privado. Además, deben ser capaces de decidir cuáles informaciones son “verdaderas” y cuáles son “falsas”, evaluándolas en base al propio punto de vista subjetivo. En efecto, mientras recorren la Web, tienen que ser conscientes de que modifican el paisaje informacional y de que cambian ellos mismos en el curso de la navegación. Así los individuos pueden desarrollar su propia autonomía, o sea elaborar unas reglas para cruzar los territorios virtuales y adquirir una postura personal.

A la par de cualquier otra tecnología, las tecnologías digitales no son ni buenas ni malas en sí, y como hemos visto ni siquiera neutrales: depende del uso que se hace de ellas y del método con el que se desarrollan. Desde luego, siendo híbridos capaces de amplificar algunos aspectos de la vida real, permiten desvelar las contradicciones entre la “Naturaleza” y la “Cultura”.

Hay que conjurar otro peligro, pues: la imagen de la Web como experiencia desmaterializada, carente de concreción física, que conduce a menudo a un rechazo luddista y reactivo de la innovación.

En esta perspectiva, la realidad virtual del ciberespacio bulliría de interacciones poco significativas, operadas por personalidades online sin memoria de las disparidades materiales de la vida concreta: sexo, raza, riqueza, posición social serían apartadas en la navegación fluida y sin atrito de identidades ficticias.

Esta idea –solo aparentemente materialista- es la que sostienen intelectuales y observadores que, desde lo alto de sus cátedras, opinan sin haber tenido nunca la humildad de preguntar a los “niños” nacidos a la vez de las tecnologías digitales.

Por el contrario, la realidad virtual es física hasta el punto que no podría existir sin las máquinas mecánicas, de silicio y circuitos, y sin las máquinas biológicas, es decir los usuarios: conectada, múltiple y desterritorializada no son sinónimos de “inmaterial”. Además, una actitud semejante oculta un profundo miedo de ver cambiar el mundo sin conseguir encontrar el propio lugar y una desconfianza radical en las posibilidades de transformación y crecimiento de los individuos.

Las tecnologías digitales representan, pues, una posibilidad de liberación solo si vinculadas con el desarrollo de alter-ego digitales complejos, conscientes, capaces de interactuar de forma imprevisible. Es posible utilizar múltiples lenguajes para construir un terreno de encuentro: entre otros, Ippolita ha considerado indispensable el método experimental de la investigación científica, la riqueza inagotable de la tradición humanista, la fuerza dialógica de las pasiones políticas, la moralidad de la narración del método feminista, la curiosidad sin límites típica de la actitud del hacker. La confianza en la posibilidad de modelar tecnologías en base a los deseos de los individuos es indispensable para crear redes de verdad libres, digitales y no solo.

El caos de los mensajes contradictorios, el ruido de fondo a veces insoportables y la amplitud casi inconcebible de la Red pueden sin lugar a duda provocar temor, pero la exploración acaba de empezar.

BOX / APÉNDICES

Apéndice I – El fin del mundo en una galleta

El plazo de Google es el 17 de enero de 2038. Esta fecha es interesante porque es precisamente dos días antes del fin del mundo. El 19 de enero de 2038 es una fecha esencial para todo el mundo POSIX, que comprende todos los sistemas operativos procedentes del sistema UNIX (GNU/Linux, BSD, Solares, ecc.) y todos los protocolos de Red UNIX style (http, ftp, etc.). En los sistemas POSIX el tiempo está regulado por razones históricas por el número de segundos transcurridos desde el primero de enero de 1970. La unidad de medida para este número es un “**intero marcado**”, valor que en informática ocupa en memoria 32 bits.

Si un programador crea una variable de tipología “**intero marcado**” para memorizar un valor numérico, esto puede ser como mínimo -2147483648 y como máximo 2147463647. Un número muy grande, pero que se convierte en un valor muy pequeño si lo transformamos en segundos. En 32bits solo caben 136 años.

Cuando Ken Thomposon, casi por juego, inventó Unix, nunca hubiera imaginado que iba a revolucionar tanto la informática y sobre todo nunca hubiera pensado que una serie de sus elecciones (el filesystem, la hora, la codificación de los caracteres, las llamadas a función, el lenguaje de programación, etc.) iban a convertirse en ley para todo programador del globo en muy poco tiempo. Thompson decidió en 1972 que la fecha 0 de su sistema operativo sería el 1 de enero de 1970.

2147463647 segundos después del 1 de enero de 1970, será exactamente el 19 de enero de 2038 a las 3:14 de la mañana, último segundo del mundo Unix, de internet (que funciona gracias a protocolos Unix). Después de ese segundo estaremos proyectados en el 13 de diciembre de 1901 a las 20:45. Seguramente éste será un problema que tendrá que ser gestionado en los próximos años y requerirá un cambio que marcará una nueva época en la gestión del tiempo y de todo lo demás en los sistemas Unix. Sin duda un problema algo más serio y fundamentado del tan decantado “Millennium bug” en el paso entre 1999 y 2000.

El plazo de Google es pues la máxima fecha que se puede configurar en un *cookie* y que conlleva que nuestro browser de nuestros sistemas operativos, no eliminará nunca aquel *cookie* y las informaciones ahí guardadas. Para mayores informaciones sobre el problema de las fechas se vean:

- <http://www.2038bug.com/index.html>
- <http://www.iee.org/Policy/Areas/SCS/problemdates.cfm>
- http://en.wikipedia.org/wiki/Unix_epoch~~

Apéndice II – Interzonas: influencias y dominios en los mundos digitales

Los proyectos Open Source de Google están hospedados también en el portal sourceforge.net; no hay cambios sustanciales, ni siquiera cambios gráficos relevantes, en el paso entre el sitio de Google.com y el de sourceforge.net. Los distintos proyectos se quedan inalterados casi en los mínimos detalles, tanto en el código, como en el interfaz. El único elemento digno de nota es el cambio del server hospedante: de hecho se pasa de una sección con el nombre del proyecto en el sitio de code.google.com a una página hospedada en sourceforge.net en un dominio de tercer nivel, por ejemplo nombreproyecto.sourceforge.net. Existen, pues, dos accesos separados e independientes para alcanzar la misma Interzona. Sin embargo, estas zonas pertenecen a dominios distintos: Google.com y sourceforge.net

El concepto de dominio está a la base de la estructura de las redes. En ámbito informático, con el término “dominio” se entiende el nombre de un server dentro de una escala jerárquica extremadamente compleja gestionada por una Red de ordenadores llamados DNS (Domain Name Server). Estos ordenadores se encargan de gestionar la asociación server – dominio y de permitir a todo el mundo de alcanzar unas direcciones esparcidas en la Red gracias a su nombre.

Por ejemplo, cuando se quiere llegar a www.google.com, un programa específico dentro de nuestro ordenador pregunta a uno de los DNS qué server responde al nombre www.google.com. En sustancia, la Red de los DNS se encarga de buscar dentro de gigantescas mapas y bases de datos el server que le pedimos, asociando la dirección alfabética que nos resulta comprensible a una dirección numérica (el número IP) que es comprensible para las máquinas. Nos comunica luego el resultado, y nos conecta a Google o a cualquier otra dirección solicitada.

Simplificando mucho este complejo mecanismo jerárquico, podemos decir que los dominios están estructurados por niveles. En una dirección web típica, los niveles se leen de derecha a izquierda. Los primeros niveles son pues los genéricos .com, .org, .net, y los nacionales como .it, .es, .fr, .de, .tv. En el segundo nivel hallamos los nombres reales de los server, asociados a uno solo de los dominios de primer nivel: google.com, eleuthera.it, somasuite.net, alekoslab.org., etc.. Nada le impide a un server tener más dominios de segundo nivel ligados a distintos dominios de primer nivel: google.it, google.de, google.com son todos dominios de segundo nivel válidos. No es obligatorio que estos dominios de segundo nivel se refieran a una misma entidad: por ejemplo, ippolita.net es el sitio del proyecto ippolita, ippolita.it es el sitio de la poetisa Ippolita Avalli, Ippolita.com es el sitio de Ippolita que fabrica y vende joyas.

Desde el segundo nivel en adelante la gestión corresponde a quien gestiona el segundo nivel. Por ejemplo, code.google.com está gestionado directamente por Google, así como www.google.com, y también moon.google.com.

No existe una regla precisa que impone la forma en la que se dan los nombres de tercer nivel: nos baste pensar en una dirección evocativa, pero absolutamente correcta, como del.icio.us: el sufijo .us indica que pertenece al dominio nacional USA, el nombre “icio” indica el server, el prefijo “del” de tercer nivel completa el juego de palabras, de forma que la dirección total signifique “delicioso” en inglés. Sin embargo, existe la costumbre de asociar algunos dominios a servicios específicos. Una dirección del tipo www.server.net hospedará tendencialmente la primera página del servicio de server.net, mientras mail.server.net hospedará servicios de correo electrónico.

La imagen tal vez más adecuada para describir esta jerarquía es la de una casa: los servicios tienen una puerta de acceso directo (mail.server.net) mientras que el contenido de los distintos servicios puede ser hospedado dentro de una de las habitaciones interiores (www.server.net/fotografie/), o sea en una de las directory del server principal.

La gestión de los dominios de tercer nivel de Google es incongruente, y por ello extremadamente abierta y fluida. De hecho es curioso notar que a la Luna se la ha concedido un dominio de tercer nivel (moon.google.com) mientras que a Marte no, relegado en una directory secundaria (www.google.com/mars): ambos estos servicios permiten explorar mapas, sin diferencias significativas en el servicio que proponen.

La estructura de Sourceforge.net es más rígida, así como la de Freshmeat.net (que Google de momento no utiliza) y de otros proyectos análogos. Sourceforge.net ofrece los proyectos hospedados como servicios para los usuarios por su misma naturaleza de portal; por esta razón, todos los proyectos de Google se encuentran en un dominio de tercer nivel. Se trata de un instrumento para las comunidades de desarrollo que dedica a cada una de ellas el ambiente *ad hoc* rico en servicios. Las distintas comunidades piden los servicios que necesitan y Sourceforge.net les da un espacio completo y personalizado. Por lo tanto Google se comporta como la comunidad de Gaim o de cualquier otro software: y, sin embargo, Google no es precisamente una comunidad cualquiera de desarrolladores independientes....

Otras interzonas se crean para gestionar los campos de influencia de las correspondientes comunidades y la eventual interacción entre proyectos que de ellas dependen. El proyecto Open Source Mozilla y la página predefinida para su browser Firefox es un pequeño ejemplo de gran visibilidad pública. Cada vez que un usuario instalará el browser Mozilla-Firefox la primera página online que verá al empezar internet será la página de Google modificada gráficamente de forma que incluya también el logo de Firefox.

Google facilita también toolbar para que se integren en Firefox. Las Toolbars (barras de herramientas) son bloques gráficos (en general en forma de línea o columna) que, gracias a una serie de botones y otros objetos gráficos, activan funciones particulares del programa que los hospeda. Las extensiones para Firefox permiten añadir toolbar gratuitas dentro del propio browser. Lo que el público difícilmente sabe es que tales productos no son para nada Open Source. En efecto, la toolbar de Google no es un producto descargable de los clásicos sitios de extensiones para Mozilla, porque bajo una licencia propietaria. El funcionamiento de la toolbar está oculto dentro de una pequeña librería usada solo por este instrumento y disponible sin código fuente. Existen distintos estudios ilegales de "*reverse engineering*" que muestran todo funcionamiento de esta extensión y la forma en la que se solicitan datos del sitio de Google.

La integración entre el propio terminal (ordenador, palmtop, ecc.) y los Google-server es cada vez más estricta; de hecho una de las novedades entre las extensiones para Mozilla Firefox se llama Google Browser Sync: bookmarks, cronología, cookies, password guardadas e incluso fichas y ventanas dejadas abiertas durante la última sesión, se guardan online en los server de Google y luego sincronizadas automáticamente en las distintas posiciones usadas y en las nuevas sesiones de navegación. El control sobre el usuario se extiende a cualquier dispositivo capaz de acceder al web.

Finalmente, podemos individuar una ulterior interzona en el frente de las licencias de uso. Google ha intentado adecuarse al estándar de las demás empresas Open Source que tienden a utilizar versiones más o menos retocadas de la BSD 2.0 (excepción hecha por la Novell). Esta

licencia permite compartir el código pero también un fuerte control sobre los programas y sus versiones. Por un lado esto consiente la creación de comunidades, mientras que por otro deja todos los márgenes para un control general con respecto a la global hilera productiva. Cualquier programador independiente que desarrolle gratuitamente, utilizando esta licencia, podrá ser legalmente expropiado en cualquier momento de su propia creación.

Por mucho que algunos responsable de Google hayan comunicado un posible cambio de licencia a favor de la licencia Apache, MIT, o incluso GPL, la BSD queda siendo la licencia estándar para todos los proyectos hospedados y promovidos por el motor de búsqueda. BSD constituye, por lo tanto la interzona de patente útil para atraer a sí los códigos libres del mundo Open Source y hacer de ellos propiedad exclusiva de la propia empresa.

LICENCIA



Reconocimiento-NoComercial-CompartirIgual 2.0

Usted es libre de:

- copiar, distribuir y comunicar públicamente la obra
- hacer obras derivadas

Bajo las siguientes condiciones:



Reconocimiento. Usted debe reconocer y dar crédito al autor original.



NoComercial. Usted no puede utilizar esta obra para fines comerciales.



CompartirIgual. Si usted altera, transforma o crea a partir de esta obra, sólo puede divulgar la obra resultante bajo los términos de una licencia idéntica a ésta.

- Para cualquier utilización o distribución, usted debe informar claramente a terceros sobre los términos de la licencia de esta obra.
- Cualquiera de estas condiciones puede ser modificada si usted consigue el permiso del titular de los derechos de autor.

Su derecho a un uso justo y legítimo de la obra, así como otros derechos no se ven afectados por lo anterior.

Esto es un resumen legible por humanos del [Texto Legal \(la licencia completa\)](#).

[Descargo](#)

Puedes consultar y descargar la versión digital del libro del sitio

<http://www.ippolita.net/google/>