

Sir Tim Berners- Lee

Creó la Web
y ahora trabaja en Internet 2.0.

MARK FRAUENFELDER. FOTOGRAFÍAS DE ASIA KEPKA

CREAR LA WEB (WORLD WIDE WEB) no hizo a Tim ni rico ni famoso instantáneamente. En parte porque la Web se desarrolló empleando tecnologías relativamente humildes. La invención de Berners-Lee se basó en un programa de búsqueda de información denominado Enquire, llamado así por el título de un libro de la época victoriana, *Enquire Within upon Everything*, que había escrito en 1980 como programador contratado por el Centro Europeo de Investigación Nuclear (CERN) en Ginebra, Suiza. Y también en parte porque Berners-Lee, cuando más de una década después terminó de escribir las herramientas que definieron la estructura básica de la Web, hizo lo que nadie hubiera podido imaginar: las regaló, con la bendición del CERN y sin ninguna contrapartida. Mientras otras personas han ganado millones con sus invenciones, este discreto programador se empeñó en fundar el *World Wide Web Consortium (W3C)* en el MIT, que aún dirige, para promover el desarrollo y los estándares globales de la Web. Berners-Lee está obteniendo al final su recompensa: en julio de 2004 la reina Isabel II le nombró caballero y el mes anterior recibió el *Millennium Technology Prize*, galardón finlandés dotado con un millón de euros, concedido "por los relevantes logros tecnológicos, que directamente promueven la calidad de vida de las personas, se basan en valores humanos y apoyan un desarrollo económico sostenible".



La nueva casa de la Web:
Tim Berners-Lee delante de su
nueva y colorista oficina, en el
Stata Center del MIT, diseñado
por Frank Gehry

Ahora, en las nuevas oficinas del MIT del Centro Ray y Maria Stata diseñado por Frank Gehry, este inglés de 49 años se ocupa de revisar centenares de proyectos del W3C. Además está personalmente involucrado en el desarrollo de su segunda gran idea: la Web semántica, que añade etiquetas identificativas a la información de las páginas Web y crea uniones entre ellas para que los ordenadores puedan descubrir datos de forma más eficiente y formar nuevas asociaciones entre elementos de información, creando de hecho una base de datos global y distribuida. A pesar de haber estado entre los objetivos de la invención inicial de Berners-Lee, la Web semántica ha estado 15 años en desarrollo y ha sufrido buenas dosis de escepticismo. De todas formas, Berners-Lee cree que pronto ganará aceptación, facilitando a los ordenadores la extracción de significado a partir de información dispersa, con la misma facilidad con la que en estos momentos Internet permite unir documentos individuales.

La Web semántica - unida a otras especificaciones y herramientas que se están desarrollando en el W3C, incluyendo estándares de accesibilidad para personas discapacitadas y software para dispositivos móviles - es una parte de la gran visión de Berners-Lee de una "única Web de significado sobre todos los temas y para todo el mundo". Pero, ¿es que estamos tejiendo una red enmarañada? A pesar de su optimismo sobre el futuro, a Berners-Lee le preocupa que cambios no suficientemente pensados en la organización y la gestión de la Web puedan comprometer su funcionalidad inherente y su "universalidad". El padre de la World Wide Web compartió con nosotros sus preocupaciones y sus sueños el día antes de volar a Helsinki para aceptar el premio *Millennium*.

TECHNOLOGY REVIEW: Durante varios años, Vd. ha promovido algo que llama Web semántica, pero la gente no parece sentirse muy interesada. ¿Por qué no?

TIM BERNERS-LEE: No es la primera vez que he tenido este problema de cambio de paradigma. Al principio, la gente no entendía realmente por qué la Web era interesante. La veían a pequeña escala, y a esa escala no es interesante. Lo mismo ocurre con la Web semántica.

TR: ¿Cómo supera esto?

B-L: En estos momentos estamos justo empezando a poner aplicaciones en la Web semántica, una a una, y uniéndolas donde nos parece útil. Pero lo que resulta más atractivo es el efecto de red. La visión es que llegaremos a una masa crítica, donde todas las cosas empiecen a estar integradas en un todo increíblemente grande. En ese momento, el incentivo para añadir más cosas crece de forma exponencial porque el valor de lo que ya está ahí también lo hace.

Como inicialmente pocas personas están muy motivadas para conectar una gran cantidad de datos en la Web semántica, todo debe hacerse por personas que están convencidas y que entienden que merece la pena hacer un esfuerzo para conseguir que el tema despegue.

TR: Entonces, por favor explíquenos: ¿Por qué merece la pena todo este esfuerzo inicial?

B-L: La amenaza habitual de la Web semántica es que existe una gran cantidad de información: información financiera, información meteorológica, información de empresas, en bases de datos, hojas de cálculo y sitios web que se puede leer pero que no se puede manipular. La cuestión es que estos datos existen, pero los ordenadores no saben qué son y cómo se relacionan. No se pueden escribir programas para utilizarlos. Sin embargo, cuando exista una Web con datos semánticos globales e interesantes, será posible combinar los datos que Vd. sabe que están disponibles con otra información de la que Vd. no conoce su existencia. Nuestras vidas se enriquecerán con esos datos a los que antes no teníamos acceso y podremos escribir programas que ayuden realmente porque serán capaces de entender datos de cualquier origen, en lugar de simplemente presentarlos en la pantalla.

TR: ¿Cómo interpreta la Web semántica los datos?

B-L: Supongamos que Vd. está explorando la Web y que encuentra un anuncio de un seminario y que decide asistir. Por supuesto, existe todo tipo de información en la página, que es accesible para cualquier persona, pero el ordenador no sabe lo que significa. Así que Vd. debe abrir un nuevo registro en su calendario y pegar allí la información. A continuación debe coger su agenda de direcciones y añadir nuevos registros con todas las personas relacionadas con el seminario. Y finalmente, si Vd. desea completar la tarea, debe encontrar la latitud y la longitud del local donde se realizará el seminario y programarlas en su dispositivo GPS (Sistema de posicionamiento global) para que le sea fácil llegar.

Es muy laborioso hacer todo esto a mano. Lo que a Vd. le gustaría es ser capaz de decirle simplemente al ordenador: "Voy a ir a ese seminario". Si existiese una versión de esa página en Web semántica, llevaría incorporadas etiquetas de información que le dirían al ordenador "esto es un evento" y su fecha y hora. Y añadiría automáticamente el viaje a su agenda, incluiría también los nombres de las personas en su directorio y programaría su GPS para dirigirle allí. Mantendría la relación entre el evento y diferentes personas que intervienen. Y esas personas dispondrían de páginas personales en la Web semántica que contendrían información sobre cómo puede Vd. contactar con ellas.

De este modo, su agenda de direcciones puede pasar de ser un repertorio cerrado de datos privados a convertirse en una visión de datos de personas relacionadas en todo el mundo.

PEQUEÑA HISTORIA DE LA WEB

1945

En el *Atlantic Monthly*, Vannevar Bush, director de la oficina de Investigación y Desarrollo Científico de Estados Unidos, describe el Memex, un dispositivo hipotético para enlazar microfichas documentales.

1968

Un estudio del Instituto de Investigación de Stanford muestra un "oNLine System" (NLS) entre cuyas prestaciones se incluye el hipertexto para los desplazamientos y la edición y el correo electrónico. Para utilizarlas inventa el ratón.

1980

Tim Berners-Lee, consultor en el Centro Europeo de Investigación Nuclear, CERN, escribe un software que permite enlazar documentos electrónicos.

1990

Berners-Lee apoda su programa global de hipertexto "WorldWideWeb". Número de sitios webs existentes: uno.

TR: Entonces, ¿la Web semántica automatiza simplemente muchas de las funciones que ejecutaría de forma manual un secretario?

B-L: No. Un secretario emplea criterios con una determinada inteligencia que no estamos reproduciendo aquí. El secretario tendría una capacidad mental humana que de forma inmediata le permite identificar relaciones en el conjunto completo de su experiencia. Por ejemplo: "Le he hecho una reserva para Tiawicha, porque ese fin de semana hay allí un concurso de flores. Lo he pensado y creo que....., bueno, quizás a Vd. le gustaría" es un proceso de pensamiento humano.

Lo que estamos elaborando se parece más a darle a Vd. un programa que es capaz de realizar todo lo que se podría hacer con programas que desarrollaría su departamento informático, pero para los que no tiene tiempo. Sin embargo, sigue siendo un programa. De la misma forma que la Web es sólo un documento.

En el futuro, la Web semántica será un gran sitio para desarrollar inteligencia artificial (IA) en el verdadero sentido de la palabra. Pero ahora mismo estamos haciendo algo bastante mecánico, a pesar de que empleamos elementos y piezas de la maquinaria desarrollada por la comunidad IA a lo largo de los años.

TR: Parece una tarea increíblemente compleja. ¿Cómo funciona la tecnología?

B-L: La tecnología de la Web semántica aborda el problema en dos etapas. La más elemental está en el ámbito común de los formatos de datos. Vd. puede coger una base de datos, un calendario, una libreta de direcciones, un extracto bancario, o un parte meteorológico, en fin, cualquier cosa que contenga datos puros, y conseguir que la máquina lo escriba en el lenguaje básico de la Web semántica, en lugar de hacerlo en alguno de los formatos específicos de una aplicación. Esto resuelve el problema "sintáctico".

De todas formas, esto no resuelve aún la parte "semántica". Para esta segunda etapa, la Web semántica primero da nombres a los conceptos básicos incluidos en los datos: fecha y hora, un evento, un cheque, una transacción, temperatura y presión, un lugar. Estos elementos se definen para significar lo que realmente significan en el sistema que genera los datos, por ejemplo: "La fecha de transacción que aparece en el extracto bancario", y así sucesivamente. A este conjunto de conceptos le llamamos ontología. Después, cuando existen conexiones entre ontologías, como cuando la fecha y hora de una fotografía es el mismo concepto que la hora de un parte meteorológico, escribimos reglas para obtener ventajas de esas conexiones. Esto nos permite, por ejemplo, solicitar al agente de la Web semántica fotos hechas en días soleados. Poco a poco, conexión a conexión, los datos se relacionan, se entrelazan. Lo más extraordinario es la reutilización de los datos para descubrir cosas nuevas: alguien coloca allí los datos para una finalidad concreta y otra persona los emplea de forma diferente.



1993

Marc Andreessen libera el Mosaic Web browser, que se convierte en la base del Netscape.

1994

Se funda el Consorcio World Wide Web (W3C). El número de sitios web alcanza 10.000.

Berners-Lee presenta la idea de la Web semántica

1998

El W3C libera las especificaciones del eXtensible Markup Language (XML) que permiten marcar el texto de las páginas Web con etiquetas descriptivas, críticas para la Web semántica

2000

A finales del año se han identificado 25.675.581 de sitios web.

2004

Se finalizan los estándares que permiten a los ordenadores intercambiar información de la Web semántica.

La reina Isabel II nombra caballero a Berners-Lee.

TR: Ha dicho que la "fase uno" de la Web semántica está terminada. ¿Puede explicarlo?

B-L: La forma de trabajar de la Web semántica es definiendo nuevos lenguajes para que los ordenadores intercambien información. La fase uno ha sido llevar estos primeros lenguajes, tanto sintáctica como semánticamente, a un estado que les permite convertirse en estándares respaldados por los miembros del W3C. Como la clave es la interoperatividad, no se puede hablar de aplicaciones de Web semántica si los programas se quedan trabajando solamente con sus propios formatos de datos, sin ser capaces de intercambiar datos con otros programas. Ahora existen los cimientos y el que quiera hacer una aplicación nueva y publicar datos lo puede hacer, y los programas de cualquier otra persona serán capaces de leer esos datos.

“ Lo más extraordinario es la reutilización de los datos para descubrir cosas nuevas: alguien allí los coloca con una finalidad concreta y otra persona los emplea de forma diferente.”

TR: ¿Qué tipo de aplicaciones se están haciendo para la siguiente fase de la Web semántica?

B-L: En las ciencias relacionadas con la vida están sucediendo cosas muy interesantes. Los grandes desafíos como el cáncer, el sida, y el descubrimiento de fármacos contra los nuevos virus, requieren la ayuda de una inmensa cantidad de datos de diferentes campos que se solapan: genómica, proteínica, epidemiología y otros más. Algunos de estos datos son públicos, otros son propiedad muy reservada de las empresas farmacéuticas y otros son información privada del paciente. El desafío de la Web semántica para conseguir la interoperatividad entre estos campos es muy grande, pero implica inmensos beneficios potenciales.

TR: Pero, ¿se trata sólo de un tema de intercambiar datos de una multitud de campos?

B-L: No. También existen desafíos en relación al mantenimiento de la privacidad y la propiedad intelectual, al mismo tiempo que se hace un uso efectivo de la información. Por ejemplo, cuando se investiga sobre un nuevo fármaco, se podría querer enlazar datos epidemiológicos con factores externos como el tiempo, viajes y demografía, para encontrar cómo se transmite una enfermedad y qué tipo de personas tienen factores de riesgo. A partir de aquí, se puede intentar conectar todo esto con un rasgo genético y empezar a buscar qué proteínas lleva asociadas y qué es lo que éstas bloquean o promueven en la biología de las células humanas. En consecuencia, uno podría querer conectar los elementos químicos que influyen en estas vías con síntomas de enfermedades y también con posibles elementos químicos que se pudieran emplear como

fármacos. Hay mucho que ganar y eso es lo que hace que mucha gente se muestre extraordinariamente interesada en trabajar con la Web semántica en el ámbito de las ciencias médicas.

TR: ¿Existe alguna aplicación que demuestre cómo la Web semántica puede realizar esas conexiones?

B-L: Si quiere jugar con la Web semántica, Vd. puede hacer un fichero amigo de un amigo (friend-of-a-friend file, FOAF). En un fichero FOAF (que contiene en formato estándar los datos que componen una página web personal), se puede publicar información variada sobre uno mismo, sobre su organización, sobre sus publicaciones, lugares y fotografías. Se puede tener un puntero que diga: "esta es mi fotografía" u otros datos sobre esa fotografía, como por ejemplo quién más aparece en ella.

Para crear un fichero FOAF, se debe rellenar un formulario como el que figura en www.ldodds.com/foaf/foaf-automatic.html. A partir de esta información se genera un fichero de texto, legible por la Web semántica, que se puede añadir a su sitio web personal. Existen sitios web semánticos que pueden extraer esa información y ofrecerle cosas como una lista de fotografías que le relacionan con otras personas. Yo estoy a tres fotografías de Frank Sinatra porque aparezco en una fotografía con Bill Clinton, que ha sido fotografiado con uno de los Kennedy, que a su vez, fue fotografiado con Frank Sinatra. Esto es una aplicación tonta, pero que muestra realmente la potencia de la reutilización de la información.

TR: ¿Puede describir un ejemplo más serio?

B-L: Es sorprendente contemplar la industria que está dedicada a implementar estos estándares. Juegos de herramientas de HP e IBM, aplicaciones de autor de Adobe, soluciones para la gestión inteligente de contenidos de Profium y Brandsoft, y motores de búsqueda de Network Inference están todos trabajando para crear una Web semántica a diferentes escalas. Éstas y otras tecnologías están siendo adoptadas por comunidades que a su vez revolucionan la forma en que esos grupos colaboran y se comunican. Esto es lo que está ocurriendo en las ciencias médicas sobre las que hablamos antes.

En el Reino Unido, el Semantic Web Environmental Directory (Directorio del Medio Ambiente de la Web semántica, SWED), es un prototipo de una nueva clase de directorios de organizaciones y proyectos de medio ambiente. En lugar de centralizar el almacenamiento, la gestión y la propiedad de la información, el SWED simplemente recolecta datos y los emplea para crear el directorio. Desde una perspectiva social, existe una aplicación apodada Fatcats de FoafCorp (un proyecto de Web semántica que extiende el formato "amigo de un amigo" a entidades corporativas) que le permite a Vd. seleccionar una empresa y le muestra quién está en su consejo, desplegando un gráfico de personas conectadas. Cuando selecciona una de las personas, le muestra todos los consejos a los que pertenece. Vd. puede empezar a explorar las esferas de influencia de la cultura corporativa americana.

Lo extraordinario de todo esto es cuando se encuentra que una de esas personas tiene un fichero FOAF y Vd. empieza a pasar de la cultura corporativa a la cultura personal, de ahí a las fotografías, luego a la información del tiempo, además a las reservas de vuelos, y después a las reservas de restaurantes y a descubrir qué tipo de vino se toma en la comida.

“Cuando seleccionamos a una persona, la web semántica muestra todos los consejos de administración a los que pertenece. Podemos así empezar a explorar las esferas de influencia de la cultura corporativa americana.”

TR: Vd. habla a menudo de la importancia de la "universalidad de la Web", ¿a qué se refiere?

B-L: Una de las propiedades fundamentales de la Web es el hecho de que es un solo espacio, y es un espacio de consenso. Debe ser independiente del hardware que se utilice. Debe ser independiente del software que se emplee o del sistema operativo sobre el que funcione. También debe ser independiente de la cultura en la que Vd. esté, de si Vd. está escribiendo un maravilloso documento, cuidadosamente editado, o de si garabatea algo en la parte posterior de un socorrido sobre. Y debe ser independiente del lenguaje que se esté utilizando, del juego de caracteres, de si sus letras van de arriba abajo, de izquierda a derecha, o de derecha a izquierda. Además, las personas deben poder acceder a esa información, incluso aunque tengan alguna discapacidad. En el W3C a este concepto le llamamos "una Web para todos, en todas partes, sobre cualquier cosa".

TR: ¿Y existe una amenaza a esta universalidad?

B-L: Había una propuesta para crear un dominio especial de alto nivel llamado ".mobi". Todos los sitios web que trabajasen con teléfonos móviles se colocarían en ese dominio; sería el lugar de los contenidos Web para los dispositivos móviles. Pero, por principio, debe existir una única URL o dirección WEB. Segregar los contenidos en un dominio ".mobi" es una forma errónea de hacerlo. Disponemos de todo tipo de estándares en el W3C para permitir a un sitio web funcionar de forma óptima cuando Vd. lo consulte desde un teléfono móvil o desde una pantalla grande. Pero, obviamente, si se pone un ".mobi" al final de un nombre de dominio, se está diciendo: "Esto es un sitio especial con cosas que se pueden ver en un teléfono móvil".

TR: ¿Qué pasa con otros dominios de alto nivel, ".biz", ".info"... , que se han propuesto aligerar el abigarramiento de nombres en el dominio .com?

B-L: Añadir nuevos dominios de alto nivel no ayudará. Lo que las personas recuerdan son la serie de caracteres entre "www" y ".com". Si se pone un ".info" o un ".biz" detrás, eso sólo les confundirá. Significa que tiene que recordar todo en lugar del nombre de la marca entre las "www" y el punto.

Además, por supuesto, existe un sistema de pago por el registro de transacciones financieras. Las pequeñas empresas o los

particulares que tienen un dominio pueden pensar que, a fin de evitar confusiones, deberían comprar estos otros. Sólo el pago anual para una familia es un añadido importante a la cuenta de Internet.

TR: Existe una lucha de poder entre Naciones Unidas y el ICANN, (Internet Corporation for Assigned Names and Numbers), que gestiona la asignación de los nombres de dominio y las direcciones de Internet. ¿Cuál es su opinión?

B-L: Algunos países están preocupados, con razón, porque el ICANN funciona bajo un contrato con el Departamento de Comercio de los Estados Unidos. Internet es algo internacional, y, a pesar de que pueda estar dirigida cuidadosamente por el ICANN teniendo en cuenta los mejores intereses de todo el mundo, existe en algunos países un sentimiento muy arraigado de que el estar financiado el ICANN por el gobierno americano significa que está controlado por los Estados Unidos y que esa situación no es la más correcta.

Mi impresión es que esta asimetría debe suprimirse cuidadosamente. Es importante que se aprecie que se trabaja de forma correcta. Sin embargo, el hecho es que el ICANN se puso en marcha y está en funcionamiento y no debería destruirse de repente. Hacer algo que represente de una forma equilibrada a todos los interesados, requiere sin duda una gran experiencia y una evaluación constante. Puede que el ICANN deba tener mayor financiación por parte de Naciones Unidas, pero no creo que deba convertirse de la noche a la mañana en un organismo de la ONU.

Se produce una gran confusión en esta área cuando se usa el término "gobierno de Internet". Se empieza hablando sobre nombres de dominios, que efectivamente es un área muy concreta, y se termina hablando sobre privacidad, derechos de autor, confidencialidad, condiciones comerciales, y todo tipo de elementos del sistema legal ordinario. Las personas no deberían de pensar que el ICANN dirige todo lo que sucede en Internet. ICANN sólo desempeña un papel muy concreto.

TR: ¿Cree que la World Wide Web será su colaboración más importante?

B-L: Mi papel ha tenido que cambiar necesariamente desde el de diseñador solitario, pasando por el agitador de comunidades, hasta el de arquitecto avanzado y facilitador de consenso en el W3C. Pero sospecho que la Web será mi contribución más importante, aunque eso requiriera estar en el lugar adecuado en el momento oportuno. El error, de todas formas, es pensar que esto está terminado. La Web semántica es sólo la aplicación de los principios de diseño de la Web a los datos; pasarán muchas decenas de años antes de que podamos decir que hemos implementado realmente la idea de la Web de una forma completa, si alguna vez llegamos a ello.

TR: Además de la Web semántica, ¿tiene algún otro sueño o deseo para el futuro de la Web?

B-L: ¡Pues sí, realmente muchísimos! Siempre he querido que la Web fuera un medio más creativo y flexible, con sistemas para anotaciones y grupos de editores y así sucesivamente. Estoy entusiasmado con los nuevos dispositivos portátiles que podemos usar en la Web, con las tecnologías basadas en el habla, y con muchas otras cosas. Una vez que has empezado con la idea básica de la Web, se hacen posibles una gran cantidad de temas. ◇