

La World Wide Web celebra vint anys amb ganes de continuar progressant tecnològicament. Una de les innovacions que marcarà més el futur de la xarxa serà la implantació dels cercadors semàntics, que permeten de relacionar conceptes de manera intel·ligent per oferir més bon servei a l'usuari. Ara bé, de moment, l'algoritme de cerca perfecta encara és lluny.

## Cercadors intel·ligents, el futur d'internet

**S**ovint, quan cerquem qualsevol cosa per internet, ens podem veure frustrats de no trobar allò que volem de la manera més ràpida possible. Ens hem hagut d'acostumar a formular la recerca seguint els paràmetres que marquen Google, Yahoo i Microsoft, que es reparteixen

el 95 per cent del total de consultes que es fan a través de motors de recerca. Per tal de prendre quota de mercat al gegant Google, que acapara el 80% de demandes a internet, les altres companyies inverteixen en innovació per donar més bon servei a l'usuari.

L'objectiu final és desenvolupar un algoritme intel·ligent que pugui interpretar les preguntes que presenta l'internauta. D'aquesta manera, el motor de cerca seria capaç de saber què es demana exactament i podria oferir la resposta més adequada possible. En poques paraules, es tracta de crear un cercador que relacioni els conceptes que desitja trobar qui fa la consulta. Un exemple pràctic: si l'usuari introdueix "amics de Dalí a Nova York", un motor de cerca actual donaria resultats per a tots tres conceptes. Un de semàntic, en canvi, oferiria pàgines que continguessin els tres criteris de cerca plegats.



EL CREADOR DEL WORLD WIDE WEB, TIM BERNERS-LEE (AL CENTRE), AMB BEN SEGAL I ROBERT CAILLIAU, DOS DELS SEUS COL·LABORADORS, DURANT LA CELEBRACIÓ DEL VINTÈ ANIVERSARI D'INTERNET.



Bill Gates, president de Microsoft. L'aposta de la companyia és el nou cercador Kumo (en detall).

Els cercadors semàntics formen part de l'anomenada web semàntica, encaminada a substituir, d'aquí a un cert temps, la web 2.0, caracteritzada per les xarxes socials i per la interacció entre usuaris. El puntal d'aquesta nova manera d'entendre internet mitjançant la relació entre conceptes és l'intercanvi d'informació de manera que l'entenguin les màquines. D'aquesta manera es pretén ampliar la capacitat dels sistemes informàtics d'operar entre si i també reduir la mediació dels operadors humans en els processos intel·ligents de flux de dades.

El pare de la idea és Tim Berners-Lee, creador del concepte *web* i del World Wide Web Consortium (W3C). Proposa que la xarxa semàntica serveixi per a ampliar la capacitat de la *www* mitjançant els estàndards, els llenguatges d'etiquetatge i més eines de processament de la informació relacionades. Allà on el llenguatge *html* (el més utilitzat a internet actualment) i l'hipertext es queden curts, la web semàntica combina tecnologies com l'*xml* (per a etiquetar conceptes) per a proporcionar descripcions que suplementen o reemplacen el contingut dels documents

web. L'etiquetatge de les dades dipositades a internet permet als gestors de la informació d'interpretar els documents i dotar-los de contingut.

Però de moment la fita a la qual aspiren els cercadors semàntics i, per extensió, la web semàntica, és llunyana. Tot i que algunes empreses ja fan servir, internament, alguns cercadors basats en les etiquetes i la indexació de continguts, extrapolar aquest procés a la World Wide Web no és un camí gens fàcil. Petits projectes competeixen amb les grans empreses, que saben que el futur és en el tren de la semàntica.

**Kumo, l'aposta de Microsoft.** Bill Gates i els seus no saben ser al segon lloc en cap àmbit tecnològic, i això inclou el dels cercadors. Amb un 4,71% de quota de mercat, la companyia de Silicon Valley cerca noves fórmules per quedar-se una part del pastís que ara és a les mans de Google. Només en matèria de publicitat, basada en gran mesura en els enllaços patrocinats —els que surten emmarcats en groc i en primer lloc en efectuar qualsevol recerca—, Google s'embutxaca 17.000 milions d'euros.

La guerra per a captar usuaris del cercador líder no ha fet sinó començar. Microsoft ja prova de manera interna el Kumo, el seu nou motor de cerca. Tal com explica una entrada del mes de març passat al bloc de tecnologia del *New York Times*, la companyia americana no pot continuar gastant milers de milions de dòlars per intentar batre Google només imitant-lo. Cal innovar, i un dels vice-presidents de l'empresa de Bill Gates, Yusuf Mehdi, vetlla personalment per la nova posició del cercador de Microsoft. És una mostra de la importància que es dona a la renovació del cercador de l'empresa líder de programari del món.

L'aposta pel cercador Windows Live no va donar fruits. Google es manté imbatible i això va fer veure que l'única aposta capaç de fer un xic llum al final del túnel era el progrés tecnològic. Un progrés que, segons que sembla, ha de situar Microsoft a l'avantguarda d'internet. El secret del nou Kumo és la incorporació de criteris de cerca semàntics. Això vol dir que s'estableixen relacions entre conceptes, de manera que la resposta que es dona a l'usuari és molt més concreta. Posem-ne un exemple: si es



**Stephen Wolfram, creador del sistema de computació simbòlica Mathematica, ha anunciat que el mes de maig llançarà el Wolfram Alpha, un cercador intel·ligent.**

vol saber, amb el Kumo (i amb qualsevol cercador semàntic) quants dòlars són 14 euros, és més fàcil d'escriure "14 euros en dòlars" que no pas "convertidor de divises". Amb aquesta última consulta hauríem de fer dos passos abans de trobar el valor dels nostres euros en moneda americana. El Kumo ens estalviaria temps i esforç, en aquest cas.

Des de Microsoft, però, s'insisteix que la majoria de funcions que es proven no s'aplicaran al mercat dels Països Catalans fins el 2010. El problema bàsic dels cercadors semàntics és que funcionen molt bé en entorns tancats, com per exemple una empresa o una administració pública. El motiu és que els límits són clars i els conceptes, relacionats. Fer el salt a la web, amb tones d'informacions de tota mena, és la frontera que encara no s'ha superat. Així i tot, incorporant criteris semàntics ja s'arriba molt lluny, diuen els que l'han provat.

**Wolfram promet respostes exactes.** No tan sols les grans companyies s'afanyen per no perdre el tren de la tecnologia. S'han engegat diversos projectes que tenen l'eix principal en la innovació. Un d'aquests projectes és

el del físic Stephen Wolfram, creador del sistema de computació simbòlica Mathematica. De moment, ja ha anunciat que el mes de maig llançarà el Wolfram Alpha, un cercador intel·ligent que poques persones han presenciat. Wolfram promet que el seu motor de cerca és capaç de donar respostes exactes a partir de llenguatge natural. És a dir, allò en què tothom treballa però que ningú no és capaç de trobar.

Per a crear el seu motor de cerca, el físic britànic s'ha basat en dues tecnologies creades abans per ell mateix. La primera és l'esmentada, Mathematica, una eina que serveix per a representar qualsevol cosa per mitjà de llenguatge simbòlic. També permet de fer càlculs de tota mena gràcies a la seva traça a crear algorismes. L'altra tecnologia aplicable al cercador de Wolfram és l'NKS, que implica un paradigma per entendre com tota mena de complexitat es pot presentar a partir de normes senzilles. Armat amb aquestes dues eines, Wolfram es veu capaç d'aplicar regles de càlcul a tota la informació acumulada a internet. L'investigador britànic va més enllà de la web pròpiament semàntica, basada en etiquetes (*tags*, en

anglès) que serveixen per a organitzar els conceptes.

El físic britànic proposa –i el mes de maig veurà si ho ha aconseguit– de combinar Mathematica i NKS per tal d'aplicar mètodes i models, a manera d'algorismes, i així calcular immediatament la informació requerida per l'usuari. Tanmateix, el físic anglès adverteix que cada mètode i cada model, com també cada informació, té unes característiques pròpies. Malgrat el gran avenç que significa la combinació de tecnologies, Wolfram encara veu problemàtica la interacció entre usuaris o ordinadors, i viceversa.

El conflicte està en el fet que màquines i humans no parlem el mateix idioma. Les persones fem servir el llenguatge natural, i fer que els ordinadors l'entenguin és una tasca per a molts impossible. Encara som molt lluny del dia en què les màquines entenguin d'una manera sistemàtica grans volums fets amb llenguatge natural dipositats a internet. Així i tot, Wolfram no veu cap més solució que continuar treballant a fer entendre als ordinadors la informació elaborada per ser dipositada a internet. La màquina, doncs, s'ha d'adaptar a l'home i al coneixement que genera.

La solució que ha trobat Wolfram és a mig camí de fer entendre com parlen les persones als ordinadors. Si som capaços de fer que tot el coneixement sigui calculable, amb criteris matemàtics, ens podem estalviar la tasca de fer entendre als ordinadors, textualment, el llenguatge natural. N'hi ha prou, explica Wolfram, de representar les preguntes que la gent formula expressant-se amb normalitat mitjançant formes precises que encaixin amb els càlculs del motor de cerca.

El mateix físic britànic, però, veu molt difícil d'aconseguir aquesta fita, i afirma que el seu motor de cerca haurà d'estar en contínua evolució. La quantitat ingent d'informació que es genera cada dia necessitarà nous algorismes matemàtics per ser trobada pels usuaris mitjançant l'ús de llenguatge natural. En tot cas, al maig es veurà si, com promet, Wolfram Alpha sap donar respostes exactes a les recerques dels internautes.

*Oriol March Ledesma*