

IGOR LETURIA AZKARATE
Informatikaria eta ikertzailea

web

SEMANTIKOA

INTERNETEN BENETAKO AHALA IRISTEKO DAGO ORAINDIK

1990ean, Sir Tim Berners-Lee CERNeke ikertzaileak World Wide Web-a asmatu zuen, hipertestuaren kontzeptua eta Interneteko TCP/IP eta DNS protokoloak konbinatu. Asmakizun hark gu guztion bizitza eta mundu osoa aldatu zituen. Hala ere, aldaketa hori ez da ezer izango Berners-Lee azkenaldian lantzen ari den ideia berria, web semantikoa alegia, gauzatuz gero sortuko denaren aldean.

Interneten informazio eta aukera guztiak erabiltzea ez da beti erraza. Batetik, informazio-kantitate ikaragarria dago, eta, informazio-uholde horretan ito gabe murgiltzeko, ezinbestekoak ditugu bilatzaileak, hala nola Google edo Elebila. Baina, horiekin ere, sarritan ez da erraza behar duguna aurkitzea. Esate baterako, emaitza asko daudenean, ez da erraza bereizten, bilatzeko eman diogun hitzaren esanahi posibleetatik, guk nahi dugun adierari buruzko orriak zein diren ("Texas" hitza bilatzean, adibidez, AEBko estatuari, musika-taldeari, liburuari edo beste gauza askori buruzko orriak, denak batera, agertuko dira). Edo maiz ez dago modu onik informazio onaren eta txarraren artean bereizteko, bilatzaileek lehenengo posizioetan erakusteak ez baitu beti kalitatea ziurtatzen. Bestetik, mota bateko zerbitzu guztiak bateratzeko modurik ez dago kasu askotan, eta zerbait erosi nahi dugunean, adibidez, prezioak eta aukerak konparatzeko, hainbat dendaren webguneetan begiratu beste aukerarik ez dugu.

ARAZOEN ITURRIA, HTML FORMATUA

Arazo horiek guztiak amaraunaren hasierako diseinutik beretik datoz. Weba, finean, Interneten azpiegituran kokatutako hipertestu-bilduma bat besterik ez da, hau da, elkarri erreferentzia egiten dioten dokumentuen bilduma bat, eta, hipertestua kodetzeko, Berners-Leek berak sortutako HTML formatua erabiltzen da. Bada, formatu horren ezaugarriak eta mugak dira arazo guztien kausa.

HTML edo *HyperText Markup Language* testu-markaketarako lengoia bat da, alegia, testu bat edo haren zatiak nola deskribatu azaltzen duen kontentzio bat. HTML lengoian, deskribapen-markatze hori testuan bertan `<>` ikurren artean sartzen diren etiketen bidez egiten da. Adibidez, HTMLn, testu-zati bat lehen mailako izenburua dela adierazteko, `h1` etiketa erabiltzen da

(`<h1>Lehen mailako izenburua</h1>`), eta, letra etzanez idazteko, em etiketa (`kurtsibazko testua`). Horrelako beste etiketa mordoa daude, estekentzat, irudientzat eta abarrentzat. Nabigatzaileek markatze hori interpretatzen dute, eta erabiltzaileari orria modu egoki batean erakusten diote.

Baina HTMLko etiketa gehienak testuaren egitura eta itxura deskribatzeko dira, gizaki batek testua modu egokian ikusi eta ulertu ahal izateko. Ez daude pentsatuta makinaren tratamendu automatikoa hobetu edo errazteko. Eta bilatzaileak eta Interneteko beste tresnak makinak besterik ez dira. HTML etiketekin soilik ez dute informazio nahikorik ongi lan egiteko, eta testua bera ezin dute ulertu pertsonok egiten dugun bezala. Hala, testua eta HTML etiketak soilik izanik, bilatzaile batek ezin du jakin "Java" hitza agertzen den orri bat uharteari edo programazio-lengoiari buruzkoa den, edo, produktu bati buruz ari den orri batean, jendeak ongi edo gaizki hitz egiten duen hari buruz, edo orri bat produktu jakin bat saltzen duen on line denda batena dela...

WEB SEMANTIKOAK DU KONPONBIDEA

World Wide Webaren sortzailea bera ari da konponbidea bilatu nahian. Hain zuzen ere, web semantikoarena da Tim Berners-Leek zuzendutako World Wide Web Consortium nazioarteko erakundea azkenaldian gehien lantzen ari den gidalerroetako bat: hasierako diseinu haren akatsak zuzendu nahi dituen egitasmoa.

HTML formatuaren etiketatzeak dokumentuak eta haien arteko loturak deskribatzeko balio du. Web semantikoan, aldiz, objektuak, pertsonak eta horrelakoak, eta haien arteko erlazioak deskribatzen dira. Etiketek orriaren itxura eta egitura azaldu beharrean, orriko elementuen esanahia etiketatzen da. Web semantikoan etiketa



123RF

Etiketek orriaren itxura eta egitura azaldu beharrean, orriko elementuen esanahia etiketatzen da: musika-taldeak diren, pertsonak diren...

jakin batzuk leudeke, adibidez, musika-taldeak deklaratu eta deskribatzeko, beste batzuk pertsonak deklaratu eta deskribatzeko, beste batzuk pertsona batzuek musika talde bat osatzen dutela deskribatzeko, eta horrela gauza guztiekin.

Hala, bilatzaileek ezberdindu ahal izango lukete “Scorpions” hitza duten orrietatik zein diren musika-taldeari dagozkionak eta zein animalari, eta emaitzak esanahi posible ezberdin horien arabera taldekatuta erakutsi. Edo musika-taldeetan espezializatutako bilatzaileak sortu ahal izango litrateke. Edo on line denda ezberdinetan produktu bera saltzen dela nahiko erraz detektatu ahal izango litzateke, eta prezio-aukera denak batera

erakutsiko litzuketen zerbitzuak eraiki. Edo, antzerki-ekitaldiak eta non diren detektatuta, herrialde bateko antzerki-emanaldi guztien egutegi bat automatikoki osatu ahal izango genuke. Edo erabiltzaileek web orri edo produktu bati emandako puntuazioak semantikoki etiketatuta, bilatzaileek edo dendek puntuazioaren araberrako rankingak ere egin ahal izango litzukete. Eta horrelako beste milaka eta milaka aplikazio izango litzuke web semantikoak, oraindik imajinatu ere egin ezin ditugunak.

Beraz, web semantikoa izango litzateke pertsonentzat kodetzen den HTML sarearekin batera existituko litzatekeen sare paralelo bat, makinek ulertzeko moduko ezagutza-base bat, semantika adierazteko formatuetan kodetua. Lengoia naturaleko webaz gain, testu egitura-tuko beste bat izango genuke. Ezagutzaren errepresentazio hori makinek ulertu ahal izango lukete, modu eraginkorrean tratatu, ezagutza berria inferitu...

Zalantzarik ez dago web semantikoak aldaketa handia ekar dezakeela Internetera. Baina nola egiten da ahala errealitate bihurtzeko? Zer tresna behar dira? Eta zein dauzkagu?

Pertsonentzat kodetzen den HTML sarearekin batera existituko litzatekeen sare paralelo bat da web semantikoa, makinek ulertzeko moduko ezagutza-base bat. Horri esker, informazioa modu eraginkorrean tratatu eta gurutzatu ahal izango da. Web semantikoak adimena eranstendio Internet-sareari.

Jarraituko du...