

SEMANTİK WEB'E HOŞGELDİNİZ*

Tim Berners-Lee
Çeviren: Sevim Özdemir

Bugünlerde, hemen hemen hayatlarımızın her alanında sayısal bilgi yaratımı şaşırtıcı oranlara ulaşmıştır. Bütün verilerin arasında hastalıkların nasıl tedavi edileceğine, nasıl para kazanılacağına ve dünyamızı nasıl daha etkili biçimde yöneteceğimize ilişkin anahtar bilgiler bulunuyor. Ancak elimizin altındaki bu hazineyi analiz ettiğimiz, birleştirdiğimiz, paylaştığımız ve onu kullanırken gittiğimiz yolu biçimlendiren teknik araçlar ve sosyal pratikler fena halde geçerliliğini yitirmiş durumda. İyi haber şu ki (AJAX, XMI, RDF ve OWI gibi adları olan) birkaç teknik yenilik, verilere ilişkin yeni sosyal düzenlemelerle birlikte dünyayı saran ağı (world wide web-www) "semantik ağ (web)" diye adlandırdığımız şeye doğru ilerletmektedir.

Daha iyi bir veri birleşmesine doğru ilerleme dünyayı saran ağı (www) bu derece başarılı hale getiren aynı temel teknolojiyi kullanarak gerçekleşecektir: Bağlantı. Bugün, bilgiyi çabuk bulmayı da içeren ağın (web) gücü insanların dokümanları standart formatta yayınlamaları ve onları birbirine bağlamalarından kaynaklanmaktadır. Ağın (web) değeri bağlantıların sayısı ile birlikte doğrusal bir biçimde arttığından daha fazla artar; ki buna "şebeke/ağ (network) etkisi" denmiştir. Semantik ağ (web) da gücünü aynı yöntemle, ancak dokümanlardan ziyade verileri birbirine bağlayarak elde eder.

Daha iyi veri birleşimine olan ihtiyacı değerlendirmek için, dünya çapındaki ticari ve akademik farmasötik laboratuvarlarda üretilen deneysel verilerin büyüklüğüyle ilaç buluşunun sinir bozucu derecede yavaş ilerleyişini karşılaştırdım. Yaşam bilimi araştırmacıları çoğu durumda ilaç buluşu için gerekli bilgiyi içeren tek bir laboratuvarın, kütüphanenin veya genetik veri havuzunun bulunmadığı sonucuna varmaktadırlar. Bunun yerine, hastalıklar, biyolojik süreçler

ve geniş bir kimyasal katkı maddeleri dizisinin karmaşık etkileşimleri apayrı veri tabanları, çizelgeler ve dokümanları boyunca yayılır.

Sonuç olarak, daha iyi ilaç buluşlarına doğru giden süreç verilerdeki bağlantıların keşfine olanak tanınması için yapılan kurumsal pratiklerdeki değişimin yanı sıra verilerin birleşimini ve paylaşımını sağlayan teknolojilere bağlıdır. Bu tüm ilaç şirketlerinin basitçe verilerini salıverecekleri anlamında değildir, ancak daha ziyade onların



Semantik Ağ (Web), internette birbirlerinden "bağlantısız" bir şekilde bulunan milyonlarca veriyi bütünleştirecek. Semantik Ağ, verilerin değişik şekillerde "bağlamlanarak" depolanmasını amaçlıyor. Sistemde bilgiler, kavramlar etrafında birbirleriyle ilişkilendiriliyor.

* Tim Berners-Lee, The Economist-The World in 2007 içinde "Welcome to the Semantic Web".

kendi entelektüel varlıklarıyla diğerlerinin birleşimi aracılığıyla yaratılan daha büyük değerlere olanak veren daha esnek lisans modellerini keşfetmeleridir.

Bu tip bir birleşmenin gerçekleşmesi için gerekli teknik bir adım verilerin "semantik web" standartlarını (RDF, OWL, SPARQL) kullanarak yayınlanması ve veriyi ifade eden terimlerin tanımlarıyla birlikte onların birbirlerine bağlanmasıdır. Örneğin, daha geniş bir biyolojik süreçte belirli bir kimyasalın davranışına ilişkin deneysel sonuçlar yayınlandığında, biri biyolojik patikayı ve kimyasal tanımlamak için kullanılan kelimeleri göstermelidir. Sonrasında bir başkası, bu verileri -örneğin, araştırma literatüründe diğer deneylerin belgelediği verilerle- birleştirmek istediğinde bu kişi makalenin anahtar sözcükleriyle kimyasal isimleri karşılaştırmak için aynı kelimeleri kullanabilmelidir.

İçinde Para da Var

Daha iyi bir veri birleşimine ihtiyaç duyan ve bundan yarar sağlayacak olan sadece bilim insanları değildir. Finans-hizmet endüstrisini düşünün; başarılı yatırım stratejileri, artan biçimde çeşitlenen bilgi kaynakları grubundan (haberler, piyasa verileri, tarihsel eğilimler, ürün fiyatları) modelleri ve eğilimleri bulmaya dayanır. Önde gelen finansal bilgi sağlayıcıları şimdilerde kullanıcılara kendi sahip oldukları -kendi portföyleri ya da kendi iç piyasa modellerine ilişkin- bilgiler ile bilgi servislerince iletilen bilgileri birleştirmeleri olanağı veren hizmetler geliştirmektedirler. Biricik değer yaratımı birleşim sisteminin kendisindedir, kendi kendilerine edindikleri ham veride ya da hatta çoğu açık-kaynak bileşenlerini inşa edecek olan yazılım araçlarında değil.

Bu birleşimin anahtarı, tanımlayıcı sözcüklerle birlikte bilgiyi birbirine bağlayan ortak veri formatları kullanmaktır. Semantik ağ (web) herkesin aynı kelimeleri kullanmasını gerektirmez, artık

tüm dünyanın aynı dili konuşmasını bekleyemeyiz. Bunun yerine, bu sistem bir terimler bütünü bir diğerine çeviren araçlar sağlar. Bu çeviriler ağ (web) çapında çok çok daha büyük bilgi toplamalarını birleştirecektir.

Semantik ağ (web) teknolojisi özel kullanıcı toplulukları için önemli ölçüde fayda sağlayabilir. Fakat en heyecan verici buluşlar, çeşitli kaynaklardan düşen verilerin rastlantısal bütünleşmesi ve birleşmesinden gelecek. Semantik ağ (web) atılımını almış olan yatırım şirketi, örneğin, onların ilaç buluş başarılarına dayanan ilaç şirketlerinin piyasa değerlerinin analiziyle ilgileniyorsa, yatırım şirketince yapılan piyasa araştırması verileriyle farklı şirketlerin süre giden araştırmaya katkı yaptıklarını gösteren kanıtı birleştirmelerinin mümkün olduğunu görecektir.

Dünyayı saran ağın (World Wide Web) ilk aşamasında, yenilik ve yeni değer yaratımı değişen kurumsal pratiklere olduğu kadar bu tip yeni teknolojilerin basit gelişimine de bağlıdır. Teknoloji veri birleşiminin potansiyel yararlarını elde etmek için gereklidir, fakat yeterli değildir. Dolayısıyla, okuyucular 2007 için bu makaleden şu ev ödevini almalırlar: sorumlu olduğunuz verilerin bir envanterini yapın ve eğer onları birleşik bir iç ağ (intranet) veya semantik ağ (web) standartlarını kullanan genel ağda (internet) paylaşırsanız hangi kısımları büyük ihtimalle yeniden kullanırdı, bunun üzerinde düşünün. On yıl önce ağın (web) ilk günlerinde olduğu gibi veri birleşimindeki ilerleme de katkıda bulunan herkese faydalı olacaktır, genellikle de beklenmedik bir şekilde. ◀

WEB'İN BABASI

Ağın (web) babası ya da kaşifi olarak anılan Tim Berners-Lee, 1989'da HTML işaretleme dilini geliştirerek, Dünyayı Saran Ağ (www) olarak tanımlanan bilgi paylaşım sistemini kurmuş olan bilgisayar profesörüdür. Böylece bilginin açık bir şekilde herkese ulaştırılabilmesine olanak sağlamıştır. İlk ağ tarayıcısı yazılımını 1990'da geliştiren Berners-Lee, 1994 yılında ağ ile ilgili standartları dünya çapında belirleyecek açık bir kurum olan W3C'yi kurmuştur.

Londra'da doğan Berners-Lee Oxford Üniversitesi'nden 1976'da mezun oldu. Kaynak makinesi, TTL kapıları ve M6800 işlemcisi ve eski bir televizyonlar ilk bilgisayarını kendisi yapan Berners Lee, yazılım mühendisi danışmanı olarak 1980 yılında Cenova'daki Avrupa Parçacık Fiziği Laboratuvarı'nda (CERN) "Enquire" (Soruşturmak) adında, beynin yaptığı gibi rastgele ilişkilendirmelerle bilgi saklamaya yarayan bir program geliştirdi. Hiçbir zaman yayınlanmamış olan bu program Web'in gelecekteki gelişimi için bir temel oldu. Daha sonra world-wide-web (www-Dünyayı Saran Ağ) adlı metinötesi (hypertext) projesini ortaya koydu. Daha önceki çalışması olan Soruşturmak adlı programı temel alan bu proje, metinötesi belgelerle insanların bilgilerin birleştirerek çalışmalarına olanak tanımak için tasarlandı.

İlk www sunucusunu (httpd) ve www adındaki ilk istemcisini 1990'da yazan Berners-Lee, "HTML" olarak bilinen metinötesilerle belge formatlama dilinin ilk sürümünü yazdı. Bir kaç ay içinde İnternet'te yayınladı. İnternet üzerinden bir çok kişinin katkılarıyla Berners-Lee "URI, HTTP ve HTML"yi geliştirmek için çeşitli çalışmalar yaptı. MIT Üniversitesi'ndeki Bilgisayar Bilimleri Laboratuvarının Yöneticisi Michael Dertouzos'un desteğiyle WWW Konsorsiyumu'nu (World Wide Web Consortium) kurdu. Dünyanın çeşitli yerlerinden üniversiteler ve enstitülerden ekipler bu çalışmalara katıldılar. Time Dergisi tarafından 1999'da 20. yüzyılın en büyük 100 zekası arasında sayılan Berners-Lee, bir çok ödül ve ünvan sahibi oldu.