

AĞ 3.0 - ANLAMSAAL AĞ

Bekir Doğan - Erhan Kesken

Genelağ (İnternet) üzerinde oluşan sosyal ağlar, yaygınlaşan elektronik günlükler ve kişisel yayınlar ile birlikte çığ gibi büyüyen sanal veri yığınının çöp haline gelmemesi ve aramalarda tekelere bağımlılığın son bulması için, içeriklerin ve kavramların birbirleriyle ve kişilerle ilişkilendirilebildiği anlamsal bir ağın (semantic web) oluşturulması şart.

“web-3.0” ve “semantic web” çoğu zaman birlikte anılan kavramlar. Ağa (web) sürüm biçen resmi bir otorite olmamasına rağmen “web-1.0”, “web-2.0” ve “web-3.0” şeklinde isimlendirmeler, pazarlama kavramları olarak sürekli gündemde. Bu sürüm numaralarıyla ifade edilen dönemleri genelağ üzerinde yaşanan ekonomik değişimler sonucu atlanan çağlar olarak tanımlamak da mümkün. “web-1.0” döneminde genelağ üzerinden ticaret ile tanıştık. “web-2.0” ise, e-ticarete katılan e-topluluklar ve beraberinde gelen zengin ve dağıtık içeriğin ekonomik gidişatı belirlediği çağdı. Şirketlerin değerleri kullanıcı sayılarıyla ölçülür hale geldi. Zamanla kişiselleştirilebilirlik ağda yükselen bir değer haline geldi. Genelağdaki kimliklerin ve kullanıcı alışkanlıklarının giderek belirginleşmesi ile birlikte görünümür kişiye göre şekillendirilebilir kılındı, ancak sadece görüntüde değil işlevde de kişiselleştirme bir ihtiyaç halini aldı. Arama motorları, arama yapan kişinin kimliğine göre farklı sonuçlar getirme mücadelesine giriştiler. Peki ağ bu değişime gerçekten hazır mı? Halen ağdaki sayfamızın arama motorları tarafından fark edilmesi için özel anahtar kelimelerden faydalanıyoruz. İnsanın okuyacağı ile makinenin okuyacağı içeriği ayırmış durumdayız. Aramalarda gerçek anlamda bir kişiselleştirme için içeriği kişiyle bir şekilde ilişkilendirmek şart. Bir insa-

nın basitçe cevap vereceği soruları bir arama motoruna sorup da düzgün bir cevap almak çok mümkün değil, bu ilişki bilgisinin eksikliği nedeniyle. Genelağ üzerindeki içeriğin birbiriyle ilişkilendirildiği ve cümlelerle ifade edilebilir hale geldiği, genelağın dev bir veritabanına dönüştüğü, makinelere soru sorabildiğimiz, makinelerin birbirleriyle konuşarak sorularımıza cevap arayabildiği, özetle makinelere konuşmayı öğrendiği, servis ve sunucu merkezli yaklaşımların yerini kullanıcı merkezli dağıtık bir yapıya bıraktığı yeni çağ “web-3.0”, ve bu çağın temelini oluşturan altyapı da anlamsal ağ (semantic web) olarak tanımlanıyor.

İçerikleri anlamlı bir şekilde ilişkilendirebilmek için öncelikle sözlere ve bu sözlerin tanımına bakabileceğimiz sözlüklere ihtiyacımız var. Günümüzde ağdaki sayfamızın insan dışında da bazı ziyaretçileri var. Yayınlarımızın kolayca takibini mümkün kılan RSS okuyucular, gezegenler; yazdıklarımızın diğer insanlar tarafından bulunabilmesini sağlayan arama motorları mevcut. İnsan dışındaki en eski ziyaretçi kitlesi olan bu arama motorlarının gözüne çarpmak için sayfamızda insanlara hitap etmeyen anahtar sözcükler barındırıyoruz. İçerik hakkında içerik olarak da tanımlayabileceğimiz bu anahtar sözcükleri (metadata) belirlerken, insanların genelağ üzerinde bir arama yaparken edindikleri bazı alışkanlıkları takip ederek ilerliyoruz. Anlamsal ağlarda ise içerik (data) ile anahtar sözcüklerin (metadata) iç içe geçmesi hedefleniyor. İçerik ile içerik hakkındaki bilgilerin birlikte ifade edilebildiği yollar aranıyor. Bu durumda anahtar sözcükler şu anda olduğu gibi sayfanın geneline dair değil, içerik bazında bir ilişkilendirme imkanı sunar hale gelecekler. Amaç

sayfanızda tanımlanmamış herhangi bir içeriğin kalmaması. Peki bu tanımlar nereden geliyor? Tanımlara, işlevi dolayısıyla sözlük olarak nitelendirilebilecek belirtilerden (spec) bakılıyor. Aynı belirtileri kullanarak yazılan uygulamalar birbirlerinin dilini konuşabiliyorlar (FOAF (Friend Of A Friend) incelenebilecek güncel sözlüklere güzel bir örnek[8]).

Sözlükler oluştuktan sonra, sayfalarımızda kullanılagelen içerik-etiket eşleşmelerini, temel dilbilgisi kuralları üzerine kurulu cümlelere dönüştürebiliriz. İçerik ilişkilendirmelerimizi, kuracağımız cümlelerin yüklemi olarak düşünmek mümkün. İnsanların kişisel bilgilerini yayınladığımız bir sayfamız olsun. Seçtiğimiz sözlükte “kişisel bilgi” sözcüğünün tanımı yapılmış mı diye bakıyoruz, yapılmışsa cümlemizin öğeleri tamamlandı anlamına geliyor. Burada özne içeriğin kendisi, nesne sayfamız, yüklem ise “kişisel bilgidir”. İnsanın kolayca anlayabileceği bir cümle olarak ifade edersek, “bu içerik bu sayfada bir kişisel bilgidir” cümlesi karşımıza çıkıyor. “Kişisel bilgi” sözcüğünün sözlükteki tanımı şu şekilde yapılmış olsun: “ad-soyad bilgisinin mutlaka olması gereken, bunun yanında takma isim, genelağ adresi, e-posta adresi, posta adresi, enlem-boylam pozisyonu, fotoğrafı, sıfatı, notlar ve çeşitli kimlik numaraları gibi kişiye dair diğer bilgileri de içerebilen bilgiler bütünü.” Diyelim ki sayfamızda “Ali Veli” diye bir isim geçiyor, “Ali Veli” ismini bir “kişisel bilgi”nin içine dahil edip “ad-soyad” olduğunu belirttiğimizde ikinci cümlemizi kurmuş oluyoruz: “Ali Veli bu kişisel bilgiye dair bir ad-soyad”. Kullandığımız “kişisel bilgi” ve “ad-soyad” etiketlerini sayfanın gösteriminde de kullanarak, hem görsel hem metin olarak insan tarafından anlaşılır kılmak

çok zor değil. Bunun yanında “sayfa”, “kişisel bilgi” ve “ad-soyad” sözcükleri arasında kurduğumuz ilişki sayesinde hem insan hem makineler için anlam ifade edecek yeni cümleler türetmek mümkün hale geliyor. (Örneğin “Ali Veli bu sayfadaki bir insan ismidir.”)

Kurulan ilişkiler ve cümleler ile bir sayfa üzerinde tıpkı bir veritabanı sorgusuna benzer sorguların çalıştırılması mümkün. Şekil 1’de kıtalar ile ülkeler arasında “içerme” ilişkisini ve ülkeler ile şehirler arasında “başkent olma” ilişkisini tutan bir sayfa üzerinde çalıştırılan örnek bir SPARQL (SPARQL Protocol and RDF Query Language)[6] sorgusunu görüyoruz. Afrika’yla “içerme” ilişkisi kuran ülkelerin isimleri ve bu ülkelerle “başkent olma” ilişkisi bulunan şehir isimleri sorgulanmış. Günümüzde, veritabanlarının yaygın kullanımı nedeniyle, veri sorgulama denilince insanların aklına ilk olarak SQL sorgularının geliyor olması, SPARQL ile SQL arasındaki söz dizimi benzerliğinin sebebini ortaya koyuyor. Burada önemli olan bu tip bir soyutlamaya izin veren altyapının oluşturulması.

Sorgu dili soyutlamasındaki son nokta ise, iletişimi günlük konuşma dili seviyelerine çekebilmek. Örneğin şekil-1’deki SPARQL sorgusu yerine “Afrika’daki ülkeler ve başkentleri nelerdir?” diye bir sorunun makineler tarafından cevaplandırılabilir hale gelmesi anlamsal ağ çalışmalarındaki nihai hedef. Sadece insanın makinelere değil, makinelerin de birbirlerine soru sorabileceği bir altyapıdan söz ediyoruz. Şekil-1’deki sorguda tek bir sayfadan değil iki sayfadan veri çekildiği durumu düşünelim. Sayfalardan biri kıtalar ile ülkeler arasındaki “içerme” bilgisini tutsun, diğeri de ülkeler ile

şehirleri arasındaki “başkent olma” ilişkisini. İlk sayfa, “Afrika’da hangi ülkeler vardır?” sorusuna cevap verebilecektir, her bir ülke için diğer sayfaya “Bu ülkenin başkenti nedir?” sorusunu soracak ve edindiği bilgiler ile kullanıcının sorusuna doğrudan istediği yanıtları vermesi mümkün olacaktır.

Makinelerin kendi arasında haberleşmesi çok yeni bir şey olmamasına rağmen, bu haberleşmenin teknik olmayan insanlar tarafından anlaşılır hale gelmesi anlamsal ağlarla mümkün. Örneğin yukarıdaki örnekte sayfalardan biri diğerine ülke sayısı kadar “Bu ülkenin başkenti neresidir?” sorusunu soruyor. Burada şebekeyi (network) dinleyen insanın ne konuşulduğunu anlaması için herhangi bir protokolün belirtimini okumasına, okuduğu belgeyi anlayabilmek için uzun uzun eğitimler almış olmasına gerek kalmıyor, çünkü konuşulan protokol gündelik hayattaki da kullana geldiği basit dil bilgisi kuralları üzerine kurulmuş bir protokol.

Anlamsal ağa dair oluşumlar, diğer yandan günümüzde olduğu gibi merkezde servislerin ve sunucunun olduğu değil, kullanıcının yer aldığı bir yapıyı zorluyor. Örneğin e-posta hesabımızı tuttuğumuz servisle, kişi listemizi tuttuğumuz servisin aynı firma tarafından veriliyor olması gibi bir kısıtın kalmaması gerekiyor. Kişi listemizde bulunan adreslere, e-posta uygulamamızdan kolayca erişip hızlı bir şekilde mesaj gönderebilmek amaçlanıyor. Bunun için de kullanıcı kimliklerinin her servis için ayrı ayrı oluşturulduğu şu anki yapının yerini tek bir kullanıcı kimliğinin olduğu, kimliğin servislere tanıtıldığı bir düzene bırakması bekleniyor. Kullanıcılar

içerik sorgularken ve yayınlarken kullanacakları dile seçtikleri sözlüklerle karar verebilecekler. Dağıtık, sonu olmayan ve güvenli bir paylaşım ortamının oluşturulması anlamsal ağların gereksinimlerinden biri.

Anlamsal ağ çalışmalarına “W3C”[1] çalışma grupları ve ortaya koydukları standartlar yön veriyor. Ağı veriye dönüştürme amaçlı bu çalışmalar genel olarak iki odak altında ilerliyor. Bir odak, RDF (Resource Definition Language)[3] ve OWL (Web Ontology Language)[4] gibi dillerle anlam bilimsel cümleler kurulması ve uygulamaların bu cümleler güdümünde geliştirilmesi. “W3C”nin ortaya koyduğu standartlar ışığında RDF ve OWL yaratabilmek, yaratılmış olanlar üzerinde SPARQL ile sorgulama yapabilmek için geliştirilmiş pek çok araç var, pek çoğu da geliştirilmeye devam ediliyor. Anlamsal ağ ile ilgili yayınlanan belge sayısı ve çalışan geliştirici toplulukları gayet tatmin edici boyutlarda. Anlamsal ağ çalışmalarında bir diğer odak noktası olan GRDDL (Gleaning Resource Descriptions from Dialects of Languages)[5] ile var olan içeriği bir dizi değişime uğratan anlam bilimsel formatlara çevirme yöntemi üzerinde duruluyor (bkz. Şekil 2). Biçim imleri (tag) içindeki öznitelikleri (attribute) etiket olarak kullanıp, daha sonra bu etiketlere bakarak kolayca RDF’e ya da başka herhangi bir formata dönüştürmek mümkün. Bu etiket mantığı üzerine kurulu en bilinen yöntem, fikir babalığını “W3C”de söz sahibi Tantek Çelik isimli bir Türk’ün yaptığı “mikroformat”[7] yöntemidir. “Tails Export”[9] isimli “firefox” eklentisi sayesinde genelağda dolaşırken girdiğiniz sayfanın herhangi bir mikroformata uyumlu içerik bulundurup bulundurmadığını ve bulunduruyorsa bu içeriklerin neler olduğunu görebilirsiniz. Anlamsal ağ kavramlarından faydalanan sayfaları ve portalları incelemek için “smile”[10] projesi kapsamında geliştirilen ve firefox eklentisi olarak çalışan “piggy bank”[11] uygulamasını indirip deneyebilirsiniz. Anlamsal ağlarla ilgili şu ana kadar yapılan çalışmalardan örnekler vermek gerekirse, tanınmış sözlükler için;

```
PREFIX abc: http://example.com
SELECT ?capital ?country
WHERE {
  ?x abc:cityname ?capital.
  ?y abc:countryname ?country.
  ?x abc:isCapitalOf ?y.
  ?y abc:isInContinent abc:africa.
}
```

Şekil-1: SPARQL örneği

