

Bilginin Depolanması ve Organizasyonuna Giriş: Dijital Kütüphane Rafları

Giriş

Dijital kütüphaneler, raflarında çok çeşitli bilgi kaynağı topluluğu bulunduran bilgi merkezleridir. Teknolojik gelişme basılı ve elektronik ortamlarda kayıtlı bilgilerin birlikte depolanmasına olanak tanır. Dijital kütüphane rafları, aslında birbiriyle ilişkili bilgilerin bir arada sınıflandırıldığı dosya topluluklarından oluşmaktadır. Geleneksel formdan dijital dönüşürülen veya doğrudan dijital formda oluşturulan dijital bilgiler farklı dil, tür ve dosya yapılarında depolanırlar. Bu nedenle kullanıcıların geleneksel yapıdaki kütüphane raflarında bulunan bilgi kaynaklarına erişmek için kütüphaneyi ziyaret etmesine gerek yoktur. Aranılan bilgi ses, görüntü, video, fotoğraf, çoklu ortam veya metin yapılarından birinde veya birkaçında bir arada bulunabilir. Dijital kütüphanelerde yer alan bilgi kaynakları bir veri tabanında veya klasörde okunabilir halde depolanır ve bir arabirimle kullanıma sunulur. Böylelikle oluşan veri harmanındaki bu veri yapılarının tanımlanması (identifiye), yönetilmesi ve erişimlerinin sağlanması için dijital ortamda oluşturulmuş veya geleneksel katalogdan aktarılmış bir başka veri topluluğuna daha ihtiyaç vardır. İşte bu “veri hakkındaki yapılandırılmış veri topluluğu” metadata, diğer bir deyişle üstveri olarak tanımlanmaktadır. Üst verinin nasıl oluşturulacağı, standartları ve erişim protokollerinin bilinmesi depolanan bilginin organizasyonu için gereklidir.

Bilginin Depolanması ve Organizasyonu

Dijital kütüphanelerde bilgi, yaygın bir şekilde web sunucuları olarak adlandırılan bilgisayar sistemlerinde depolanır. Bir dijital kütüphane birden çok hiyerarşik derme grubuna bölünmüş bir koleksiyon serisini tek bir sunucu veya bir dizi dağıtık sunucu üzerinde depolayarak paylaşımına sunar. (Bkz 1. şekil). Ancak dijital ortamda kayıtlı bilgi sağlayan çok büyük yayınevleri web sunucuları ile kullanıcılarına aktardıkları çok geniş veri gruplarını yönetmek için “ilişkisel veri tabanları” nı kullanırlar. Bu türden veri tabanları birbirine bağlı veri gruplarının yine birbiriyle ilişkili tablolarından oluşan bir model yaklaşımı ile çalıştırılır. (Arms, 2001: 247).

Bir dijital kütüphanedeki veri harmanı ile ilgili hizmetler bilgi kaynaklarının depolanması, saklanması ve erişimlerinin sağlanmasına yöneliktir. Web sunucusu HTTP ve TCP protokolleri sayesinde verileri kullanıcıya ulaştırır. Her sunucu, TCP ile sağlanan bir portalı kullanır. Diğer bir deyişle TCP ile belirlenen genel bir ağ adresi üzerinden bilgiler ulaştırılır. Sunucunun gerektiği zaman aranılan bilginin var olduğunu belirleyebilmesi ve erişimi sağlayabilmesi için sorgulama yapılacak bilgi kaynağının önceden web sunucusunda tanımlanması gerekir. Sorgulanan bilgi parçasının benzeri olan her bilgi aynı klasör ya da aynı veri tabanında sınıflandırılacağından, her bilgiye o bilgiyi tanımlayan üst veri bilgileri ile erişim sağlanır. (Şekil 2).

Üstveri kayıtları genel dermeden bağımsız bir bölümde kayıtlı tutulabileceği gibi HTML temelli web kaynaklarında olduğu gibi kaynakla beraber de kayıt edilebilir.

Üstveri kayıtları aranan bilginin var olan tüm dijital formlarına erişimi sağlamalıdır. Aşağıda üst veri standartlarına bağlı olarak bir dijital kütüphanede bulunabilecek bilgi kaynağı türleri sıralanmıştır:

- HTML dosyaları
- Web siteleri
- Dijital görüntüler
- Veritabanları
- Kitaplar ve diğer basılı kaynaklar
- Müze koleksiyonları
- Arşiv belgeleri
- Üst veri formları, kütüphane katalogları
- Öz ve Atıf indeksleri, diğer indeksler

Üstveri yapıları geleneksel kataloglamadaki bibliyografik alanlar yanında çok daha ayrıntılı bilgileri ve o bilginin kayıtlı olduğu ağ adresi ya da adreslerini de içerebilir. Üst veri terimi özellikle internet ortamındaki bilgi ve bilgi kaynaklarını tanımlamak amacıyla kullanılmaktadır. Bu anlamda üstveri işlemleri katalogçular açısından geleneksel kataloglama ile aynı amaca hizmet etmelidir; çünkü yapılan iş kataloglamanın felsefesinden uzak değildir. Ancak bilgi merkezlerinde oluşturulan OPAC kayıtları o kütüphanedeki kaynakları tanımlarken, bu kaynaklara aynı kütüphanede nasıl erişilebileceğine ilişkin yer bilgilerini de içerir. Oysa üst veri kayıtlarında “yer” bilgisi belirli bir kütüphane ile sınırlı değildir. Web temelli kaynaklar dijital ya da sayısal nesnelere olarak ele alınmaktadır. Dijital nesnelere uluslararası üstveri işlemlerinin yürütülerek paylaşımının yapılabilmesi için standart tanımlayıcılarla temsil edilmesi gerekmektedir. Bu tanımlayıcılar karakter diziliminden ya da bir dizi numaradan oluşmaktadır. Her bir web kaynağını tanımlayan tekbiçim kaynak buldurucu (Uniform Resource Locator) zaman içinde değişebileceğinden Dijital Nesne Tanımlayıcısı (Dijital Object Identifier) olarak Tekbiçim Kaynak Numarası (Uniform Resource Number) ile web temelli her kaynağa bir sabit numara verilerek o kaynağa erişimin sürekliliği sağlanmaktadır. Öte yandan internet üzerinde birden fazla adreste kayıtlı aynı üst veri formlarına erişmek de mümkündür. Diğer yandan kataloglama standartlarına göre kaynaklar birbirinden bağımsız ya da tek tek tanımlanmaktadır. Aynı biçimde üst veri kayıtlarında da bir veri tabanı başlı başına tek bir kaynak olarak ele alınabilir. Ancak Heery bir veri tabanındaki her bir verinin ayrı ayrı tanımlanıp tanımlanamayacağını; yine bir web sitesinin tek başına mı yoksa kendi içindeki bölümler düzeyinde mi üst veri yapılarının ilişkilendirileceğini sorgulamaktadır. (1996). Günümüzde bu sorun gerek Dublin Core’da ve gerekse MARC uyumlu üstveri yapılarında herhangi bir bilgi içeriğinin ana kaynakla ilişkilendirilmesine olanak tanınması ile çözümlenmiştir.

Üstveri türleri ve standartlaşma:

Bibliyografik üstveri geleneksel kataloglamada olduğu gibi kaynakların tanımlanmasını sağlayarak birbiriyle ilişkili bilgilerin bir arada sunulmasına ve böylece gereksinim duyulduğunda farklı erişim uçları ile erişimlerini sağlamaya hizmet eder. Bu anlamda bugün bibliyografik üstveri ile yapılmak istenen 1904 yılında Cutter tarafından ortaya koyulan ilkelerden farklı değildir. O halde hangi standartta olursa olsun bilginin

organizasyonunu günün gereksinimlerini karşılayacak bir yapıya dönüştürürken Cutter 'ın tanımladığı amaçları bugün hala geçerliliğini sürdürmesi bakımından yeniden gözden geçirmek gerekir:

- Yazarı, başlığı ve/veya konusu bilinen bir kaynağın erişimini sağlamak
- Bir kütüphanede yer alan her türdeki bütün kaynakları göstermek (Bir kaynağın o kütüphanede bulunan farklı türdeki bütün sürümlerini de bir araya getirmek)
- Birbiriyle ilişkili kaynakları bir araya getirmek. (Cutter 1904).

Dijital ortamda anılan bu işlevlerin yerine getirilebilmesi için kaynaklar ile üstveri kayıtları arasında bağlantı (hyperlink) kurulması gerekir. Metin biçiminde oluşturulan üstveri kayıtları tanımladıkları dijital bilgi kaynaklarından bağımsız organize edilerek depolanabilirler. Bazı dijital kütüphaneler ve web kaynaklarında ise üstveri kayıtları ile kaynakların içeriği dijital ortamda bir arada erişime açılabilir.

Yapısal üstveri, bir bilgi kaynağını oluşturan bölümlerle o bilgi kaynağının kendisi arasındaki ilişkinin kurulmasına yönelik bilgilerdir. Bir kaynağı oluşturan sayfa veya dosyaların her birinin tanımlanarak erişimin sağlanması üstveri kayıtlarının yazılımda nasıl yapılandırıldığı ile ilgilidir.(Caplan: 2003, 5).

Yönetmel üstveri, üstveri kayıtlarının oluşumu ve yönetilmesine ilişkin bilgilere verilen isimdir. Üst veri kayıtlarının ne zaman ve nasıl oluşturulduğu, erişimlerinden kimin sorumlu olduğu, arşivlenmesinde izlenen yöntemlerin neler olduğu, kullanım ile ilgili şifre ve benzeri kısıtlamaların olup olmadığı gibi konularda alınan kararlar yönetmel üstveri kapsamına girmektedir. Bu türden bilgiler herkesin erişebileceği bilgiler olmamakla beraber bilgi kaynaklarının yönetiminden sorumlu kişilerce izlenebilmektedir. (Caplan, 4).

Üstveri bilgilerinin ağ ortamında etkili paylaşımı ve değişimi için veri toplulukları arasındaki anlam ve kullanım birliği (Semantik: sözdizimi)nin sağlanması yanında verilerin bilgisayarca anlaşılır biçimde anlamlandırılması, düzenlenmesi (Syntax) ve bu amaçla geliştirilen XML(Extensible Markup Language), DTD (Document Type Definition) RDF (Resource Description Framework) gibi kodlama dil ve standartlarının kullanımına gereksinim duyulmaktadır. XML gibi standartlar internet üzerinde hangi dilde ve hangi ortamda olursa olsun veri paylaşımı sağlar; ancak tek başına üstverilerin kullanımı ve değişimi için yeterli değildir. Örneğin bu sorunu aşmak için başlatılan ilk uygulamalardan biri, Yahoo tarama motorundaki bilgilerin robot kullanılmaksızın kütüphaneciler tarafından konu kataloglaması ile standart erişim başlıkları altında sınıflandırılması olmuştur. Bugün farklı üstveri şemalarında temsil edilen ve bugün internette yaygın kullanılan bir dil olan XML ile tanımlanmış üstverilerin değişimi Kaynak Tanımlama Çerçevesi(RDF) yöntemi ile gerçekleştirilmektedir. Bu anlamda RDF, üstverinin tanımlanarak paylaşılmasına izin veren standart bir yapıdır. Aşağıda, RDF'in bir üstveri şeması ile gerçekleştirilen kullanımına ilişkin olarak kısa bir örnek verilmektedir. İlk adımda, elektronik kaynağın genel tanıtımı tekbiçim kaynak buldurucu altında yapılarak DC ve MARCXML gösteriminde temel giriş niteliğindeki alanlara yönelik sözdizimi verilmektedir:

Elektronik bilgi kaynağı: "Women studies" başlıklı bir web sitesi.

DC de başlık gösterimi:

```
<rdf:rdfo>
<rdf:description href= "www.sju.stjohns.edu/library/staugustine/womenstudies.html">
    <dc>Title>Women Studies</DC>Title>
    <dc:Creator>Özlem Bayram</DC:Creator>
    <dc>Type>Web sitesi</DC>Type>
</rdf:description>
</rdf:rdfo>
```

MARCXML gösteriminde başlık alanı gösterimi:

```
<rdf:rdfo>
  <rdf:description href= "www.sju.stjohns.edu/library/staugustine/womenstudies.html">
    <marc:datafieldtag = "245"ind1= "0" ind2= "0">
      <subfield code= "a">Women Studies </subfield code>
      <subfield code= "c">compiled by Ozlem Bayram and Stephen Gelfman</subfield code>
      <subfield code= "u">www.sju.stjohns.edu/library/staugustine/womenstudies.html
    </subfield code>
  </rdf:description>
</rdf:rdfo>
```

Yukarıda örnek olarak seçilen web sitesi elektronik kaynak olarak MARC 21 'de 856 alanında 856#40\$u<http://www.sju.stjohns.edu/library/staugustine/womenstudies.html> biçiminde temsil edilecektir.

Farklı şemalarda kullanılan üstveri yapıları arasındaki semantik uyumun sağlanarak farklı yazılımların bir arada çalışabilmesi (interoperability) için ISO tarafından standart geliştirme amaçlı çalışmalar yürütülmektedir. Bu tür çalışmalarda, bilgi kaynaklarının tanımlanması için kullanılan veri öğelerinin özelliklerinin ne olacağı, üstveri şemalarında hangi göstergelerle nasıl yapılandırılıp kullanılacağı ve farklı dillerdeki karşılıklarının belirlenmesi gibi uluslararası veri paylaşımına yönelik standartlar belirlenmektedir. Böylece farklı üstveri şemalarının bu standartlara uygunluğu üzerinde çalışılarak ISO 'dan tescil (metadata registry) alınmaktadır. (ISO/IEC 11179-5:2005).

Üstveri şemalarının iki temel amacı bulunmaktadır: Birincisi bibliyografik kayıtların elektronik ortamda temsil edilmesine olanak tanıyarak kaynakların bulunmasını sağlamaktır. İkincisi ise bibliyografik kayıtların dil farkı gözetilmeksizin uluslar arası boyutta paylaşılmasını sağlamaktır. Elektronik kaynakların ortaya çıkış ve kullanım özelliklerine göre belirli alanlarda uygulanmak üzere çeşitli üst veri şemaları tasarlanmıştır. Avrupa Birliği'nin Beşinci Çerçeve Programı kapsamında Information Society Technologies tarafından eğitim, coğrafya ve çevre içerikli kaynaklar için üstveri projeleri ile XML şemaları geliştirilmektedir (<http://www.schemas-forum.org/>). Bu türden çalışmaların yanı sıra koruma ve arşiv amaçlı üstveri projeleri de bulunmaktadır. Cedar projeleri (www.leeds.ac.uk/cedars) ve The Encoded Archival Description (EAD) bu projelere örnek oluşturmaktadır. (www.loc.gov/ead/ead.html). Ancak Dublin Core ve AACR2/MARC üstveri yapıları çoklu ortamlar, ses ve görüntü dosyalarını da kapsamak üzere her tür web kaynağı düşünülerek tasarlanmıştır. Yaygın olarak kullanılan bu iki şemaya, sadece tanıtım amacıyla değinilecektir.

Dublin Metadata Core Element Set (Dublin Core: DC): DC, özellikle kütüphanelerde yaygın olarak kullanılan MARC 21 için bir alternatif hareketin sonucunda ortaya çıkmıştır. DC' de Web kaynaklarını tanımlamak için on beş öğe kullanılmaktadır. Bu öğeler MARC yapısındaki kadar ayrıntılı değildir ve hiçbir alan zorunlu olmadığı için daha esnek ve kolay uygulanabilir bir yapı sunmaktadır.

- Başlık (örneğin bir web sitesinin adı)
- Yazar (web sitesini oluşturan kişi veya kurum)
- Konu (web sitesinin anahtar terim ya da standart konu başlığı ile belirlenecek konusu)
- Kaynağın tanımı (web sitesinin içeriğinin tanımlanması, “içindekiler” tanımının doldurulması)
- Yayıncı
- Emeği geçenler
- Tarih
- Tür (bu örnekte “web sitesi”)
- Format
- Tanıtıcılar (URL ve/veya URN)
- Kaynak
- Dil: (Örneğimizde Web sitesinin hazırlandığı dil veya diller)
- İlişki: Kaynağın bölümlerinin ana sayfa ile ilişkisi
- Kapsam:
- Haklar (Kaynağın yönetimi ile ilgili hak ve sorumluluk ifadeleri)

Bu öğelerin hiç biri zorunlu alan değildir; bu nedenle hepsi isteğe bağlı olarak kullanılabilir. Şemanın XML tasarımı <http://dublincore.org> adresinde bulunmaktadır.

MARXML(MARC 21): ABD Kongre Kütüphanesi tarafından özellikle geleneksel kütüphane kataloglarının üstveri yapılarına dönüştürülmesi için geliştirilmiştir. Şemanın XML tasarımı www.loc.gov/marc/ adresinde bulunmaktadır. MARC21 de Anglo American Kataloglama Kuralları 'nda olduğu gibi ISBD niteleme alanları temel kabul edilerek web kaynakları için ISBD(ER) geliştirilmiştir. ISBD (ER)'de ISBD den farklı olarak elektronik bilgi kaynağının türü ve genişliği (örneğin 3 dosyadan oluşan bir ses kaydı gibi) bilgilerinin yer aldığı alanlar ile fiziksel niteleme alanında verilen bilgiler monografılardan farklı olarak web kaynakları için düzenlenmektedir. Üst veri şemalarında elektronik kaynakların adı ve sorumluluk bildirim bilgilerine (creator) ek olarak kaynağa ilişkin standart tanımlayıcılar (URL ve URN) , format (kaynağın türü) ve kaynağın diline ilişkin bilgiler de zorunlu alan olarak eklenmelidir. MARC 21 de 856 alanı ve MARXML gösterimi standart kataloglama bilgisi gerektirdiği için esnek bir kullanım olanağı sunmamakla eleştirilmektedir. Bu bakımdan gelişmiş ülkelerde özellikle kütüphane merkezli olmayan çevrelerce hem ticari kaygılarla ve hem de hatalı uygulamaya açık olması nedeniyle katalogcu olmayanlar tarafından tercih edilmemesi dikkat çekicidir.

Sonuç:

Üstveri şemalarının kullanılması verilerin organize edilerek depolanmasına destek sağlamakla beraber mevcut verilere erişim için tarama ve sorgulama sistemi kurulmadan sağlıklı erişim sağlanamayacaktır. Bu bakımdan üstveri işlemleri sadece web kaynaklarının dijital ortamda tanımlanmasından ibaret değildir. Önemli olan dijital ortamda kayıtlı bilgilerin birbirleriyle ilişkilendirilerek arayan herkes tarafından erişilebilir kılınmasıdır. Konu başlıkları, kavramsal dizin, otomatik dizinleme ve algoritma temelli dizin yapıları veri depolarında harmanlanmış verilere en kolay, çabuk ve doğru erişimi gerçekleştirmek için geliştirilmektedir.

Yukarıda oluşturduğumuz örnekten de anlaşılacağı gibi kataloglamada kullanılan ISBD standart alanları temel alınarak elektronik kaynaklar için gerekli ek alanlara da yer veren üstveri modellerinin XML gösteriminde ülkemizde de kendi gereksinimlerimize uygun

olarak ISO standardının belirlenmesine yönelik çalışmaların yürütüleceğinden kuşku duyulmamaktadır.

Kaynaklar

Arms, William Y. (2001). Digital libraries. Cambridge, Massachusetts: Massachusetts Institute of Technology. 287 s.

Caplan, Priscilla. (2003). Metadata Fundamentals for All Librarians. Chicago: American Library Association. 192 s.

Cutter, C.A. (1904). Rules for dictionary catalog: selections 62 – 71 s. İçinde Foundations of Cataloging. Ed. By M. Carpenter and C. Svenonius. Littleton: Libraries unlimited. 1985.

Dublin Core Standards. URL: <http://dublincore.org>. Erişilen Tarih: 4 Eylül 2005.

Heery, Rachel (1996). Review of metadata formats. Program 30(4) October. Elektronik adres: URL: www.ukoln.ac.uk/metadata/review Erişilen tarih: 28 Ağustos 2005.

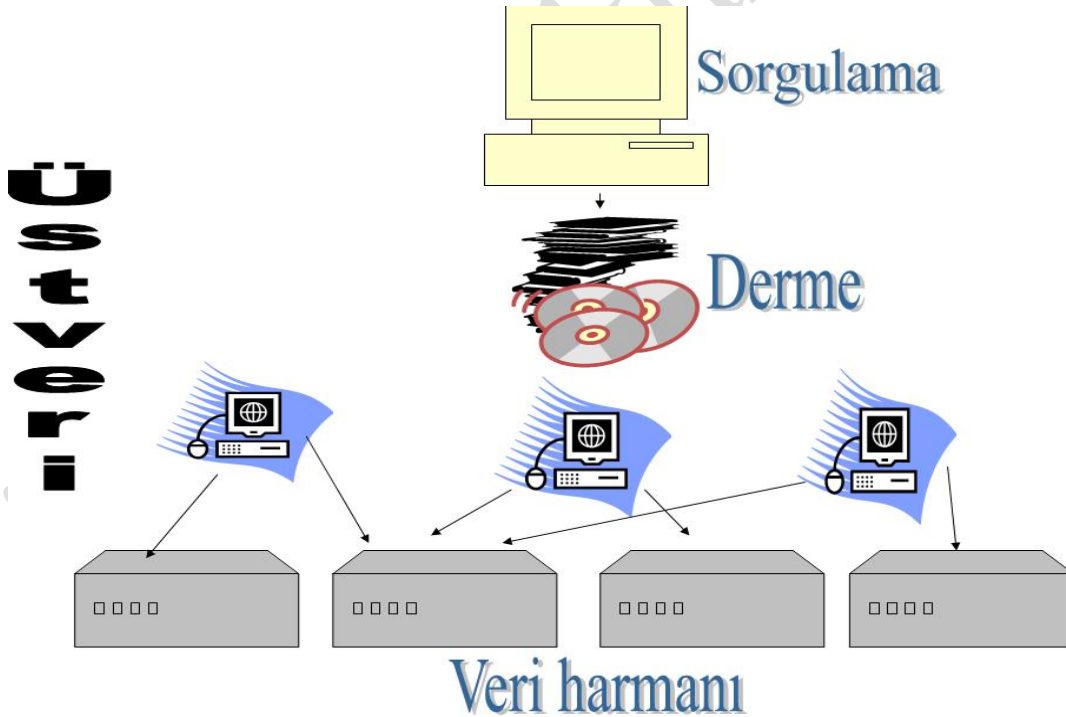
ISO/IEC 11179-5:2005(E). Information Technology- Metadata Registries (MDR) 2nd. Ed.

MARC Standards. URL: <http://www.loc.gov/marc/> Erişilen Tarih: 4 Eylül 2005.

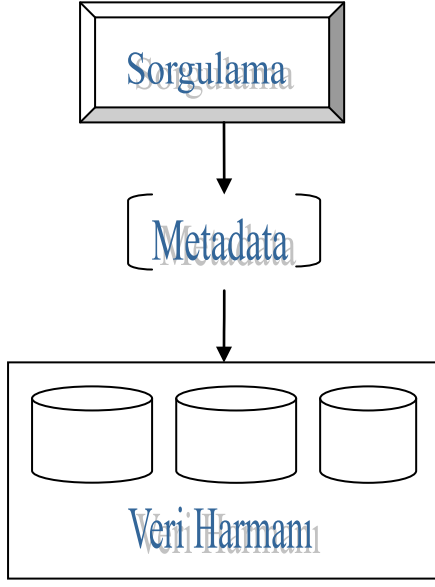
Forum for Metadata Schema Implementers. The Schemas Project. International Society Technologies. URL: <http://www.schemas-forum.org/> Erişilen Tarih: 13 Eylül 2005.

Ek:

1. Şekil



2.Şekil



Özlem Gökkurt