

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
BİLGİ VE BELGE YÖNETİMİ
ANABİLİM DALI**

**KÜLTÜREL MİRASIN KORUNMASINDA WEB ARŞİVLEME VE ULUSAL
KÜTÜPHANE UYGULAMALARI**

Yüksek Lisans Tezi

METEHAN SAMANCI

Ankara-2018

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
BİLGİ VE BELGE YÖNETİMİ
ANABİLİM DALI**

**KÜLTÜREL MİRASIN KORUNMASINDA WEB ARŞİVLEME VE ULUSAL
KÜTÜPHANE UYGULAMALARI**

Yüksek Lisans Tezi

METEHAN SAMANCI

Tez Danışmanı
Prof. Dr. Özlem GÖKKURT DEMİRTEL

Ankara-2018

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
BİLGİ VE BELGE YÖNETİMİ
ANABİLİM DALI**

METEHAN SAMANCI

**KÜLTÜREL MİRASIN KORUNMASINDA WEB ARŞİVLEME VE ULUSAL
KÜTÜPHANE UYGULAMALARI**

Yüksek Lisans Tezi

Tez Danışmanı: Prof. Dr. Özlem GÖKKURT DEMİRTEL

Tez Jürisi Üyeleri

Adı ve Soyadı

İmzası

.....
.....
.....
.....
.....

.....
.....
.....
.....
.....

Tez Sınavı Tarihi

Tez Doğruluk Beyanı Belgesidir.

**T.C.
ANKARA ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü'ne,**

Prof. Dr. Özlem GÖKKURT DEMİRTEL danışmanlığında hazırladığım “Kültürel Mirasın Korunmasında Web Arşivleme ve Ulusal Kütüphane Uygulamaları (Ankara.2018) ” adlı yüksek lisans ☐ - doktora/bütünleşik doktora ☐ tezimdaki bütün bilgilerin akademik kurallara ve etik davranış ilkelerine uygun olarak toplanıp sunulduğunu, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

18.06.2018

Metehan SAMANCI

ÖNSÖZ

İnsan yaşamının bir yansıması olarak üretilen bilgi ve kültür varlıkları çeşitli ortamlara kaydedilmekte ve gelecek kuşaklara birer miras unsuru olarak aktarılmaktadır. Bilgi ve kültür varlıklarını koruma ve onları sonraki nesillere aktarma sorumluluğu ise kültürel miras kurumlarına aittir. İnternet'in dünya genelinde yaygınlaşmasının bir sonucu olarak, binyıllar boyunca fiziksel ortamda üretilen, saklanan ve sunulan bilgi form değiştirmiş ve dijital ortamdaki yerini almıştır. Dijital bilginin üretim ve aktarımında yaşanan büyük artış kültürel miras kurumlarını dijital kültür mirasının sahiplenilmesine yönelik adımlar atmaya zorlamıştır.

Günümüzde dijital ortamda üretilen ve dijitalleştirilen bilgi ve kültür varlıkları genellikle Web üzerinde sunulmakta ve Web mirasını oluşturmaktadır. Kültürel miras kurumları arasında yer alan ulusal kütüphaneler ise Web mirasını koruma, arşivleme ve erişime sunma amacıyla Web arşivleme uygulamalarında bulunmaktadır. Fakat Web büyük hacmi, dinamik yapısı ve karmaşık doğası nedeniyle diğer dijital miras ürünleri arasında arşivlenmesi en zor alanlardan biridir. Bu bağlamda, araştırmamızda kültürel miras ve Web arşivleme konuları derinlemesine ele alınmış ve dünya genelinde ulusal kütüphaneler tarafından sürdürülen Web arşivleme uygulamaları incelenmiştir. Ayrıca Türkiye Milli Kütüphanesi'ne yönelik bir değerlendirmeye de yer verilmiştir.

Araştırma sürecinde güler yüzü, bilgisi ve doğru yönlendirmeleriyle yanımda olan kıymetli danışmanım Prof. Dr. Özlem GÖKKURT DEMİRTEL'e ve bana sürekli destek olan, çok sevdiğim ve değer verdiğim aileme teşekkür ederim.

İÇİNDEKİLER

ÖNSÖZ.....	vii
İÇİNDEKİLER.....	viii
SİMGELER VE KISALTMALAR DİZİNİ.....	xi
TABLolar RESİMLER VE ŞEKİLLER DİZİNİ.....	xvi

BİRİNCİ BÖLÜM

GİRİŞ

1. GİRİŞ.....	1
1.1. Konunun Önemi.....	1
1.2. Araştırmanın Amacı.....	2
1.3. Araştırma Problemi ve Hipotezler.....	3
1.4. Araştırmanın Kapsamı.....	4
1.5. Yöntem ve Veri Toplama Teknikleri.....	5
1.6. Literatür İncelemesi.....	6
1.7. Araştırmanın Düzeni.....	9
1.8. Yararlanılan Kaynaklar.....	10

İKİNCİ BÖLÜM

KÜLTÜREL MİRASIN KORUNMASI

2. KÜLTÜREL MİRASIN KORUNMASI.....	11
2.1. Geleneksel Kültür Mirası.....	12
2.1.1. Kültür Kavramı.....	12
2.1.2. Kültürel Miras.....	14
2.1.3. Kültürel Miras Kategorileri.....	14
2.1.3.1. Somut ve Somut Olmayan Kültürel Miras.....	15
2.1.3.2. Belgesel Miras.....	16
2.1.4. Kültürel Mirasın Yönetimi.....	17
2.1.5. Kültürel Mirasın Korunmasında Uluslararası Sözleşmeler.....	19
2.1.6. Kültürel Mirasın Ekonomik Değeri.....	21
2.2. Dijital Kültür Mirası.....	22
2.2.1. Dijital Miras.....	23
2.2.2. Dijital Mirasın Korunmasında Kütüphanelerin Rolü.....	24
2.2.3. İnternet Kültürü ve Web Mirası.....	27
2.2.4. Web Mirasının Korunması.....	29

ÜÇÜNCÜ BÖLÜM

WEB ARŞİVLEME

3.	WEB ARŞİVLEME.....	31
3.1.	Web Derleme Stratejileri.....	35
3.1.1.	Yığın Derleme.....	35
3.1.2.	Alan Adı Derleme.....	36
3.1.3.	Seçimli Derleme.....	36
3.1.4.	Diğer Derleme Stratejileri.....	37
3.1.5.	Derleme Stratejilerinin Ölçeklendirilmesi.....	38
3.2.	Web Arşivleme Yaşam Döngüsü Aşamaları.....	40
3.2.1.	Değerlendirme ve Seçim.....	40
3.2.2.	Sağlama.....	42
3.2.3.	Organizasyon ve Depolama.....	44
3.2.4.	Tanımlama ve Üstveri.....	45
3.2.5.	Erişim ve Kullanım.....	46
3.2.6.	Koruma.....	50
3.2.7.	Archive-It Web Arşivleme Yaşam Döngüsü Modeli.....	54
3.3.	Yasal Derleme ve Telif Hakları.....	58
3.3.1.	Web Arşivleme için Yasal Dayanak.....	58
3.3.2.	Yasal Derleme ve Telif Hakları.....	61
3.4.	Web Arşivleme Yazılım Araçları.....	63
3.5.	Web Arşivleme ile İlişkili Standartlar.....	66
3.5.1.	Robot.txt Protokolü.....	66
3.5.2.	WARC Dosya Formatı.....	67
3.5.3.	HTML-XML-CSS-JavaScript ve HTTP.....	69
3.5.4.	ISLI.....	69
3.5.5.	Üstveri Standartları.....	71
3.5.6.	Açık Arşiv Bilgi Sistemi Referans Modeli.....	72
3.6.	Web Arşivlemede Mevcut Durum ve Zorluklar.....	73
3.6.1.	Yasal Zorluklar.....	74
3.6.2.	Seçim Zorlukları.....	74
3.6.3.	Web Mimarisi ve Tarayıcı Sınırlılıkları.....	76
3.6.4.	Sosyal Medya Arşivleme.....	77
3.6.5.	Virüsler ve Zararlı Yazılımlar.....	78
3.6.6.	Veri Tekilleştirme.....	79
3.7.	Web Arşivlemede İstatistik ve Kalite Göstergeleri.....	82
3.8.	Web Arşivlerinde İşlevsellik.....	85
3.8.1.	Arama Parametreleri ve Sonuçların Listelenmesi.....	85
3.8.2.	Politikayla İlgili İşlevsellikler.....	86
3.8.3.	Veri Madenciliği ve Kişiselleştirilmiş Hizmetler.....	87
3.8.4.	Kayıp Web Sitelerinin Yeniden Yapılandırılması.....	88
3.9.	Web Arşivlemenin Evrimi.....	88
3.10.	Web Arşivleme Girişimleri.....	91

3.10.1.	Web Arşivlemeye Dayalı Araştırma Projeleri.....	91
3.10.2.	Kuruluşlar.....	93
3.11.	Ulusal Kütüphanelerin Web Arşivleme Uygulamaları.....	95
3.11.1.	Kuzey Amerika Kıtası Ulusal Kütüphane Web Arşivleri.....	97
3.11.2.	Avrupa Kıtası Ulusal Kütüphane Web Arşivleri.....	98
3.11.3.	Asya Kıtası Ulusal Kütüphane Web Arşivleri.....	104
3.11.4.	Okyanusya Kıtası Ulusal Kütüphane Web Arşivleri.....	105
3.12.	Bulguların Değerlendirilmesi ve Araştırma Hipotezlerin Test Edilmesi...	106
3.13.	Türkiye Milli Kütüphanesine Yönelik Bir Değerlendirme.....	116

4.	SONUÇ VE ÖNERİLER.....	119
	KAYNAKÇA.....	127
	EKLER.....	144
	ÖZET.....	163
	ABSTRACT.....	164
	ÖZGEÇMİŞ.....	165

SİMGELER VE KISALTMALAR DİZİNİ

Çalışmada kullanılan simgeler ve kısaltmalar açıklamaları ile birlikte aşağıda yer almaktadır.

Kısaltmalar	Açıklama
ABD	Amerika Birleşik Devletleri
ARCOMEM	The Collect all ARchives to COmmunity MEMories (Toplumsal Hafıza için Tüm Arşivleri Toplama)
AR-GE	Araştırma- Geliştirme
ARPAnet	Advanced Research Projects Agency Network (Gelişmiş Araştırma Projeleri Dairesi Ağı)
CSS	Cascading Style Sheets (Basamaklı Stil Sayfası)
çev.	Çeviren/Çevirenler
DDC	Dewey Decimal Classification (Dewey Onlu Sınıflama)
DNS	Domain Name System (Alan Adı Sistemi)
DPC	Digital Preservation Coalition (Dijital Koruma Koalisyonu)
Ed.	Editör/Editörler
E-LIS	E-prints in Library and Information Science (Kütüphane ve Bilgi Biliminde Elektronik Baskılar)
FQDN	Fully Qualified Domain Name (Tam Nitelikli Alan Adı Kaydı)
FTP	File Transfer Protocol (Dosya Aktarım Protokolü)
GMT	Greenwich Mean Time (Greenwich Ortalama Zamanı)
HTML	HyperText Markup Language (HiperMetin İşaretleme Dili)
HTTP	Hyper Text Transfer Protocol (Hiper Metin Transfer Protokolü)

IA	Internet Archive (İnternet Arşivi)
IFLA	International Federation of Library Associations and Institutions (Uluslararası Kütüphane Dernekleri ve Kurumları Federasyonu)
IIPC	International Internet Preservation Consortium (Uluslararası İnternet Koruma Konsorsiyumu)
IMF	Internet Memory Foundation (İnternet Bellek Vakfı)
Internet	International Network (Uluslararası Ağ)
IP	Internet Protokolü (Internet Protocol)
ISAD(G)	General International Standard Archival Description (Uluslararası Arşivsel Tanımlama Standardı)
ISBN	International Standard Book Number (Uluslararası Standart Kitap Numarası)
ISLI	International Standard Link Identifier (Uluslararası Standart Bağlantı Tanımlayıcı)
ISO	International Organization for Standardization (Uluslararası Standartlaştırma Örgütü)
ISSN	International Standard Serial Number (Uluslararası Standart Süreli Yayın Numarası)
ITU	International Telecommunication Union (Uluslararası Telekomünikasyon Birliği)
IWA	International Web Archiving Workshop (Uluslararası Web Arşivleme Çalıştayı)
LAWA	Longitudinal Analytics of Web Archives (Web Arşivlerinin Boylamasına Analizi)
LCSH	Library of Congress Subject Headings (Kongre Kütüphanesi Konu Başlıkları)
LIWA	The Living Web Archives (Yaşayan Web Arşivleri)
MARC	MAchine-Readable Cataloging (Makinece Okunabilir Kataloqlama)

MD	Message Digest (İleti Kütüğü)
METS	Metadata Encoding and Transmission Standard (Üstveri Kodlama ve Aktarma Standardı)
NA	National Archives (Ulusal Arşivler)
NDSA	National Digital Stewardship Alliance (Ulusal Dijital Yönetim İttifakı)
NEDLIB	Networked European Deposit Library Project (Ağ Tabanlı Avrupa Derleme Kütüphanesi Projesi)
NWA	Nordic Web Archives (Kuzey Avrupa Ülkeleri Web Arşivleri)
OAIS	Open Archival Information System (Açık Arşiv Bilgi Sistemi)
OCLC	Online Computer Library Center (Çevrimiçi Bilgisayar Kütüphanesi Merkezi)
ODTÜ	Orta Doğu Teknik Üniversitesi
PDF	Portable Document Format (Taşınabilir Belge Formatı)
PREMIS	Preservation Metadata: Implementation Strategies (Metadata Koruma: Uygulama Stratejileri)
t.y.	tarih yok
TCP/IP	Transmission Control Protocol/Internet Protocol (İletim Kontrol Protokolü/ Internet Protokolü)
UKWAC	United Kingdom Web Archiving Consortium (Birleşik Krallık Web Arşivleme Konsorsiyumu)
UN	United Nations (Birleşmiş Milletler)
UNDP	United Nations Development Programme (Birleşmiş Milletler Kalkınma Programı)
UNESCO	United Nations Educational, Scientific and Cultural Organization (Birleşmiş Milletler Eğitim, Bilim ve Kültür Örgütü)
URI	Uniform Resource Identifier (Tekbiçimli Kaynak Tanımlayıcı)

URL	Uniform Resource Locator (Tekbiçimli Kaynak Konumlayıcı)
URN	Uniform Resource Name (Tekbiçimli Kaynak Adı)
vb.	ve benzeri
vd.	ve diğerleri
W3C	World Wide Web Consortium (Dünya Çapında Web Konsorsiyumu)
WALCM	Web Archiving Life Cycle Model (Web Arşivleme Yaşam Döngüsü Model)
WARC	Web ARCive
WAS	Web Archive Services (Web Arşiv Servisleri)
Wayback Machine	Wayback Makinesi
Web	World Wide Web, WWW, W3 (Dünya Çapında Ağ)
WEF	World Economic Forum (Dünya Ekonomik Forumu)
XML	eXtensible Markup Language (Genişletilebilir İşaretleme Dili)

TABLolar, RESİMLER VE ŞEKİLLER DİZİNİ

TABLolar DİZİNİ

Tablo 1. Web Arşivleme Yazılım Araçları.....	63
Tablo 2. İlişkili Performans Göstergeleriyle Birlikte Web Arşivleme Süreci.....	84
Tablo 3. Dünya Geneline Web Arşivleme Uygulamalarına Yer Veren Ulusal Kütüphanelerin Derleme Stratejileri, Web Arşivleme Programlarına Başlama Yılları, Erişim Yöntemleri ve Arayüz Dilleri.....	108
Tablo 4. Ulusal Kütüphanelerde Web Arşivleme Uygulamalarına Yer Veren Ülkelerin Küresel Rekabet Endeksi, İnsani Gelişim Endeksi, Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi Sıralamaları.....	114

ŞEKİLLER DİZİNİ

Şekil 1. Web Arşivleme Üretim Şeması.....	40
Şekil 2. Türkiye Milli Kütüphanesi Ana Sayfasına Yönelik Yapılan Bir URL Araması.....	47
Şekil 3. Archive-It Web Arşivleme Yaşam Döngüsü Modeli.....	55
Şekil 4. ISLI Bağlantı Modeli.....	70
Şekil 5. Açık Arşiv Bilgi Sistemi Referans Modeli.....	73
Şekil 6. Kongre Kütüphanesi Web Arşivleri'nde Banner Kullanımı.....	86
Şekil 7. Tarihteki İlk Web Sitesi.....	89
Şekil 8. Web Arşivleme Yapılan Ülkelerin Coğrafi Dağılımı.....	90
Şekil 9. Web Arşivleme Uygulamalarına Yer Veren Ulusal Kütüphanelerin Coğrafi Dağılımı.....	96
Şekil 10. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Programlarına Başlama Yılları ve Sayısal Artışları.....	96
Şekil 11. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Web Derleme Stratejileri Yüzdesel Dağılım Grafiği.....	111
Şekil 12. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Erişim Yöntemleri Yüzdesel Dağılım Grafiği	112

GİRİŞ

1.1. Konunun Önemi

İnsanlık varoşundan beri sahip olduđu bilgiyi kaydetmekte, bir miras unsuru olarak korumakta ve gelecek nesillere aktarmaktadır. Bilginin korunmasında ve aktarımında kilit rolde bulunan ulusal kütüphaneler ise kültürel miras kurumları olarak geniş bir yelpazedeki bilgi ve kültür nesnelerine ev sahipliğı yapmaktadır.

Günümüzde bilgi kayıt ve sunum ortamlarının değışmesiyle bilginin boyut ve çeşitliliğinde hızlı bir artış yaşanmaktadır. Bu durum başta ulusal kütüphaneler olmak üzere kültürel miras kurumlarını bilgi nesnelerinin kalıcı olarak korunmasına ve aktarımına yönelik yeni önlemler almaya ve farklı hizmetler geliştirmeye zorlamaktadır. Ulusal kültürel mirasın toplayıcısı ve koruyucusu konumunda olan ulusal kütüphaneler fiziksel materyallerin dijitalleştirilmesi ve bilginin dijital ortamda üretilmesi neticesinde çok büyük boyutta ve çeşitlilikte dijital koleksiyonlara sahip olmuştur. Web hem dijitalleştirilen hem de dijital olarak üretilen nesneleri içinde barındırdığından, Web siteleri ulusal kütüphanelerin dijital koleksiyonlarında gün geçtikçe daha fazla yer almaktadır.

Web, Internet üzerinde yayınlanan birbiriyle bağlantılı hiper-metin dokümanlardan oluşan bir bilgi sistemi olarak tanımlanabilir [1]. Sosyal iletişim ve etkileşim için bir araç olarak kullanılan Web ayrıca içerisinde metin, görsel, ses ve video gibi farklı türde medyaları barındıran bir bilgi hazinesidir. Web, 1991 yılındaki icadından bu yana geçen zaman içinde çok sayıda bireyin ve yayıncının bilgi dağıtımını yaptığı ve birbirleri ile iletişim kurduğu güçlü bir platform haline gelmiştir. Yerel, ulusal ve evrensel kültürün bir kaydını tutması ve özgün yapısı onu dijital kültür mirasının önemli bir kaynağı haline getirmiştir.

Sürekli olarak değışen ve büyüyen Web'in boyutunu ifade etmek zordur. Arama motorları aracılığıyla görüntülenebilen ve yüzey Web olarak ifade edilen Web'in büyüklüğünün yaklaşık olarak 4,78 milyar sayfa olduğu tahmin edilmektedir [2]. Arama motorları tarafından erişilemeyen ve derin Web olarak ifade edilen diğer bölümle birlikte düşüldüğünde, yüzey Web, Web'in tamamı içinde yalnızca %4'lük bir bölüme karşılık gelmektedir [3]. Bu devasa boyuttaki bilgi yığını Web sitelerinin kısa ömürlü doğası ve

hassas yapısı nedeniyle yok olma tehdidi altındadır. Bu tehdit bilimsel, kültürel veya tarihsel değeri yüksek, benzersiz birçok kaynağı barındıran Web'in arşivlenmesine yönelik zorunluluğu arttırmaktadır.

Kültürel mirasın korunması ve gelecek nesillere aktarılması amacıyla Web'in arşivlenmesi gerekliliği birer hafıza kurumu olan ulusal kütüphaneleri ön plana çıkarmaktadır. Web arşivlerinin ulusal kültür mirasının bir parçası olarak görülmesi ve Web sitelerinin birçok ülkede yasal derleme kapsamında değerlendirilmesi, Web arşivleme konusunda ulusal kütüphanelerin yetki ve sorumluluğunu arttırmıştır. Avrupa genelinde Web materyallerini arşivleyerek koruyan kurumların %58'inin ulusal kütüphaneler veya denetimi altındaki kurumlar olması ulusal kütüphanelerin aldığı bu sorumluluğu doğrular niteliktedir (Lasfargues, Martin ve Medjkoune, 2012:119).

Türkiye Milli Kütüphanesi'nin mevcut derleme mevzuatı gereği dijital kültürel miras varlıklarını arşivleme sorumluluğu ve ülkemizde Web arşivleme alanında henüz herhangi bir uygulamanın gerçekleştirilmemiş olması bu araştırmanın ortaya konulmasındaki başlıca nedenlerdir. Ayrıca ülkemizde Web arşivlemeye yönelik yürütülen bilimsel çalışmaların sınırlı sayıda olması ve Web arşivlerinin bilimsel araştırmalara kaynaklık etme potansiyeli de araştırmanın ortaya konulmasındaki diğer nedenler olarak sıralanabilir. Bu bağlamda araştırmamızda başta ulusal kütüphaneler olmak üzere dünya genelinde kültürel miras kurumlarının Web arşivleme faaliyetleri incelenmiş ve edinilen bilgilerle Türkiye Milli kütüphanesi'ne yönelik bir değerlendirmede bulunulmuştur.

1.2.Araştırmanın Amacı

İnsan kültür mirasını korumak ve hâlihazırda Web'den kaybolan veya çok geçmeden ortadan kaybolacak içeriğe erişim sağlamak için dünyadaki pek çok kültürel miras kurumu Web kaynaklarını arşivlemeye başlamıştır (Niu, 2012a). Ülkemizdeki kültürel miras kurumları ise Web arşivleme konusunda bugüne dek herhangi bir uygulama gerçekleştirilmemiştir. Söz konusu durumun devam etmesi Web üzerinde bulunan yerel, ulusal veya evrensel nitelikteki değerli kültürel miras varlıklarının gözden kaybolması veya sonsuza dek yitilmesi anlamına gelecektir. Gerek sosyal gerekse kamusal alanda Internet'in aktif olarak kullanıldığı ülkemizde [4] kültürel mirasın sonraki nesillere aktarımında yaşanabilecek problemler ulusal hafızanın önemli bir bölümünün kaybedilmesine neden olacaktır.

Bir araştırma ve derleme kütüphanesi olan Türk Milli Kütüphanesi'nin yürüttüğü elektronik derleme faaliyetleri kapsamında bilgi ve kültür mirası ürünlerine eksiksiz olarak koleksiyonunda yer vermesi, toplumsal hafızanın ve ulusal kültürün güçlendirilmesi ve bilgi toplumu inşası adına büyük önem arz etmektedir. Bilginin yaklaşık olarak %90 oranında sayısal ortamda üretildiği günümüzde (Kirschenbaum, Ovenden ve Redwine, 2010:2) Web kaynaklarının arşivlenmemesi ve bunun neticesinde

gelecek nesillere aktarılamaması büyük bir bilgi kaybına ve ulusal kültürün zayıflamasına neden olacaktır.

Araştırmamızın amacı kültürel mirasın korunması bağlamında dünya genelinde ulusal kütüphanelerin Web arşivleme uygulamalarının incelenmesi, karşılaştırılması ve Web arşivleme faaliyetlerinin Türkiye Milli Kütüphanesi'nde uygulanmasının gerekliliğini ortaya koymaktır. Araştırmamızın diğer amaçları ise;

- Dünya genelinde ve ulusal kütüphaneler özelinde Web arşivleme uygulamalarının mevcut durumunu izlemek ve yönelimleri ölçmek,
- Gelecekte Türkiye Milli Kütüphanesi'nde Web arşivleme uygulamalarına yer verilmesi durumunda benimsenecek yaklaşımları ve atılacak adımları belirlemek,
- Web arşivleme sürecinde önemli görülen derleme stratejileri, yaşam döngüsü aşamaları, yazılımlar, standartlar, zorluklar, istatistik ve kalite göstergeleri gibi alanlarda açıklayıcı bilgiler sunmak ve değerlendirmeler yapmak,
- Web arşivlemeye ilişkin kapsamlı bir analiz gerçekleştirmek ve muhtemel uygulayıcılara bir öngörü kazandırmak,
- Web arşivleme konusunda Türkçe literatürü desteklemek ve gelecekteki araştırmalara katkı sağlamak,
- Web arşivlemeye yönelik kavramları analiz etmek ve araştırmacıların kullanımına sunmak üzere önemli görülen kavramları içeren bir kavram listesi oluşturmak olarak sıralanabilir.

1.3.Araştırma Problemleri ve Hipotezler

Bir araştırma kütüphanesi ve ulusal derleme kütüphanesi olan Türkiye Milli Kütüphanesi, yürütmekte olduğu dijitalleştirme ve elektronik derleme faaliyetleri aracılığıyla dijital arşivleme uygulamalarına yer vermektedir. Kurum nezdinde mevcut derleme mevzuatı gereği elektronik ortamda oluşturulan kitap, gazete, dergi ve tez gibi çevrimiçi dijital materyallerin arşivlenmesine yönelik çalışmalar yürütülmekte fakat Web materyalleri henüz arşivlenmemektedir. Kültürel miras kapsamında değerlendirilen ulusal Web mirasının Türkiye Milli Kütüphanesi tarafından arşivlenmemesi büyük bir bilgi boşluğuna ve dolayısıyla ulusal hafızanın bir bölümünün kaybedilmesine neden olacaktır. Ulusal Web mirasının kaybedilme tehdidi karşısında Türkiye Milli Kütüphanesi'nde Web arşivleme uygulamalarına henüz yer verilmemesi, çalışmamızda sunulan araştırma problemlerinin ve hipotezlerin belirlenmesinde etkili olmuştur. Söz konusu araştırma problemlerinin yanıtlanması ve hipotezlerin test edilmesi sürecinde Web arşivleme alanında üretilen bilimsel yayınlardan, çeşitli istatistiklerden ve dünyadaki ulusal kütüphanelerin uygulama örneklerinden yararlanılmıştır. Çalışma kapsamında belirlenen araştırma hipotezleri ve araştırma problemleri aşağıda sunulmaktadır:

Araştırma Temel Hipotezi: “Türkiye Milli Kütüphanesi tarafından Web arşivleme uygulamaları gerçekleştirilmezse ulusal Web mirası korunamaz.”

Temel Araştırma Problemi: “Türkiye Milli Kütüphanesi ulusal Web mirasını koruma amacıyla Web arşivleme uygulamasına yer vermeli midir? Eğer yer verecekse diğer ulusal kütüphanelerle karşılaştırıldığında seçeceği web derleme ve erişim stratejileri neler olmalıdır?” şeklinde belirlenmiştir.

Araştırma Alt Hipotezi 1: “Küresel rekabet düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Araştırma Alt Hipotezi 2: “İnsani gelişme düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Araştırma Alt Hipotezi 3: “Bilgi ve iletişim teknolojileri gelişmişlik düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Alt Araştırma Problemi: “Ülkelerin küresel rekabet, insani gelişme ve bilgi ve iletişim teknolojileri gelişmişlik düzeyi ile ulusal kütüphanelerinin Web arşivleme uygulamalarına yer vermesi arasında bir ilişki var mıdır?”

1.4. Araştırmanın Kapsamı

Araştırmamızın kapsamını kültürel miras ve Web arşivleme konularına ilişkin yerli ve yabancı bilimsel literatür ve Web arşivlemeye yönelik olarak dünya genelinde sürdürülen uygulamalar oluşturmaktadır. Kültürel mirasın korunması bağlamında Web arşivlemeye yönelik olarak kuramsal çalışmaların ve uygulamaların analiz edilmesi, anlaşılması ve aktarılması gerekliliğinden hareketle, çalışmamızda teorik bilgilerle birlikte ulusal kütüphanelerin Web arşivleme uygulamaları da aktarılmıştır. Web arşivlemeye ilişkin konuları daha iyi yorumlama gereksinimi sonucunda gerektiğinde ulusal kütüphaneler dışında bazı kuruluşların Web arşivleme faaliyetlerine de değinilmiştir. Kaynak tarama sürecinde, Web arşivlemenin gelişimini anlamak ve bugün ulaştığı noktayı belirlemek için farklı türlerdeki kaynaklardan yararlanılmış ve yıl aralığı geniş tutulmuştur. Aynı süreçte, teorisyenlerin ve uygulayıcıların teknik bilgiye gereksinim duyabileceği öngörüsüyle yayımlanan teknik dokümanlardan ve çok sayıda Web sitesinden yararlanılmıştır.

Araştırmamızda ilk olarak kültürel miras konusuna değinilmiş ve kültürel miras kapsamında dijital kültürel miras ve Web mirası konuları ele alınmıştır. Ardından çalışmamızda yasal derleme, derleme stratejileri, yaşam döngüsü aşamaları, yazılım araçları, standartlar, zorluklar, istatistik ve kalite göstergeleri ve çeşitli işlevsellikler Web arşivleme bağlamında ele alınmıştır. Web arşivlemeye ilişkin önemli görülen konulara

detaylı bir biçimde yer verilen bu bölümde araştırmanın kapsamı geniş tutulmuştur. Sonrasında Web arşivlemenin evrimine, önemli görülen araştırma projelerine ve uluslararası girişimlere yer verilmiş, 27 Ulusal Kütüphane'nin Web arşivleme uygulamaları coğrafi lokasyonlarına göre dört kıta altında gruplandırılarak aktarılmış ve çalışma kapsamında belirlenen hipotezler test edilerek araştırma problemlerine yanıt aranmıştır. Araştırmamızda ayrıca Türkiye Milli Kütüphanesi'ne yönelik bir değerlendirme yapılmış ve son olarak sonuç ve öneriler bölümüne yer verilmiştir.

Çalışmamızda sunulan “Web arşivleme” ve “Web derleme” kavramlarının birbirinden farklılıklarının ortaya konularak açıklanması, bu işi yapacak kurumlar bağlamında karışıklığa ve yanlış anlaşılmalara yer verilmemesi adına önemlidir. İçerisinde Web derleme, arşivleme, depolama, koruma ve erişime sunma gibi temel alanları barındıran Web arşivleme dijital arşivlemenin bir alt alanıdır. Web derleme ise hem kurumların Web sitelerini toplamada kullanacağı tarama yaklaşımını hem de ulusal kütüphanelerin yasal derlemeyi gerçekleştirmede yararlanacağı yasal izni ifade eder.

1.5. Yöntem ve Veri Toplama Teknikleri

Araştırmamızda, amacı araştırma konusu olguları ve bu olgular arasındaki ilişkileri saptama, sınıflama ve kaydetme olarak ifade edilen (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2008:6) betimleme yönteminden yararlanılmıştır. Araştırma konusu olguların saptanması ve açıklanması için betimleme aracı olarak ise gözlem tekniği tercih edilmiştir. Gözlem, kendiliğinden oluşan ya da bilinçli olarak hazırlanan olayları, belirttikleri sırada sistemli ve amaçlı bir biçimde inceleyerek bilgi toplama olarak ifade edilebilir (İslamoğlu, 2009:104). Çalışmamızda gözlem tekniğinin tercih edilmesinin nedeni, Web arşivlemeye ilişkin faaliyetlerin ve ulusal kütüphanelerin Web arşivleme uygulamalarının çoğunlukla Web üzerinde görünür olması ve söz konusu tekniğin hızlı ve düşük maliyetle bilgi toplamaya olanak sağlamasıdır. Bu bağlamda Web arşivlemeye ilişkin konular geniş bir bakış açısıyla ele alınarak betimlenmiş ve dünya genelinde Web arşivleme uygulamalarına yer veren ulusal kütüphaneler belirlenerek gözlem tekniğiyle incelenmiştir. Örneklemin belirlenmesinde akademik dokümanlardan ve Web sitelerinden yararlanılmıştır. Web üzerinde gerçekleştirilen taramaların sonucunda, örneklemin kapsamını Web arşivleme uygulamalarına yer veren ulusal kütüphaneler oluşturmuştur.

Veri toplama sürecinde ilk olarak Türkiye Milli Kütüphanesi'nin ve Ankara Üniversitesi'nin abone olduğu elektronik veri tabanlarında arama yapılmıştır. Ardından Internet ortamında bulunan açık erişim materyaller aranmış ve Web siteleri ziyaret edilmiştir. Gerçekleştirilen aramalar sonucunda makale, tez, bildiri, kitap bölümü, standart, teknik rapor, çalıştay raporu, Web sitesi vb. akademik, teknik veya bilgilendirici dokümanlar sağlanmıştır. Söz konusu yayınlar bilimsel ve/veya teknik doküman olma

özelliğine ve ihtiva ettiği içeriğin aranan bilgiyi sağlama niteliğine göre kategorilere ayrılarak dikkatle incelenmiş ve araştırma konusuyla kapsam bakımından uyumluluk gösterenler değerlendirmeye alınmıştır. Veri toplama sürecinde başta ulusal kütüphaneler olmak üzere Web arşivleme alanında faaliyet gösteren kurumların Web siteleri de incelenmiştir. Elektronik veri tabanlarında ve Internet ortamında Türkçe ve İngilizce dillerinde kaynak taramaları gerçekleştirilmiştir. Tercih edilen arama terimleri: “kültür, kültürel miras, kültürel mirasın korunması, dijital kültür mirası, Web, Web mirası, Web arşivleme, Web arşivleme standartları, Web arşivleme ve ulusal kütüphaneler, Web arşivleme uygulamaları, culture, cultural heritage, preserving cultural heritage, digital cultural haritage, Web heritage, Web archiving, Web archiving standards, Web archiving and national libraries ve Web archiving practices” olarak sıralanabilir.

Web arşivleme konusunda literatürde yer alan Türkçe kaynakların niceliksel azlığı nedeniyle çalışmada ağırlıklı olarak İngilizce dilindeki kaynaklardan yararlanılmıştır. Sağlanan kaynaklar konu içeriğine göre dosyalar halinde gruplandırılmış ve ardından kaynak okumaları gerçekleştirilmiştir. Araştırmamıza katkı sağlayacağı düşünülen kaynak içerikleri yapılan göndermeler aracılığıyla sunulmuştur. Kültürel miras ve Web arşivleme konularına yönelik olarak gerçekleştirilen çeviriler, yapılan okumalar, incelenen Web siteleri ve sunulan istatistikler, çalışma kapsamında belirlenen araştırma problemlerinin yanıtlanmasına ve hipotezlerin test edilmesine katkı sağlamıştır.

1.6. Literatür İncelemesi

Web, kısa ömürlü doğası, sınırlarının net olmayışı ve boyutsal büyüklüğü nedeniyle arşivlenmesi zor bir alan olarak karşımıza çıkmaktadır. Buna rağmen, Web arşivleme faaliyetleri yerel, ulusal ve evrensel nitelikteki kültürel miras unsurlarının korunması amacıyla dünya genelinde hızla yaygınlaşmaktadır. Web arşivlemeye ilişkin yabancı literatüre bakıldığında dünya genelinde veya Avrupa ve Kuzey Amerika kıtaları özelinde faaliyetlerin ve uygulamaların mevcut durumunu ele alan ve gelişimini yansıtan birçok çalışmanın yürütüldüğü görülmektedir. Hakala 2004 yılındaki çalışmasında, Avrupa’da Web arşivlemenin gelişimine yönelik kısa bir tarihçeye yer vermiş ve Kuzey Avrupa ülkelerinin Web arşivleme uygulamalarında yaşadıkları teknik zorlukları aktararak çözüm önerilerinde bulunmuştur. Lasfargues ve arkadaşları’nın 2012 yılındaki çalışmasında, Web arşivlemenin önemi vurgulanmış ve Web arşivleme alanında Avrupa’nın mevcut durumu yaşanan sorunlarla birlikte ele alınmıştır. Niu 2012 yılındaki ilk çalışmasında Web arşivlerinin işlevselliklerine yer vermiş ve aynı yıl yayımladığı ikinci çalışmasında Web arşivleme hakkında genel bir değerlendirmede bulunmuştur. Pennock 2013 yılındaki çalışmasında Web arşivleme faaliyetinin zorluğundan bahsetmiş ve zorluk yaşanan çeşitli alanlara ve birkaç vaka çalışmasına yer vermiştir. Costa, Gomes ve Silva 2016 yılında yürüttükleri bir çalışmada, 2010 ve 2014 yıllarında uyguladıkları anketlerin sonuçlarını karşılaştırmış ve dünya genelinde Web arşivleme faaliyetlerinin gelişimini incelemişlerdir. Bailey, Grotke, McCain, Moffatt ve Taylor’un 2017 yılındaki

alışmasında ise, 2013 ve 2016 yıllarında yrtlen anket alışmaları neticesinde ABD’de srdrlen Web arşivleme faaliyetlerinin gelişimi deęerlendirilmiřtir.

Bu alışmaların ortak özellięi kamu veya niversite gzetiminde veya ticari nitelikteki kuruluřlar tarafından srdrlen Web arşivleme faaliyetlerini ele almaları ve kuruluřlar arasında net bir ayırım gzetmeden elde edilen bilgileri ve istatistikleri sunmalarıdır. Dnya genelinde ulusal ktphanelerin Web arşivleme uygulamalarını ele alan ve onları birbirleriyle karřılařtıran bir alışmaya ise henz rastlanmamıřtır. Bu noktada Web arşivleme konusunda ulusal ktphane uygulamalarının mevcut durumunu ele almasıyla alışmamız dięerlerinden farklılařtırmakta ve nem kazanmaktadır.

Son 20 yılda Web arşivleme faaliyetlerine ynelik artan bir ilgi vardır. Bunun bir sonucu olarak arařtırmacıların alışmalarında Web arşivlemenin farklı alanlarına yneldikleri sylenebilir. Bu alışmaların bir blmnde Web arşivlemeye ynelik yasal sreler ve telif ve mlkiyet hakları konuları ele alınmaktadır. Beunen ve Schiphof 2006 yılındaki alışmasında Web arşivlemenin yasal ynlerini ele almıř, mlkiyet hakları konusunu ve Hollanda Milli Ktphanesi’nin Web arşivleme politikasını deęerlendirmiřtir. Lasfargues ve arkadaşları’nın 2008 yılındaki alışmasında Fransa’da geerli olan derleme yasasına atıfta bulunulmuř, Fransa Milli Ktphanesi’nde yrtlen Web arşivleme faaliyetleri ve kullanılan arşiv tarayıcılarına iliřkin bilgiler ve teknik zmler sunulmuřtur. Glanville 2010 yılındaki alışmasında etik ve yasal konulara yer vermiř ve Hollanda ve Avustralya Ulusal Ktphaneleri rneklerinde Web arşivleme srecinde karřılařılan zorluklara deęinmiřtir. Cadavid 2014 yılındaki alışmasında Singapur’da Web arşivlemenin gelişimini deęerlendirmiř ve telif haklarına ynelik grř ve nerilerini aktarmıřtır.

İngilizce literatrde ayrıca kltrel mirasının korunması konusunu ele alan eřitli yayınlar bulunmaktadır. Fabian 2014 yılındaki alışmasında dijital mirasın neminden ve koruma srecinde ulusal ktphanelerin zerine dřen sorumluluktan bahsetmiřtir. Council of Europe’un 2005 yılında yayınladıęı karar metninde kltrel mirasa iliřkin hakların Evrensel İnsan Hakları Beyannamesi’nde tanımlandıęı řekliyle kabul edilmesi gerektięi aktarılmıř ve kltrel mirasın korunmasının nemi ve gereklilięi vurgulanmıřtır (Council of Europe, 2005:1). Birleřmiř Milletler Eęitim, Bilim ve Kltr rgt (UNESCO) tarafından yayımlanan 1982, 2003, 2013, 2015 tarihli dokmanlarda ise kltrel miras, belgesel miras (documentary heritage) ve dijital kltrel miras varlıklarının artan nemine deęinilmiř ve koruma amacıyla atılması gerekli grlen adımlara yer verilmiřtir.

İngilizce literatrde Web arşivlemenin teknik ynlerini ele alan kaynaklar da mevcuttur. Guenter ve Myrick’in 2007 yılındaki alışmasında, Kongre Ktphanesi Minerva Projesi rneęinde Web arşiv materyallerinin toplanması ve ynetilmesinde yařanan bazı teknik zorluklar aktarılmıř ve uygulanan stveri stratejileri incelenmiřtir.

Murray ve Hsieh'in 2008 yılındaki çalışmasında, Kongre Kütüphanesi tarafından yürütülen ve koleksiyon oluşturma, depolama ve yönetim desteği sağlayan bir Web arşivleme hizmeti olan "Web-at-Risk" adlı Proje tanıtılmıştır. Uluslararası Standartlaştırma Örgütü'nün (ISO) 2009 yılında yayımladığı standart, WARC dosya formatını ele almakta; 2012 yılında yayımladığı standart, istatistik ve kalite konularını değerlendirmekte; aynı yıl yayımladığı diğer standart, Açık Arşiv Bilgi Sistemi Referans Modeli'ni (OAIS) açıklamakta; 2014 yılında yayımladığı standart, uzun süreli koruma bağlamında dijital depolama ortamının seçimini konu edinmekte; 2015 yılında yayımladığı standart ise Uluslararası Standart Bağlantı Tanımlayıcı'yı (ISLI) tanıtmaktadır. Grotke 2011 yılındaki çalışmasında Web arşivleme sürecinde yaşanan teknik ve yasal sorunları Kongre Kütüphanesi örneğinde aktarmıştır. Kim ve Ross'un 2012 yılındaki çalışmasında Web arşivleme için dijital korumayı destekleyen uygun depolama konteyner formatının seçimine yönelik bir değerlendirmede bulunulmuş ve belirlenen üç konteyner dosya formatına yönelik bir karşılaştırmaya yer verilmiştir. Bununla birlikte Internet Archive (IA) 2013 yılında Web Arşivleme Yaşam döngüsü Modeli'ni (WALCM) yayınlamıştır. Huurdeman, Ben David ve Samar'ın 2013'te yürüttüğü bir çalışmada Web arşivlerinde erişim konusu ele alınmıştır. 2017 yılında National Archives Web arşivlemeye yönelik bir rehber hazırlamış ve kullanıma sunmuştur.

Türkçe literatüre bakıldığında kültür ve kültürel miras konularında kaynak çeşitliliğinin fazla; fakat Web arşivleme konusunda kaynak çeşitliliğinin az olduğu söylenebilir. Tonta'nın 2002 yılındaki çalışmasında kültür ve kültürel miras konuları ele alınmış ve kültürel mirasın ihmal edilme nedenleri aktarılmıştır. Zan 2006 yılında hazırladığı yüksek lisans tezinde derleme olgusu ve elektronik yayınların derlenmesi konularını ele almıştır. Aldemir ve Oğuz 2006 yılında yayımladıkları çalışmada, Web'in neden arşivlenmesi gerektiğine değinmiş, Web arşivlemede kullanılan belli başlı yaklaşımları ele almış, ulusal ve uluslararası ölçekteki Web arşivleme çalışmalarına yer vermiştir. Bayter, 2009 yılındaki çalışmasında, Türkçe Web sitelerinin kataloglanması konusunu ele almıştır. Özbağ, 2010 yılındaki çalışmasında, Web arşivlemeyi dijital koruma bağlamında incelemiş ve kültürel mirasın korunmasına ve arşivlenmesine yönelik kavramsal bir model oluşturmuştur. Oğuz 2011 yılındaki çalışmasında toplum bilimlerinde kültür kavramının ortaya çıkışı üzerinde durmuş ve kültür kavramının toplum bilimleri literatüründeki farklı kullanımlarına yer vermiştir. Oğuz 2011, yılında yayımlanan bir diğer çalışmasında ise kültür politikaları ile kütüphane arasındaki ilişkiyi ortaya koymuş ve kültürel mirasın korunması sürecinde kütüphanelerin rolüne ve önemine vurgu yapmıştır. Özel, 2014 yılındaki çalışmasında kültürel miras varlıklarının dijitalleştirilmesi konusunu ele almış ve planlama çalışmalarının ve üstveri tanımlamalarının önemini vurgulamıştır. Gökkurt ve Demirtel 2014 yılında yayımladıkları çalışmada sosyal medyada kurumsal arşivleme sorununu ele almıştır.

Bu bölümde yer verilen yerli araştırmacılara ait çalışmalar değerlendirildiğinde, çalışmaların genellikle kültür, kültürel miras, kültürel mirasın korunması, Web arşivleme, dijital koruma, sosyal medya arşivleme gibi alanları ele aldığı görülmektedir. Web arşivleme alanında Türkçe literatürün güçlendirilmesini destekleyen söz konusu çalışmalar tarafımızca oldukça değerli görülmektedir. Bununla birlikte Web arşivlemenin yasal, teknik, idari ve uygulamaya ilişkin yönlerini bütüncül bir bakış açısıyla yansıtan yeni ve kapsamlı araştırmalara ihtiyaç duyulduğu düşünülmektedir.

1.7. Araştırmanın Düzeni

Çalışmamız dört bölümden oluşmaktadır.

Birinci bölümde konunun önemi, araştırmanın amacı, araştırma problemleri ve hipotezler, araştırmanın kapsamı, yöntem ve veri toplama teknikleri, literatür incelemesi, araştırmanın düzeni ve yararlanılan kaynaklara ilişkin bilgiler sunulmaktadır.

İkinci bölümde geleneksel kültür mirası ve dijital kültür mirası konuları ele alınmaktadır. Geleneksel kültür mirası kapsamında kültür, kültürel miras, kültürel miras kategorileri, kültürel mirasın yönetimi, kültürel mirasın korunmasında uluslararası sözleşmeler ve kültürel mirasın ekonomik değerine ilişkin bilgiler aktarılmaktadır. Dijital kültür mirası kapsamında ise dijital miras, dijital mirasın korunmasında kütüphanelerin rolü, İnternet kültürü ve Web mirası ile Web mirasının korunması konularına yer verilmektedir.

Üçüncü bölümde Web derleme stratejileri, Web arşivleme yaşam döngüsü aşamaları, yasal derleme ve telif hakları, Web arşivleme yazılım araçları, Web arşivlemede standartlar, Web arşivlemede mevcut durum ve zorluklar, Web arşivlemede istatistik ve kalite göstergeleri ile Web arşivlerinde işlevsellik konularına yer verilmektedir. Ardından Web arşivlemenin evrimi, Web arşivleme girişimleri ve ulusal kütüphanelerin Web arşivleme uygulamaları aktarılmaktadır. Bu bölümde ayrıca çalışma kapsamında belirlenen hipotezler test edilerek araştırma problemlerine yanıt aranmakta ve Türkiye Milli Kütüphanesi'ne yönelik bir değerlendirmeye yer verilmektedir.

Dördüncü bölümde ise çalışmamızın sonuçları aktarılmakta ve sunulan öneriler yer almaktadır.

1.8. Yararlanılan Kaynaklar

Araştırma sürecinde yararlandığımız veri tabanları,

- British Standards Online (1927-)
- EBSCOHost Research Databases (1975-)
- Elsevier ScienceDirect (1936-)
- Emerald Insight (1967-)
- ProQuest Dissertations & Theses Global (1938-)
- ProQuest Ebook Central (1938-)
- ProQuest Periodicals Archive Online (1938-)
- ResearchGate (2008-)
- Scopus (1900-)
- Springer Link (1993-)
- Taylor & Francis (1954-)
- Wiley Interscience (1986-)
- World eBook Library (1996-) olarak sıralanabilir.

Söz konusu elektronik veri tabanlarına ek olarak araştırmacılara bilgi sunma amacıyla Web arşivleme alanında çalışmamıza temel oluşturan bazı kaynakların bu bölümde vurgulanmasının yararlı olacağı düşünülmektedir. Bu kaynaklar, ISO'nun 2009 yılında yayımladığı WARC dosya formatı standardı ve 2012 yılında yayımladığı istatistik ve kalite konularını ele alan teknik rapor taslağı, Internet Archive 'ın 2013'te yayımladığı Archive-It Web Arşivleme Yaşam Döngüsü Modeli, Costa, Gomes ve Silva'nın 2016 yılında yayımladığı Web arşivlemenin evrimini yansıtan çalışma ile Pennock'un Web arşivleme faaliyetlerini ele aldığı 2013 tarihinde yayımlanan teknoloji izleme raporu olarak sıralanabilir. Çalışmada yer alan ve Web arşivlemede kullanılan yazılım araçlarına yönelik tablonun hazırlanmasında ise bir açık kaynak uygulaması ile yazılım dizini olan Sourceforge [75] ve bir yazılım geliştirme platformu olan GitHub [147] Web sitelerinden yararlanılmıştır.

Bu kaynaklara ek olarak çalışmamızda Şencan ve Doğan tarafından bilimsel yayınlara kaynak gösterme yöntemlerini sunma amacıyla hazırlanan ve 2017 yılında yayımlanan bir rehberden de yararlanılmıştır. Çalışmamızın ekinde (Ek 1) sunulan ve Web arşivlemeye ilişkin Türkçe terminolojiyi destekleme ve güçlendirme amacı taşıyan kavram listesinin hazırlanma aşamasında bazı kaynakların terminoloji bölümlerinden ve sözlüklerden yararlanılmıştır. Hazırlanan kavram listesine kaynaklık eden ve tarafımızca yetkin olduğu düşünülen başlıca dokümanlar: Reitz tarafından kütüphanecilik ve bilgi bilim alanlarına yönelik olarak çevrimiçi bir sözlük biçimde hazırlanan ve 2002'de yayımlanan ODLIS, Ruusalepp tarafından 2003'te dijital koruma terminolojisine yönelik olarak hazırlanan AHDS, ISO'nun 2013'te yayımladığı istatistik ve kalite konularını ele alan teknik raporu ve 2017 yılında National Archives tarafından yayımlanan rehberdir.

İKİNCİ BÖLÜM

KÜLTÜREL MİRASININ KORUNMASI

2. KÜLTÜREL MİRASIN KORUNMASI

18. yüzyılda yaşanan sanayi devrimiyle birlikte insan, hayvan, su ve rüzgâr gücünün yerini buhar gücü almış, tarım ve sanayi faaliyetlerde makine kullanımı yaygınlaşmıştır. Topraksız kalan insanların fabrikalarda ücretli işçi olarak çalışma amacıyla kentlere göç etmesi kentlerdeki yerleşimi ve yaşamı farklılaştırmıştır (Aksoy ve Ünsal, 2012:146). Sanayi devrimi tıpkı tarım devrimi gibi, ekonomik, sosyal ve kültürel boyutlarıyla insanların düşünce ve yaşam biçimlerinde ve yaşadıkları çevrede köklü değişimlere yol açmıştır.

Tarım ve sanayi devrimlerinin ardından bilgi ve iletişim teknolojilerinde yaşanan gelişmeler ve Soğuk Savaş'ın ardından gelen tek kutuplu siyasal yapı, küreselleşmeyi ortaya çıkarmış ve bu durum toplumları yakınlaştırarak sosyal ve ekonomik yapıda gözlenebilir değişimlere yol açmıştır. Kültürlerarası etkileşim artarken, temel üretim araçları olan emek, sermaye ve doğal kaynakların yerini bilgi almıştır. Bilginin ekonomik bir kaynak olarak kullanımı ve hızlı veri transferine imkân veren Internet ve Web teknolojilerinin yaygınlaşması haberleşme hızında artışa ve iletişim maliyetinde düşüşe neden olarak toplumsal bir dönüşüm gerçekleşmiştir. Bu süreç bilgi toplumu olarak tanımlanan ve teknoloji kullanımı yoğun olan toplum yapısını ortaya çıkarmıştır.

İnsanlığın ilk kültür ürünlerinden bugüne kadar ortaya çıkan eserler somut ve somut olmayan kültürel miras ile doğal miras başlıkları altında toplanmıştır. Zamanla miras ürünlerinin kapsamı genişlemiş belgesel, endüstriyel ve dijital miras ürünleri de kültürel miras kapsamına dâhil edilmiştir. Günümüzde kültürel, bilimsel, sanatsal ve tarihi eserler başta olmak üzere tüm miras alanlarının korunması ve gelecek nesillere aktarılması uluslararası sözleşmelerle zorunlu tutulmuştur. Kültürel mirasın kalıcılığını sağlama amacıyla bellek kurumları olan kütüphane, arşiv ve müzeler hizmet alanlarına göre ulaşabilecekleri her türlü kültür varlığını fiziksel veya dijital ortamda toplamakta, korumakta ve gelecek nesillere aktarmaktadır.

Bu bölümde kültür, kültürel miras, somut ve somut olmayan kültürel miras ile belgesel mirasa yönelik bilgiler geleneksel kültürel miras başlığı altında toplanmış, İnternet kültürü ve Web mirası ise dijital kültürel mirasın alt başlıkları olarak ele alınmıştır. Bununla birlikte, dijital kültür mirasının korunması bağlamında kıymetli dijital nesneleri barındıran Web mirasının önemi vurgulanmıştır.

2.1. Geleneksel Kültür Mirası

Bu bölümde kültür ve kültürel miras kavramları, kültürel miras kategorileri, kültürel mirasın yönetimi, kültürel mirasın korunmasında uluslararası sözleşmeler, kültürel mirasın tarihçesi ve kültürel mirasın ekonomik değeri aktarılmaya çalışılacak ve değerlendirilecektir.

2.1.1. Kültür Kavramı

Kültür kavramı tarihsel, toplumsal ve bilimsel derinliğe sahip, farklı disiplinlere konu olan maddi ve bilişsel unsurları içerdiğinden literatürde kavrama ilişkin yapılan tanımların sayısı oldukça fazladır. Bu tanımlara bakıldığında, kültür;

- Bir toplumu ya da toplumsal bir grubu tanımlayan belirgin maddi, manevi, zihinsel ve duygusal özelliklerin bileşiminden oluşan bir bütün; sadece bilim ve edebiyatı değil, aynı zamanda yaşam biçimlerini, insanın temel haklarını, değer yargılarını, geleneklerini ve inançlarını da kapsayan bir olgu (United Nations Educational, Scientific and Cultural Organization [UNESCO], 1982:1);
- Tarihsel, toplumsal gelişme süreci içinde yaratılan bütün maddi ve manevi değerler ile bunları yaratmada, sonraki nesillere iletmede kullanılan, insanın doğal ve toplumsal çevresine egemenliğinin ölçüsünü gösteren araçların bütünü [148];
- Bir insan topluluğunun, bir halkın ya da bir ulusun düşünce ve değer birliğini meydana getiren düşünsel, sanatsal, teknik, felsefi tüm üretim ve varlıkları (Oğuz, 2011a:126);
- Sanat, edebiyat ve yaşam biçimleri ile değer, gelenek ve inançları kuşatan topluma veya bir sosyal gruba özgü maddi, manevi, entelektüel ve duygusal özellikler (UNESCO, 2001:1) olarak farklı biçimlerde tanımlanmıştır.

Bu tanımlardan yola çıkarak kültür, toplumsal yaşamı meydana getiren düşünce ve davranışlar, sonraki nesillere aktarılan miras ürünleri, öğrenilen ve edinilen beceri, alışkanlıklar ile bilgi, yetenek, inanç, dil, değer, sembol ve araçlar olarak yorumlanabilir.

Kültür kavramı aktarılabilir ve dinamik niteliği sayesinde tarih boyunca sürekli değişime uğramış ve farklı düşünce sistemlerinde yer almıştır. Başlangıçta bitki ve hayvan besleme, yetiştirme anlamına gelen kavram sonrasında insan zihninin üretimi ve halkın genel hayat tarzını belirleyen ruhani yapılanış gibi anlamlar kazanmıştır (Oğuz, 2011a:127). Tarihsel süreçte zihinsel, manevi ve estetik gelişim, toplumların yaşam biçimleri veya düşünsel ve sanatsal etkinliğin ürünleri de kültür kavramıyla açıklanmıştır.

İnsanın doğaya karşı verdiği savaşın ve toplumsallaşma sürecinin bir ürünü olan kültür, zaman ve mekân içerisinde çeşitli biçimler alır ve bu çeşitlilik toplumsal kimliklerin özgünlüğünde ve çoğulluğunda yansıma bulur (UNESCO, 2001:1). Yapısına bakıldığında kültür insan, doğa ve toplum etkileşiminin bir ürünü olarak yerel, ulusal veya evrensel değerleri ifade eder. Topluma ve bireylere göre değişkenlik gösterir.

Kültür insanın yaşamı boyunca edindiği bilgi ve tecrübeler ile zihinsel, dini ve estetik öğelerden oluşmaktadır. Öğrenilebilir ve süreklidir. Değişime ve etkileşime açıktır fakat derin köklere sahip oluşu nedeniyle kültürün yapısında ani ve toptan değişimler görülmez (Ölçer Özünel, 2013:17).

Kültür kavramı doğası gereği sürekli devinim halindedir ve koşullara göre biçimlenerek dönüşür (Ölçer Özünel, 2013:15). Miras kavramı ise geçmişten geleni ve durağanlığı yansıtır. Bir neslin kendinden sonra gelen nesle bıraktığı şey [148] veya geçmişten aldığımız, bugün bizimle birlikte yaşayan ve gelecek nesillere aktaracağımız birikimler anlamına gelen miras kavramı sıklıkla kültür kavramıyla iç içe kullanılmaktadır.

Kültür ve miras konulu çalışmalarda kültürün nesilden nesile aktarılan ve birikimli olarak çoğalan bir kavram olduğu, kültürel mirası değerli kılan şeyin onun toplum tarafından kullanılacak olmasından kaynaklandığı vurgulanmaktadır (Oğuz, 2011b:23). Bu bağlamda kültürel miras geçmiş ve gelecek arasında toplum tarafından önemli görülen değerleri aktaran tarihi bir köprü olarak düşünülebilir.

2.1.2. Kültürel Miras

Kültürü oluşturan ögeler bir toplumun gelişmişlik durumunu, eserlerini, varlıklarını kapsayan maddi kültür varlıkları ve geleneklerini, değerlerini, dilini, inançlarını, toplumsal kurallarını kapsayan manevi kültür varlıkları olarak iki grupta ele alınabilir. Bu varlıklar toplumların bilinen bir geçmişte yaşadıkları, biriktirdikleri, geliştirerek ve sürekliliğini sağlayarak gelecek nesillere aktardıkları veya aktarmak istedikleri veriler dizisi, o toplumun kültürel mirası olarak adlandırılabilir (Özel, 2014:160).

Kültürel miras, insanın binlerce yıllık yaşam deneyiminin, aklının ve yaratıcılığının bugüne ulaşmayı başarmış kalıntılarıdır (Aksoy ve Ünsal, 2012:34). Kavram ayrıca kültürümüzle, tarihimizle, kimliğimizle ilgili somut ve somut olmayan varlıkları ve değerleri ifade eder. Kaleler, saraylar gibi anıt eserler, arkeolojik alanlar, tarihi kentler ve dokular, kültürel peyzajlar kadar dil, gelenek, dans, müzik, tören gibi yaşayan ama somut olmayan değerler de kültürel mirastır (Aksoy ve Ünsal, 2012:34).

Kültürel miras geçmişle bugün arasında bağlantı kurarak, içinde yaşanan kültüre ve dünyaya bir temel oluşturur. Manevi anlamda ise insanların hayatlarını zenginleştirir (Aksoy ve Ünsal, 2012:34).

Kültürel miras tanımının kapsamı zaman içinde genişlemiş ve zenginleşmiştir. Anıt eserler üzerine odaklanan bir tanım ve koruma anlayışından, çok daha kapsayıcı ve insana ait tüm kültürel değerleri içeren bir kültürel miras anlayışına ulaşılmıştır (Aksoy ve Ünsal, 2012:34). Somut varlıkların yanı sıra somut olmayan eserler de kültürel miras kapsamında değerlendirilmeye başlanmış, daha sonra doküman formundaki eserlerden oluşan belgesel miras ürünleri de bu kapsama dâhil edilmiştir.

2.1.3. Kültürel Miras Kategorileri

Geleneksel kültür mirası bu bölümde somut kültürel miras, somut olmayan kültürel miras ve belgesel miras olarak temelde üç kategoride ele alınmaktadır.

2.1.3.1.

Somut ve Somut Olmayan K lt rel Miras

Somut k lt rel miras kavramı ve genellikle uzun yaşıyan estetik, doęal veya k lt rel maddi varlıkları nitelemeye kullanılır. Anıtlar, mimari eserler, binalar ve yerleşim yerleri gibi yerleşik k lt r ve sanat  r nleri ile kitaplar, takvim, resim, halı gibi yerleşik olmayan k lt r ve sanat  r nleri somut k lt rel miras kapsamında deęerlendirilmektedir. Somut k lt rel miras sessiz ve hareketsiz olarak nitelenen k lt rel miras  r nlerini i erir.

B t n insanlığın ortak mirası olarak kabul edilen somut k lt rel miras  r nleri Birleşmiş Milletler Eęitim, Bilim ve K lt r Teşkilatı (UNESCO) tarafından 16 Kasım 1972 tarihinde ‘‘D nya K lt rel ve Doęal Mirasının Korunmasına Dair S zleşme’’ ile D nya Mirası Listesine alınarak korunmaya başlamıştır. S zleşmenin ardından  lkemiz sınırları i erisinde yer alan somut k lt rel miras varlıklarının D nya Mirası Listesi’ne alınması i in bařvurular yapılmıştır. Bug n, D nya Mirası listesine alınan ve k lt rel miras olarak tescillenen eser sayımız 15; hem k lt rel hem doęal miras kategorisine giren eser sayımız ise 2’dir [149].

Somut olmayan k lt rel miras ise bizlere atalarımızdan miras kalan atas z  ve masal gibi s zl  gelenek ve ifadeler, kukla oyunları ve tiyatro gibi g steri sanatları, d ę n ve kutlama gibi sosyal uygulamalar, doęa ve evren ile ilgili bilgiler ve deneyimlerden oluşur. Somut olmayan k lt rel miras  r nleri geleneksel,  aędař, kapsayıcı, orijinal veya topluma  zg  olma gibi  zelliklere sahiptirler. Dinamik ve etkileşime a ık yapıları nedeniyle yaşıyan miras  r nleri olarak deęerlendirilirler.

UNESCO tarafından 2003 tarihinde Paris’te Somut Olmayan K lt rel Mirasın Korunması S zleşmesi kabul edilmiştir. S zleşme’de somut olmayan k lt rel miras toplulukların, grupların ve kimi durumlarda bireylerin, k lt rel miraslarının bir par ası olarak tanımladıkları uygulamalar, temsiller, anlatımlar, bilgiler, beceriler ve bunlara ilişkin ara lar, gere ler ve k lt rel mek nlar olarak tanımlanmıştır (Arıoęlu ve Aydoędu Atasoy, 2015:113). Somut Olmayan K lt rel Mirasın Korunması S zleşmesi, somut olmayan k lt rel miras konusunda d nya  apında farkındalık kazandırmayı ile bu mirası yaymayı ve korumayı zorunlu kılar (Pietrobruno, 2013:1260).

UNESCO Somut Olmayan K lt rel Mirasın Korunması S zleşmesi’ne g re bir k lt rel mirasın, insanlığın somut olmayan k lt rel mirası olarak ilan edilebilmesi i in řu beř kategoriden birine d hil olması gerekmektedir [5]: s zl  gelenek ve anlatımlar; toplumsal uygulamalar, rit eller ve ř lenler; doęa ve evrenle ilgili bilgi ve uygulamalar; g steri sanatları ve el sanatları geleneęi. S zleşme’de ayrıca somut olmayan k lt rel

mirasın kalıcılığını sağlama amacıyla kullanılabilecek yöntemler: kimlik saptaması, belgeleme, araştırma, muhafaza, koruma, geliştirme, güçlendirme, örgün veya yaygın eğitim yoluyla kuşaktan kuşağa aktarma ve kültürel mirasın değişik yanlarının canlandırılması olarak gösterilmiştir (Arioğlu ve Aydoğdu Atasoy, 2015:114-115).

Somut olmayan kültürel miras dünya genelinde artan küreselleşmenin bir sonucu olarak yaşanan tektipleşme karşısında kültürel çeşitliliğin korunmasına yardımcı olur. Sözleşmeye konu olan kültürel çeşitlilik kavramı, farklı kültürlerin yaşatılması adına herkese açık olan seçenekler yelpazesi sunar; sadece ekonomik kalkınma amaçlı değil, daha tatmin edici bir entelektüel, duygusal, ahlaki ve ruhsal varlık elde etmeye yönelik bir gelişimin de temel unsurlarından biridir (UNESCO, 2001:2). Sözleşmeye göre ifade özgürlüğü, medya çoğulculuğu, çok dillilik, sanata, bilimsel ve teknolojik bilgiye dijital formda da dâhil olmak üzere eşit erişim ve tüm kültürler için ifade ve yayma araçlarına erişme olanağı, kültürel çeşitliliğin garantileridir (UNESCO, 2001:2). Sözleşme kültürel çeşitliliği vurgulaması ve dijital kültür öğelerinin değerini ortaya koyması açısından önem taşımaktadır.

2.1.3.2. Belgesel Miras

Belgesel Miras kaybedilmesi veya zarar görmesi halinde yoksullaşmaya neden olabilecek, bir topluma, kültüre veya genel anlamda insanlığa önemli ve kalıcı değer ihtiva eden tekil veya grup halindeki belgelerden meydana gelir (UNESCO, 2015:2). Belgesel mirasa konu olan belge kavramı analog veya dijital bilgi içeriği ile içeriği barındıran taşıyıcıdan oluşan hareket edebilir ve saklanabilir nitelikteki nesneyi tanımlamak için kullanılır (UNESCO, 2015:2).

Belgesel miras ulusal hafızanın ve kültürel kimliğin taşıyıcısı konumundadır. Kuşaklar arasında sosyal, siyasal, tarihsel ve bilimsel bilginin aktarımını sağlar. Genellikle kütüphane, arşiv, müze gibi bellek kurumlarında saklanan belgesel miras ürünleri kırılgan ve hassas yapıdadır. Bununla birlikte, tarih boyunca ihmal, yağmalama veya savaşlar nedeniyle zarar görmüş veya yitirilmiş oldukları için arşivlenerek korunmaları ve sınırlama olmaksızın tüm insanlığın erişimine sunulmaları gerekmektedir.

UNESCO 2012 yılında belgesel mirasın korunması amacıyla Dünya Belleği Programı'nı (Memory of the World Programme) başlatmıştır. Programın temel amaçları belgesel miras ürünlerinin korunmasına yönelik mali kaynağın sağlanması, kamuoyu oluşturulması ve belgesel miras ürünlerine evrensel düzeyde erişim sağlamaktır. Program ile ayrıca insanlığın tarihî, kültürel ve sosyal belleğini oluşturan ve savaşlar başta olmak üzere çeşitli doğal felaketler veya sosyal nedenlerle ortadan kaybolma tehlikesi olan belge

ve bilgilerin insanlığın ortak değerleri olarak korunması ve bu korumanın önlemlerinden biri olarak dijital ortamda paylaşılmasını sağlamak amaçlanmaktadır. Bu amacı gerçekleştirmek için program, bilginin saklandığı ortam (analog, dijital) her ne olursa olsun, belgesel miraslarının en uygun teknikler kullanılarak korunması, belgesel miraslara evrensel erişimin sağlanması, belgesel mirasın varlığına ve önemine yönelik dünya çapında kamuoyu oluşturulmasını destekler.

Programla hedeflenenler “üretim, iletişim ve bilgi paylaşımı konularında daha geniş imkânlar yaratarak tüm insanlar arasında dijital mirasa erişim farkındalığı yaratmak; giderek artan bir şekilde dijital formda üretilen, dağıtılan, erişilen ve muhafaza edilen bilgi kaynaklarını tanımlayarak ve yeni ifadeler ortaya koyarak yeni bir miras oluşturmak; dijital mirasın kaybolma riski altında olduğunu kavrayarak, şimdiki ve gelecek nesillerin yararı için mirasın korunmasının önemini dünya çapında vurgulamak” olarak sıralanabilir.

2.1.4. Kültürel Mirasın Yönetimi

Kültürel miras dünyada ağırlıklı olarak devlet örgütlenmeleri içinde yönetilmektedir. Kültürel miras varlıkları devlete aittir ve koruma sorumluluğunu devlet icra eder. Bunun sebebi mirasın kültür, tarih, kimlik gibi ulus kavramını tanımlayan değerlerden biri olarak tanımlanmış olmasıdır (Aksoy ve Ünsal, 2012:61).

Türkiye Cumhuriyeti 1982 Anayasası’nın 63. maddesinde “Devlet, tarih, kültür ve tabiat varlıklarının ve değerlerinin korunmasını sağlar, bu amaçla destekleyici ve teşvik edici tedbirleri alır.” [6] hükmü yer almaktadır.

Türkiye’de Kültür ve Turizm Bakanlığı kültürel miras yönetiminden doğrudan ve tek elden sorumludur. Merkezde Vakıflar Genel Müdürlüğü, Milli Saraylar Genel Müdürlüğü ve Millî Savunma Bakanlığı gibi kurumlar, yerelde ise belediyeler ve il özel idareleri kültürel mirasın kamu yönetimi içindeki diğer aktörleridir (Aksoy ve Ünsal, 2012:61).

4848 sayılı Kültür ve Turizm Bakanlığı Teşkilât ve Görevleri Hakkında Kanun ile kurulan Kültür ve Turizm Bakanlığı; millî, manevî, tarihî, kültürel ve turistik değerleri araştırmak, geliştirmek, korumak, yaşatmak, değerlendirmek, yaymak, tanıtmak, benimsetmek ve bu suretle millî bütünlüğün güçlenmesine ve ekonomik gelişmeye katkıda bulunmak; tarihî ve kültürel varlıkları korumak ile görevlendirilmiştir

(2003:madde 2). Aynı Kanun'da K lt r ve Turizm Bakanlıęı'nın Merkez teŐkilatında yer alan "Mill  K t phane BaŐkanlıęı, T rkiye'nin mill  k lt r ve bilgi birikimi ile bilgi akımını saęlar." h k mleriyle yetkili ve sorumlu kılınmıŐtır. Milli K t phane'nin g revleri ise 5632 sayılı Mill  K t phane KuruluŐu Hakkında Kanun ile belirlenmiŐtır (1946:madde 2). Bakanlık Makamının 23/02/2018 tarih, 121771 sayılı onayı ile y r rl ęe giren Mill  K t phane BaŐkanlıęı Elektronik Yayınların Derlenme Usul ve Esasları Hakkında Y nergede İnternet sayfası (Web) derleme baŐlıęı altında "DeęiŐken i erięe sahip, belirli aralıklarla g ncellenen internet sayfaları (web siteleri) bu Y nergenin yayımından itibaren    (3) yıl sonra derlenmeye baŐlanır" ge ici maddesinin yer alması [156], k lt rel mirasımızın bir par ası olan Web mirasının y netimi adına  nemlidir.

K lt rel mirasın y netimine iliŐkin esaslar 2863 sayılı K lt r ve Tabiat Varlıkları Koruma Kanunu kapsamında d zenlenmiŐtır. Kanun'un amacı, "Korunması gerekli taŐınır ve taŐınmaz k lt r ve tabiat varlıkları ile ilgili tanımları belirlemek, yapılacak iŐlem ve faaliyetleri d zenlemek, bu konuda gerekli ilke ve uygulama kararlarını alacak teŐkilatın kuruluŐ ve g revlerini tespit etmek" olarak aktarılmıŐtır. (1983:madde 1).

K lt rel mirasın korunması devletin sorumluluęunda olsa da mirasın muhafaza edilmesi ve aktarımında bilgi merkezleri  zel bir  neme sahiptir. Bu kurumlar arasında ulusal hafıza kurumları olarak da anılan ulusal k t phaneler  ne çıkmakta ve bilhassa belgesel ve dijital mirasa y nelik koruma ve arŐivleme sorumluluęunu  stlenmektedirler.

Kamu sekt r nde k lt rel miras y netimi alanında s z sahibi bir dięer kurum Vakıflar Genel M d rl ę 'd r. Kurumun teŐkilat ve g revlerini tanımlayan 227 sayılı Kanun H km nde Kararnameye g re, Genel M d rl k "vakıf mallarını ekonomik bir Őekilde iŐletmek, mimari ve tarihi deęere sahip vakıf eski eserleri muhafaza ve imar etmek ve vakfa ait m esseseleri gayelerine g re yaŐatmak" ile sorumludur.

21. y zyılın ilk yıllarında  ıkarılan yeni yasalarla k lt rel miras y netiminde yerel y netimlere de yetkiler verilmiŐtır.  zellikle b y kŐehir belediyeleri ve il  zel idarelerine koruma, deęerlendirme ve denetleme alanlarında yetki devri olmuŐtur. Bu durum k lt rel mirası korumaya y nelik yaklaŐımların zaman i erisinde devlet merkezci ve yapı odaklı anlayıŐtan insan merkezci ve evrensel deęer odaklı anlayıŐa doęru bir d n Ő m g sterdięini ve hukuki zemin kazandıęını g sterir.

Ulusal ve uluslararası  l ekte somut ve somut olmayan k lt rel mirası korumaya y nelik ortaya  ıkan y netim anlayıŐı bilgi ve iletiŐim teknolojilerinde yaŐan geliŐmelere paralel olarak yeni meydan okumalarla karŐılaŐmıŐtır. İnternet ve Web teknolojisindeki hızlı ilerleme dijital miras  r nlerine y nelik yeni koruma ve arŐivleme gereksinimleri

doğurmuş, bilginin paylaşımında mülkiyete ve sorumluluğa yönelik daha fazla çalışma ihtiyacı belirmiştir.

İnternet kültürünün dünya çapında yaygınlaşması ve bunun bir sonucu olarak Web ortamındaki dijital nesnelerin boyutunda ve sayısında görülen artış kültürel mirasın korunmasından ve yönetiminden sorumlu olan kamu kurumlarına yeni sorumluluklar yüklemiştir. Bazı kamu kurumları İnternet kültürünün yönetimi ve gelişimi için plan, politika, rehberleri hazırlamakla ve ulusal kültür faaliyetlerini yönetmekle yükümlü kılınmıştır (Ministry of Culture of the People's Republic of China, 2016:135). Dijital mirasın boyutunda görülen artış, bilgi formlarının çeşitlenmesi ve belgesel mirasın dijitalleştirilmesi sonucu ortaya çıkan dijital arşivlerin yönetimi bilgi merkezlerini yeni politikalar ve stratejiler geliştirmeye yöneltmiştir. Kültürel mirasın yönetiminde karşılaşılan zorluklarla mücadelede en büyük sorumluluk şüphesiz ulusal kütüphanelere aittir.

2.1.5. Kültürel Mirasın Korunmasında Uluslararası Sözleşmeler

Kültürel miras tarih boyunca savaş, ihmal ve bilgi eksikliği nedeniyle ihmal edilmiştir. Öneminin ve değerinin anlaşılmasıyla birlikte uluslararası alanda kültür varlıklarının korunmasına ilişkin ilk düzenlemelerin savaş ve silahlı çatışma halinde kültür varlıklarının başına gelen zararlarla ilişkili olduğu görülecektir (Aksoy ve Ünsal, 2012:35). Kültürel mirasın korunmasına yönelik sözleşmeler ağırlıklı olarak UNESCO'nun öncülüğünde düzenlenmiştir. Avrupa Konseyi'nin de kültürel mirasın korunmasına yönelik çalışmaları mevcuttur. Tarihsel sürece bakıldığında;

1907 tarihli Lahey Kuralları kültürel mirasın korunmasına yönelik yapılan ilk çalışma olarak dikkat çekmektedir (Başlar, 2011:25). Ayrıca, 1949 tarihli Cenevre Sözleşmesi'nin kültürel varlıkların ve ibadet yerlerinin korunmasına yönelik 16. Maddesi'nde, silahlı çatışma durumunda halkların kültürel ve manevi mirasını oluşturan tarihi anıtlara, sanat eserlerine ya da ibadet yerlerine karşı herhangi bir düşmanca davranışta bulunmak ve bu yapıları askeri çabaları desteklemek için kullanmak yasaklanmıştır (Galatasaray Üniversitesi ve International Committee of the Red Cross, t.y.:274).

1954 tarihli Silahlı Bir Çatışma Halinde Kültür Mallarının Korunmasına Dair Sözleşme (Lahey Sözleşmesi) silahlı çatışma durumunda kültürel mülkiyetin korunmasını şart koşmakta ve bu şartlara uyulmaması durumunda cezai yaptırımlar getirmektedir (UNESCO, 1954; aktaran Tonta, 2014:2). Sözleşme'de kültürel miras,

anıtsal mimari eserler, tarihi ya da sanatsal öneme sahip yapılar ve arkeolojik sit alanları gibi taşınmaz kültür varlıklarının yanı sıra resim, heykel gibi sanat eserleri, kitaplar, arşivler, el yazmaları, bilimsel açıdan önemli koleksiyonlar gibi sanatsal, tarihi veya arkeolojik öneme sahip taşınabilir kültür varlıkları olarak açıklanmıştır (Council of Europe, 2005:3).

1964 tarihli Venedik Tüzüğü ile uygarlığa mal olmuş ve tarihe tanıklık eden kentsel ve kırsal yerleşim birimleri ile daha basit eserler de kültürel miras kapsamına alınmıştır.

1972 tarihli UNESCO Dünya Mirası Sözleşmesi bahsi geçen sözleşmeler arasında kültürel ve doğal mirasların korunması, tanıtılması ve çeşitli sebeplerle bozulan, yok olan kültürel ve doğal değerlerin yaşatılması için gerekli iş birliğinin sağlanması amacıyla hazırlanmıştır. Sözleşme ile birlikte üstün evrensel değere sahip olduklarına inanılan, özgünlük ve bütünlük şartlarına haiz kültürel miras varlıkları insanlığın ortak malı sayılarak Dünya Mirası statüsü kazanır.

2003 tarihli Somut Olmayan Kültürel Mirasın Korunması Sözleşmesi'nde kültürel çeşitlilik vurgulanmış, somut olmayan kültürel miras zihinsel, ruhsal ve ahlaki gelişimin bir öncülü olarak kabul edilmiştir.

2011'de yürürlüğe giren Avrupa Konseyi Toplum İçin Kültürel Mirasın Değeri Çerçeve Sözleşmesi herkesin kültürel mirastan yararlanma ve onu sağlama hakkı olduğunu belirtir. Sözleşmede kültürel miras hem insani gelişme, kültürel çeşitliliğin artırılması ve kültürler arası diyalogun geliştirilmesi için bir kaynak hem de sürdürülebilir kaynak kullanım ilkelerine dayanan ekonomik kalkınma modelinin bir parçası olarak tanımlanmıştır (Tonta, 2014:11).

Kültürel miras kavramı yakın bir zamana kadar nesne odaklı ve fiziksel yapı ölçeğinde dar çerçeveli bir yaklaşımla ele alınmıştır. Birinci ve İkinci Dünya Savaşları'nın yarattığı yıkım ve kentlerde meydana gelen hızlı dönüşümün ardından kültürel miras varlıklarının korunması konusunda farkındalık artmış ve kavramın çerçevesi genişlemeye başlamıştır. Böylece, kültürel miras, salt bilimsel incelemelere konu olan tarihi ve sanatsal değere sahip varlık olmanın ötesinde, toplulukların ve bireylerin kültürel kimliğini oluşturan bir unsur haline gelmiştir (Aksoy ve Ünsal, 2012:35).

2.1.6. Kültürel Mirasın Ekonomik Değeri

Kültürel miras bir yönüyle ekonomik değeri olan bir kamu malı diğer taraftan devletlerin sürdürülebilir kalkınma politikaları için bir kaynaktır. Kültüre yapılan yatırımlar turizm, ulaşım, konaklama, eğitim başta olmak üzere birçok sektörü canlandırarak üretim artışına neden olmakta ve sürdürülebilir kalkınmayı desteklemektedir. Günümüzde kültüre ilişkin sermaye, emek ve mal dolaşımı ulusal sınırları aşarak dolaşıma katılmakta, dünya tek pazar haline gelmektedir.

Kültür ekonomisi, kültür alanındaki üretim ve tüketim ilişkilerini inceleyen, kültür sektörünü ekonomi teorileri ve istatistikleriyle analiz eden ekonomi biliminin alt alanlarından biridir (Aksoy ve Ünsal, 2012:31). Kültür ekonomisi film, müzik, yayın, medya, gösteri sanatları, görsel sanatlar veya kültürel miras gibi alanları inceler. Bu ekonominin merkezinde yaratıcılık ve entelektüel sermaye girdileriyle oluşturulan faaliyet, hizmet ve ürünler yer alır. Kültürel miras, kültür ekonomisinin en değerli parçası konumundadır. Kültürel miras ürünleri özgündür ve yeniden üretilemez bu nedenle ekonomideki diğer ürünlerden ayrılır.

Kültürel miras tarihi, toplumsal ve kültürel değerlerin yanı sıra yerel ve ulusal ekonominin önemli bir bileşenidir. Kültür endüstrileri aracılığıyla üretilen ve dolaşıma sokulan küresel mal ve hizmetlerin değeri, topluma sağladığı yararlar ölçülmektedir. Kültürel mirasın ekonomik değerinin ölçülmesine yönelik birkaç örneğe bakalım;

- New York Metropolitan Müzesi'nde yaklaşık 4 ay sergilenen Picasso Sergisi'nin kişi başı ortalama 1,050 Amerikan Doları gelir bırakan 703,256 ziyaretçiye ev sahipliği yaptığı belirlenmiştir (Hayırsever Baştürk ve Dimitriadis, 2012:110).
- ABD'nin Pennsylvania Eyaleti'nde düzenlenen Salvadore Dali Sergisi'nin ekonomiye doğrudan katkısının 54,9 milyon Avro civarında olduğu tespit edilmiştir (Hayırsever Baştürk ve Dimitriadis, 2012:110).
- 2012 yılında Birleşik Krallık 'ta yer alan 28 kültürel miras varlığı için 15,7 milyon İngiliz Sterlini harcandığı ve sağlanan gelirin 76,9 milyon İngiliz Sterlini civarında olduğu hesaplanmıştır (Tonta, 2014:9). Bu çalışmadan kültürel mirasın fayda maliyet oranının yaklaşık beş kat olduğu anlaşılmaktadır.

Avrupa Konseyi ve UNESCO gibi kuruluşların kalkınma hedeflerinde kültürel mirasın korunması ve sürdürülebilir kalkınma programlarıyla entegrasyonuna yönelik maddelerin yer aldığı görülmektedir. Sürdürülebilir kalkınma yaklaşımı, içinde sosyal, çevresel ve ekonomik boyutları da barındıran bütünsel bir yaklaşımdır ve kalkınmayı yalnızca bir modernleşme süreci ya da ekonomik büyüme olarak ele almaz (Ölçer Özünel, 2013:17). Bu nedenle kültürel mirasın devletlerin ve kuruluşların sürdürülebilir kalkınma politikalarında yer alması önemlidir.

UNESCO tarafından 2013 tarihinde yayınlanan Hangzhou Deklarasyonu'nda:

Kültürün kalkınma politika ve planlarının tamamına entegre edilmesi, barışı ve sosyal uzlaşmayı güçlendirilmek için karşılıklı anlayış ve kültürün harekete geçirilmesi, sosyal kalkınmanın sağlanabilmesi için kültürel hakların herkese sağlanması, ekonomik refahın sağlanması ve yoksulluğun azaltılması için kültürün güçlendirilmesi, çevresel sürdürülebilirliği desteklemenin kültür kavramı üzerine inşa edilmesi, kültür aracılığıyla afetler ve iklim değişikliğiyle mücadele edilmesi, kültürün gelecek nesillere aktarımı için değerinin bilinmesi ve korunması, sürdürülebilir kentsel kalkınma için kültürden bir kaynak olarak yararlanılması, yenilikçi ve sürdürülebilir iş birliği modelleri geliştirmek için kültürden yararlanılması gibi önemli ifadeler yer almaktadır [7].

Bu konferansın hemen ardından somut olmayan kültürel mirasa yönelik Çin'de düzenlenen uluslararası bir konferansta (14-16 Haziran 2013, Chengdu) sürdürülebilir kalkınma stratejileri masaya yatırılmış, konferansın sonuç bildirgesinde sürdürülebilir kalkınma ve somut olmayan kültürel miras ilişkisine dair önemli tespitler yapılmış ve stratejiler belirlenmiştir (Ölçer Özünel, 2013: 19).

Kültür ürünleriyle birlikte bilginin ulusal üretimdeki ve istihdamdaki payı giderek artmaktadır. Sürdürülebilir kalkınma öncülünde entelektüel sermayeyi oluşturan bilgi varlıklarını korumak ve erişilebilir kılmak kültür ekonomisini güçlendirmektedir. Aynı zamanda bilginin akışkan, taşınabilir, değiştirilebilir, yenilenir, bölünebilir ve paylaşılabılır yapısı ekonomik faaliyetlerin her aşamasına hız, kalite, verim ve esneklik kazandırmıştır (Oğuz, 2011b:20).

2.2. Dijital Kültür Mirası

Dijital kültür mirası kapsamında dijital miras, dijital mirasın korunmasında kütüphanelerin rolü, Internet kültürü ve Web mirası ile Web mirasının korunması konuları incelenecek ve yorumlanacaktır.

2.2.1. Dijital Miras

Gün geçtikçe kültür mirası kapsamındaki kaynaklar ve bilgiler dijital formatta üretilmekte veya dijitalleştirilmekte ve bu bilgilere Web servisleri aracılığı ile erişilmektedir (Özbağ, 2010:v). Dijital ortamda üretilen ve analog kaynaklardan dijitale çevrilen materyaller birer kültür varlığı olarak değerlendirilmekte ve dijital miras kapsamında arşivlenerek korunmaktadır.

Dijital miras dijital formda üretilen veya mevcut analog kaynaklardan dijital forma dönüştürülen kalıcı değere ve öneme sahip, korunması ve muhafaza edilmesi gereken bilgi kaynakları ve kültür nesneleri olarak tanımlanabilir. Kapsamında metin, veri tabanı, hareketli ve hareketsiz görüntü, ses, grafik, yazılım ve Web sayfası gibi farklı türdeki dijital materyaller bulunur. Çok değerli kültürel miras varlıklarının taşıyıcısı konumunda olan dijital mirasın gelecek nesiller için korunması ve muhafaza edilmesi gerekir.

Dijital kültürel mirasın korunması dijital arşivleme faaliyetlerini zorunlu kılar. Dijital arşivlemenin temel amacı, dijital ortamda üretilen veya dijitalleştirilen bilginin saklanarak hizmete sunulması ve bu sayede dijital kültür mirasının gelecek kuşaklara aktarılmasıdır.

Dijital miras materyallerinin (metin, görsel, hareketli görüntü, ses kayıtları, müze materyalleri vd.) Internet üzerinden erişiminin etkin bir şekilde sağlanabilmesi ve korunabilmesi için arşivlenmesi büyük önem taşımaktadır (Özbağ, 2010:5). Arşivlenmesi planlanan dijital miras materyalinin seçimi tanımlanan politikalar ve geçerli standartlara dayandırılarak hesap verilebilir bir şekilde sürdürülmelidir. Arşivlenecek nesneler eğer dijital ortamda üretilmişse, analog kopyaları bulunmadığı için korunacak materyalin seçiminde bu nesnelere açık bir biçimde öncelik verilmelidir. Dijital arşivleme sürecinde üretimden erişime kadar dijital bilgi yaşam döngüsü boyunca tüm önlemlerin alınması gerektiği (UNESCO, 2003:75) ve dijital mirasın korunmasında bilgiye erişimin en az bilgiyi üretmek ve yaymak kadar değerli olduğu unutulmamalıdır (Oğuz, 2011:25).

Dünya genelinde birçok ülke dijital arşivleme çalışmalarını elektronik yayınları derleme kanunlarının desteğiyle yürütmektedir. Dijital mirasın korunması konusunda ulusal kamu politikasının temel bileşeni olan derleme yasalarının dijital mirası sahiplenmesi gerekir (UNESCO, 2003:76). Ülkemizde 2012 yılında yürürlüğe giren 6279 sayılı Çoğaltılmış Fikir ve Sanat Eserlerini Derleme Kanunu ve 28388 sayılı Çoğaltılmış

Fikir ve Sanat Eserlerini Derleme Yönetmeliği elektronik yayınların derlenmesine ilişkin esasları ve derlemenin kapsamını belirlemektedir.

Dijital arşivleme sürecinde kullanılan teknoloji ve yapılan iş birliklerinin yanı sıra miras ürünlerinin bibliyografik denetimi ve güvenilir arşiv sistemlerinin tasarımı dijital mirasın korunması ve sürekliliği adına oldukça önemlidir.

2.2.2. Dijital Mirasın Korunmasında Kütüphanelerin Rolü

Uluslararası Kütüphane Dernekleri ve Kuruluşları Federasyonu'nun (IFLA) dünya genelinde bibliyografik bilginin değişimini ve kontrolünü sağlama amacıyla ortaya koyduğu Evrensel Bibliyografik Kontrol anlayışı, derleme yasaları yardımıyla ulusal mirası toplayan ulusal kütüphanelere odaklanmakta ve her ülkeyi ulusal mirasın korunmasıyla sorumlu kılmaktadır (UNESCO,1974:5). Bu sorumluluktan hareketle, günümüzde giderek artan bir biçimde dijital ortamda üretilen yayınlar derleme yasaları kapsamına dâhil edilmektedir.

Kültürel mirasın korunması tüm insanlığın sorumluluğunda olmakla beraber dijital mirasın korunması, paylaşımı ve aktarımı konularında kütüphaneler özel öneme sahiptir. Kütüphanelerde yer alan özgün ve ulusal dijital kültürel miras materyalinin arşivlenerek korunması ve dolayısıyla bu materyale daha fazla insanın erişim sağlanması kültürel değerlerin yayılmasını kolaylaştıracak ve mirasın kalıcılığını arttıracaktır.

İnternet ve Web teknolojileri sayesinde dolaşıma sokulan, dijital formda üretilen veya dijitalleştirilen zengin kültür mirasının gelecek kuşaklara aktarılmasında, bilgiyi koruma ve aktarma misyonu olan kütüphanelere büyük sorumluluk düşmektedir. Ulusal kültür mirası kapsamında değerlendirilen ve sayıları hızla artan dijital formdaki bilgi kaynaklarına yönelik koruma ve arşivleme çalışmaları dünya genelinde Milli Kütüphaneler aracılığıyla veya öncülüğünde yapılmaktadır.

Kütüphanelerin koleksiyonlarındaki genellikle kısa ömürlü, belirli bir amaca yönelik üretilen, bakıma ve yönetilmeye ihtiyaç duyan dijital materyaller hızlı yazılım ve donanım eskimleri, koruma ve muhafaza konusunda yaşanan belirsizlikler ve destekleyici mevzuat eksikliği gibi nedenlerle gelecek nesiller için yok olma tehdidi altındadır. Kültürel mirasın korunmasında kültür varlıklarına yönelik envanter çalışmaları ve belgelendirme, ortaya çıkan verinin bir araya getirilmesi, yapılandırılması ve analiz

edilerek bilgiye dönüştürülmesi dijital koruma çalışmalarında önemli rol oynamaktadır. Dijital arşivi oluşturan kültürel miras materyallerinin bir araya getirilmesi, arşivlenmesi, uzun süreli korunması ve erişime sunulması bir bellek çeşit kurumu olan kütüphanelerin temel görevleri arasındadır.

Kültürel miras materyalinin yaşatılması ve uzun vadede kalıcılığının sağlanması erişim yoluyla görünürlüğü arttırarak mümkün olabilir. Tam da bu noktada dijital arşivlemeye ihtiyaç duyulmaktadır. Dijital koruma veya arşivleme, dijital materyallerin çeşitli teknikler ve stratejiler doğrultusunda saklanması ve erişilmesi için yapılan işlemlerin ve süreçlerin tamamı şeklinde tanımlanabilir (Özbağ, 2010:12). Çok katmanlı bir sistem veya basit bir bilgisayardan oluşan dijital arşiv ise bilgiye uzun süreli erişim sağlama amacıyla dijital bilgi nesnelerini saklayan depodur [8].

Dijital mirasın korunmasında bilgi merkezlerinin karşılaştığı temel zorluklardan biri dijital ortamda üretilen ve dijital ortama aktarılan kültürel miras ürünlerinin gelecek kuşaklara aktarımıdır. Sayıları giderek artan dijital nesneleri korumak ve gelecek kuşaklara aktarmak için kopyalamak, teknoloji eskidikçe bir ortamdan diğerine göç ettirmek ve bağlantıları güncelleştirmek gerekmektedir (Tonta, 2009: 757). Depolanmış dijital miras materyaline erişim, makul kısıtlamalar dâhilinde, suistimale ve önyargıya neden olmadan, yasalara uygun olacak şekilde sağlanmalıdır. Bunun için kamuya açık dijital miras materyaline erişim kısıtlamalarının kaldırılması, hassas ve gizli bilgilere ise yetkisiz ve izinsiz erişimlerin önlenmesi ve eser sahipleri ile kamunun çıkarları arasında bir denge sağlanması gerekmektedir.

Bilgi merkezleri bünyesinde sunulan dijital bilgi hizmetleri alt yapısı mevcut bilgilerin işlenmesine, depolanmasına, bir yerden bir yere iletilmesine ve bu bilgilere gerektiğinde erişilmesine olanak sağlayan teknolojileri, kuralları, standartları ve politikaları kapsar (Tonta, 2002:4). Dijital olarak üretilen bilgi ve kültür materyallerine yönelik arşivleme çalışmalarına ek olarak mevcut analog materyalin dijital ortama aktarılması ve bibliyografik tanımlamalarının yapılarak erişime sunulması dijital mirasın korunmasını destekler. Dijital mirası anlamlı ve kapsamlı hale getirmek için tanımlamak, kataloglamak, üstverileri hazırlamak ve bu verileri semantik ağ sayesinde Internet ve dünya çapında erişim için uygun hale getirmek gerekir (Fabian, 2014:28).

Kütüphaneler, arşivler ve müzeler günümüzde çeşitli dijitalleştirme projeleriyle sahip oldukları kültürel miras varlıklarını dijital ortamlara aktarmakta (Özel, 2014:160) ve çok uluslu dijital platformlarda insanlığın erişime sunmaktadır. Dijitalleştirme kültürel mirasın korunmasında ve gelecek nesillere aktarımında çok büyük bir öneme sahiptir. Dijitalleştirme herhangi bir formda yer alan bilginin çeşitli donanım, yazılım ve ağlar aracılığıyla dijital forma dönüştürülmesi olarak tanımlanabilir. Dijitalleştirmede amaç

özgün kopyalara zarar vermeden hızlı ve uygun maliyetle bilginin çoğaltılması ve hizmete sunulmasıdır. Dijitalleştirme nesnenin bütünlüğüne zarar verilmeden muhafaza edilmesine, kültürel mirasın küresel ölçekte paylaşılmasına ve mülkiyet ve sorumlulukları kaybetmeden bilgi ve görüntülerin dağıtımına izin verir (Fabian, 2014:15). Üretilen materyallerin ve materyallere ait bibliyografik künyelerinin paylaşılması dijital miras ürünlerine erişimi arttırmaktadır.

Özel'e göre dijital mirasın gelecek kuşaklara aktarımında dijitalleştirilecek kültür objelerinin değerlendirilmesi ve seçilmesi, dijitalleştirme için özgün kopyaların sağlanması, donanım ve yazılımların temin edilmesi, dijitalleştirmenin gerçekleştirilmesi, dijital koleksiyonların oluşturulması ve bilgi kaynaklarının ulusal ve uluslararası bilgi erişim sistemleri ile bütünleştirilmesi gibi tüm süreçlerin etkin bir biçimde planlanması ve yönetilmesi gerekmektedir (2014:160).

Kültürel miras varlıklarının dijital ortama aktarılması kullanım için tek başına yeterli değildir. Dijitalleştirme işleminin ardından görüntülere ait verilerin tanımlanması, erişim platformlarının oluşturulması ve söz konusu verilerin bilgi erişim sistemleri tarafından okunarak keşfedilmesi sağlanmalıdır.

Dijital kültür mirasına erişimde kütüphane hizmetlerinin önemli bir parçasını dijital kütüphaneler oluşturmaktadır. İnternet ve Web teknolojilerinde yaşanan ilerleme sonucunda kültür varlıklarının ve bilginin aktarımı dijital kütüphane platformları aracılığıyla sağlanmaktadır. Bu platformlarda yer alan dijital miras tüm ulusların, kültürlerin ve dillerin temsil edilmesine izin verecek şekilde korunmalı ve erişilebilir durumda olmalıdır.

Bilginin üretilmesi, araştırılması ve kullanılması amaçlarına yönelik bir dizi teknik yetenek ve güç sergileyen elektronik kaynaklar grubu (Subaşıoğlu, 2001:50) veya elektronik olarak erişilebilir dermesi bulunan ve herhangi bir yerden erişilebilen kütüphane olarak tanımlanan dijital kütüphaneler (Karataş Ateş, 2015:565), dünyanın birçok yerinde zamandan ve mekândan bağımsız olarak kullanıcıların bilgiye erişim taleplerini karşılamaktadır. Bu tanımda kavramın bir yüzü sanal kütüphaneleri tasvir ederken diğer yüzü depolanmış bilgiye ağ üzerinden erişimi mümkün kılan elektronik bir sistemi karşılar.

Dünyada Web tabanlı ilk dijital kütüphane girişimi 2000 yılında Kongre Kütüphanesi tarafından başlatılan ve 5 milyon kaynağın çevrimiçi erişilebilirliğini sağlamayı amaçlayan Ulusal Dijital Kütüphane Programı'dır (National Digital Library

Program) (Karataş Ateş, 2015:565). Kongre Kütüphanesi'nin bu programı daha sonra açılacak dijital kütüphanelere örnek olması bakımından önemlidir.

Programın ardından dünya genelinde uluslararası dijital kütüphane platformları yaygınlaşmıştır. E-LIS, Europeana, Internet Archive gibi projeler dijital olarak üretilen ve/veya dijitalleştirilen bilgi ve kültür varlıklarını kullanıcılarına sunarken, kültürel miras varlıklarının arşivlenerek kalıcılığının sağlanmasına ve korunmasına da yardımcı olurlar. Dijital kütüphanelerde standartların gelişmesi ve birlikte çalışabilir sistemlerin kurulabilmesi için birden çok veri tabanını aynı anda sorgulamayı destekleyen federe aramaların yapılabilmesi önemlidir. Federe arama, son kullanıcılar için tek bir arama sonuçları sayfası oluşturmak amacıyla birden çok çevrimiçi veri tabanının eşzamanlı olarak sorgulanması işlemi olarak tanımlanmaktadır [9].

Dünya genelinde Internet ağının yaygınlaşması, dijital bilgi üretiminde görülen artış ve dijitalleştirme uygulamalarının çoğalması dijital kütüphane platformlarının kullanımını arttırmıştır. Internet üzerinde dolaşıma sunulan bilgi ve kültür nesnelerin büyük oranda Web ortamında sunulması kütüphaneleri Web mirasını arşivleme konusunda cesaretlendirmiştir. Dijital kültür mirasının önemli bir parçası olan Web arşivleri de bir tür dijital kütüphane platformu olarak değerlendirilebilir.

2.2.3. Internet Kültürü ve Web Mirası

Internet, küreselleşme süreciyle birlikte başat kültürel formların ve kimliklerin oluşturulduğu kültürel bir ortam ve bunların hızlı bir şekilde dolaşıma sokulduğu teknolojik bir sistemdir (Güzel, 2006:1). Birbirine bağlı bilgisayar sistemlerinden oluşan ve sürekli büyüyen bu küresel iletişim ağı ülkeleri, insanları ve makineleri birbirine bağlamakta ve ortak değerler üzerine inşa edilen evrensel bir kültür meydana getirmektedir. Bilim, sanat, kültür, tarih ve daha birçok alanda dijital bilginin aktarımını sağlayan Internet, kültürel alanda değişim yaratabilecek bir güce ulaşmıştır.

Internet'in tarihi kökenine bakıldığında çalışmaların 1960'lı yıllarda başladığı görülür. 1969 yılında ABD Savunma Bakanlığı tarafından desteklenen "ARPAnet Projesi" bilgisayarlar arası veri iletişimini düzenleyen kurallardan oluşan TCP/IP protokollerinin kullanıldığı ilk ağıdır [10]. Internet, başlangıçta akademik bilgi aktarımı amacıyla tasarlansa da zaman içerisinde ticari ve diğer amaçlar için de kullanılmış ve 1990'lı yıllardan itibaren dünya genelinde yaygınlaşmaya başlamıştır. Dinamik yapısı nedeniyle sınırları tam olarak belirlenemese de bugün üç milyarın üzerinde kullanıcı sayısı ve bir milyarın üzerinde Web sitesiyle [11] Internet çağın en önemli kültür ağıdır.

Giderek gündelik hayatla bütünleşen Internet'te her gün binlerce kütüphaneyi dolduracak boyutta bilgi transferi sağlanmaktadır.

Internet, birbirine bağlı bilgisayar sistemlerinden oluşmanın ötesinde, makro düzeyde yeni bir kültürel, siyasal ve ekonomik alan olarak analiz edilmelidir (Güzel, 2006:2). Internet ve ona bağlı olarak gelişen Internet kültürü toplumları ve kıtaları birbirine bağlayarak yeni dijital bir kültür meydana getirmiştir. Günümüzde Internet kültürüne karşı temelde iki farklı bakış açısının ortaya çıktığı söylenebilir. İlk bakış açısına göre, egemen sınıfın ürünü olan geleneksel kültür gücü, otoriteyi veya devleti yansıtırken, Internet kültürü yeni, sınırsız, özgür, çok sesli ve demokratik bir alanı ifade etmektedir. İkinci bakış açısına göre ise, Internet hegemon güçlerin denetimi altındadır ve Web aracılığıyla dolaşıma sokulan kültür diğer kültürleri homojenleştirerek kendine benzetmekte ve ulusal kimlikleri yok etmektedir. Sonuç her ne olursa olsun Internet dijital formdaki kültür varlıklarının insanlığa sunulmasında muazzam fayda sağlamıştır. Fakat ulusal kimliklerin küreselleşme karşısında asimile olmaktan koruması Web mirasının korunmasını ve arşivlenmesini gerektirmektedir.

Internet ve Web kavramlarının literatürde sıklıkla aynı anlamda ve birbirinin yerine kullanıldığı görülür. Internet kablolar, sunucular, yönlendiriciler (router), anahtar (switch) ve bilgisayar gibi fiziksel ağın oluşumuna katkıda bulunan nesnelerden meydana gelen bir iletişim alt yapısıdır. Web ise bu alt yapı üzerinde çalışan bir sistemdir. Internet üzerinde ses aktarımı, dosya paylaşımı gibi hizmetler sunan başka sistemler/protokoller de mevcuttur. Web'de yer alan metin, ses, görüntü, hareketli görüntü gibi multimedya verilerin transferi Internet aracılığıyla bütünleşik olarak sağlanmaktadır.

WWW, W3 ve World Wide Web şeklinde ifade edilen Web kavramı dünya çapında ağ anlamına gelen bir bilgi paylaşım sistemidir. Web dinamik ve statik yapıdaki Web sitelerini kapsar. Arama motorlarının erişebildiği yüzeysel Web ve erişemediği veya sınırlı erişim sağlayabildiği derin Web olarak iki bölümden oluşur. Web, kullanıcılarının sorularına cevap bulabileceği her türlü bilgi kaynağını kapsayan bir referans kütüphanesi olarak da değerlendirilebilir (Oğuz, 2006a:289).

Dünya genelinde Internet ağının yaygınlaşmasıyla birlikte Web teknolojileri de gelişim göstermiştir. Web teknolojilerinin gelişimine bakıldığında, Web 1.0 kullanıcıların içeriği görüntüleyebildiği fakat değiştiremediği etkileşimsiz Web sürecini ifade eder ve tek yönlü iletişime olanak tanır. Web 2.0 kullanıcıların içerik oluşturabildiği ve insan-insan etkileşimine izin veren Web teknolojisidir. Bu süreçte kullanıcılar pasif ve tüketici konumdan aktif ve üretici konuma geçmiş, içerik oluşturmaya ve paylaşmaya başlamışlardır. Web 3.0 aktarılan içeriklerin anlamlandırıldığı, makine-makine

iletişimine izin veren ve kişiselleştirilebilen Web teknolojisidir. Web 4.0 ise yapay zekâya sahip işletim sistemi ve Web teknoloji mimarisi olarak tanımlanmaktadır [12].

Internet ve Web teknolojilerinin gelişmesinin bir sonucu olarak yerel ve evrensel kültürel miras nesnelerinin üretimi ve sunumu fiziksel ortamdan dijital ortama doğru kaymış ve yaşanan format değişimi ve bilginin boyutundaki büyüme, ulusları koruma konusunda yeni formüller geliştirmeye zorlamıştır. Ulusal veya evrensel ölçekteki dijital kültür mirasının korunmasında bilgi merkezleri önemli roller üstlenmişlerdir.

Web ortamında yaşanan bilgi artışı beraberinde bu bilgilerin gelecek nesillere aktarılması amacıyla arşivlenmesi gereğini gündeme getirmiştir (Oğuz, 2006a:283). Başka bir ifadeyle, dijital ortamda üretilen ve dijital ortama aktarılan bilgi ve kültür varlıklarının taşıyıcısı ve yayıncısı konumundaki Web sitelerinin gelecek nesillere aktarılması için arşivlenmesi ve korunması gerekmektedir. Dünyada birçok ülke milli kütüphaneleri öncülüğünde Web arşivleme çalışmalarını başlatmış ve bu girişimlerinin yasal bir platforma taşınması için gerekli adımlar atmıştır (Oğuz, 2006a:283). Web arşivleme teknik, yasal ve idari boyutları olan, bilgi yaşam döngüsündeki her bir aşamanın önemli olduğu çok boyutlu süreçlerden oluşur.

2.2.4. Web Mirasının Korunması

Web ve Internet teknolojisinde son yirmi yılda yaşanan gelişmeler bilginin boyutunda ve çeşitliliğinde artışa neden olmuş, korunmaya ve yönetilmeye ihtiyaç duyan dijital mirasın kontrolü ve denetimi zorlaşmıştır. Kültürel ve bilimsel değeri olan devasa boyuttaki dijital koleksiyonların korunması ve kalıcılığının ve aktarımının sağlanması zorluk yaşanan alanlar olarak karşımızda durmaktadır. Bu etkileyici ve hızla büyüyen aşırı bilgi yükünün muhafazası, genellikle bir ulusal kütüphane tarafından dikkatlice koordine edilen bir programın yürütülmesi ile mümkündür (Fabian, 2014:26).

Eşsiz bir bilgi kaynağı olan World Wide Web, dünya genelinde toplum ve bireyleri birbirine bağlayan milyarlarca sayfaya ev sahipliği yapmaktadır. Web sitelerinin ömrünün 100 günü geçmemesi [13] ve bilginin çoğunlukla (%93) dijital formda üretilmeye başlanması (Lyman ve Varian, 2003; aktaran Aldemir ve Oğuz, 2006:284) Web ortamında yer alan dijital bilginin korunması ve arşivlenmesi konusunda önemli hamleleri gerekli kılmaktadır. 2017 verilerine göre önemli arama motorlarından biri olan Google'ın indekslediği Web sayfası sayısı ortalama 45 milyar sayfa, Bing'in indekslediği Web sayfası sayısı ise ortalama 15 milyar sayfa civarındadır [14]. Arama motorlarının indekslediği bilgi kaynaklarına kolay erişim ve kaynak çeşitliliği kullanımı artırsa da düzenlenmemiş ve süzülmemiş bilgi eksik veya yararsızdır. Bu nedenle arama

motorlarının getirdiği sonuçlardaki içeriğin doğruluğu konusunda denetim mekanizmalarına ihtiyaç duyulmaktadır. Dijital kültürel mirasın bir parçası ve dinamik bir kültür hazinesi olan Web'in bibliyografik denetime, tanımlanmaya ve arşivlenmeye ihtiyacı vardır.

Küresel ölçekte çok hızlı gelişim gösteren bu bilgi ortamı, içinde bulunduğumuz çağın ve dijital kültür mirasımızın kaydını tutmaktadır. Web ile ilgili bir diğer önemli nokta Web üzerinde paylaşılan verinin boyutundaki hızlı artıştır. Basılı bilgi miktarı yaklaşık 14 yılda bir ikiye katlanırken, Web aracılığıyla erişilen bilgiler her yıl 2-3 kat artmaktadır (Tonta, Bitirim ve Sever, 2002:11). Web'in kısa ömürlü doğası ve hacminin hızla artması nedeniyle içerik kayıplarının önüne geçebilmek için onu yakalamak ve arşivlemek gerekmektedir. Birçok ülkede kamu politikalarının bir parçası olarak Web içeriğini yakalama ve arşivleme sorumluluğu bilgi merkezlerine aittir.

Dijital ortamda üretilen bilginin giderek değer kazandığı günümüzde, Web'deki bilgi miktarının boyutu kadar kalıcılığı da önemli hale gelmiştir. Dijital mirasa yönelik başarılı bir Web arşivleme çalışması gerçekleştirebilmek için elektronik bilginin yaşam döngüsündeki tüm aşamalar gözden geçirilmelidir. Web mirasının kalıcılığını sağlama adına bilgi kaynaklarının sağlanması, korunması, düzenlenmesi, arşivlenmesi, erişime sunulması, yönetimi ve denetimi bilgi merkezlerinin sistematik çalışmalarını gerekli kılar. Bu çalışmaların bir sonucu olarak bilgi merkezleri günümüzde yalnızca fiziksel yapılar değil aynı zamanda mobil ve dijital bilgi hizmetleri sunan, sanal mekânlar haline gelmiştir. Bilgi merkezlerinin yapısı ve hizmetleri kullanılan Web teknolojisine paralel şekilde gelişim göstermiştir. Üretilen bilgi miktarında görülen artış beraberinde erişim problemlerini de getirmektedir. Dijital ortamda üretilen ve dijitalleştirilen kültür varlıkları kültürel mirasın bir yansıması olarak Web üzerinde sunulmaktadır. Web mirasının korunması ise Web'in arşivlenmesini zorunlu kılar.

ÜÇÜNCÜ BÖLÜM

WEB ARŞİVLEME

3. WEB ARŞİVLEME

World Wide Web küresel ölçekte kullanılan ve bir miras unsuru olarak değerlendirilen benzersiz bir bilgi kaynağıdır (Pennock, 2013:2). Oldukça geniş bir yelpazede dijital doğan ve sonradan dijitalleştirilen kaynaklara ev sahipliği yapar (ISO, 2013:15). Bir miras unsuru olması nedeniyle tehdit altında oluşu onun kalıcı hale getirilmesini ve dolayısıyla arşivlenmesini gerekli kılmaktadır.

Web insanların iletişim kurduğu ve iş birliği yaptığı oldukça katılımcı ve yenilikçi bir alandır (ISO, 2013:15). Günümüzde cumhurbaşkanlığı seçimlerinden müzik festivallerine kadar her olay bir Web sitesinin oluşturulmasına, güncellenmesine veya kapatılmasına neden olmakta (Lasfargues ve diğerleri 2012:118) ve bloglar, forumlar, sosyal Web siteleri gibi yeni nesil iletişim ortamları olarak giderek yaygınlaşmaktadır. Web bu iletişim ortamlarına ev sahipliği yapan bir bilgi merkezi olmanın yanında kültürel yaşamın benzersiz bir sicili olarak öne çıkmaktadır.

Web içeriği doğası gereği kısa ömürlüdür ve sık güncellenir. Ancak Web'in gelişim hızı, dijital kültürel mirasa ve içerisindeki değerli içeriğe yönelik bir tehdit oluşturmaktadır. Bir araştırmada, geçen her bir yıl içinde Web sayfalarının %80'inin güncellendiği veya yok olduğu tespit edilmiştir (Ntoulas, Cho ve Olston, 2004:2). Yaşanan hızlı değişim ve kültürel miras kapsamındaki dijital kaynakların yok olma riski Web arşivlemeyi zorunlu hale getirmiştir.

Web arşivleme dijital süreklilik sağlar ve bir ulusun hafıza ve bilgisinin kaybolmasından kaçınmak için gerekli bir eylemdir (ISO, 2013:15). Web'in korunması ise somut ve somut olmayan kültürel miras unsurlarını barındıran kütüphane, arşiv ve müzelerin sorumluluğundadır. Bu bağlamda bir ulusun kültürel mirasının koruyucuları olarak, kütüphaneler, özellikle ulusal kütüphaneler, kültürel açıdan değerli materyallere gelecekte erişimi sağlamak için Web arşivlemenin önemini kabul etmişlerdir (Glanville, 2010:129).

Web arşivleme World Wide Web’de yayınlanan verinin bir araya getirilerek derlenmesi, depolanarak bir arşiv içerisinde korunması ve gelecek araştırmalar için kullanılabilir hale getirilmesi süreci olarak tanımlanabilir [15]. Bununla birlikte, Web arşivleme, World Wide Web’de kaydedilen verilerin, depolanmasını ve bir arşivde tutulmasını ve toplanan verilerin gelecek araştırmalar için kullanılabilmesini sağlayan bir veri toplama işlemidir (Niu, 2012b). Web Arşivleme ayrıca zaman içerisinde oluşan Internet kaynaklarının anlık görüntülerini seçme, yakalama, derleme, koruma ve erişim sağlama aktivitelerini ifade eder (International Organization for Standardization [ISO], 2012a: vi).

Dijital doğan ve dijitalleştirilmiş içeriği barındıran Web arşivleri kültürel miras kurumlarının yeni bir formudur (Costa, Gomes ve Silva, 2016:3). Web’i orijinal haline sadık kalarak arşivlemek, zengin doğası ve geçici yapısı nedeniyle özel bir çaba ve dikkat gerektirir. Dinamik içerik, değişkenlik, format çeşitliliği gibi zorluklara rağmen Web diğer dijital medyalar gibi kültürel mirasın bir parçası olarak hatırlanmayı hak etmektedir ve bu nedenle arşivlenmeli ve korunmalıdır.

Web arşivleme kendine has bir yazılım ve otomatikleştirilmiş süreçler aracılığıyla yapılır. Web sitelerinin toplanmasında “crawler” olarak ifade edilen tarayıcılardan yararlanılır ve toplama işlemi sırasında tarayıcılar Web sitelerini ziyaret ederek onları kopyalar ve kaydederler. Geleneksel arşivlemeyle benzer şekilde seçilen, depolanan ve kullanıma sunulan Web kaynakları belirli bir zamana ait anlık görüntüler olarak korunmaktadırlar.

Ulusal Arşivler Web Sürekliliği Girişimi’nden (Web Continuity Initiative at The National Archives) elde edilen bulgular, bir kuruluşun güncelliğini yitirdiğini veya önemsiz olduğunu düşündüğü eski içeriğe, büyük ve devam eden bir kullanıcı erişim talebi olduğunu göstermiştir (National Archives [NA], 2011:7). Bu içeriği arşivleme ve ona erişim sağlama daha geniş kapsamlı bilgi ve belge yönetimi aktivitelerinin parçası olmaktadır. Buna bağlı olarak Web arşivleme, bir kuruluşun bu bilgiyi etkin bir şekilde yönetme yeteneğine pozitif katkı sağlayabilir (NA, 2011:7).

Tarihsel derinliğine bakıldığında ilk Web arşivleme çalışmalarının 1996’da Internet Archive ve birkaç ulusal kütüphane tarafından başlatıldığı görülür. 2001’de başlayan Uluslararası Web Arşivleme Çalıştayı (IWA), deneyimleri paylaşmak ve fikir alışverişinde bulunmak için bir platform sağlamış ve 2003 yılında Uluslararası Internet Koruma Konsorsiyumu’nun (IIPC) kurulması, standartların geliştirilmesinde ve Web arşivlerinin oluşturulması adına iş birliğini büyük ölçüde kolaylaştırmıştır (Niu, 2012b).

Web arşivleme etkinliğinin amacı her biri görüntü, ses, video gibi gömülü kaynaklara sahip Web sitelerini mümkün olduğunca eksiksiz bir şekilde toplamaktır [16]. Arşivleme etkinliğinin odağında Web sayfaları yer aldığından Web sayfası ve Web sitesi terimleri iyi bir şekilde anlaşılmalıdır [15]. Bir Web sayfası, aynı anda oluşturulması beklenen bir veya daha fazla Web kaynağının kümesidir ve diğer kaynaklar kümeye yerleştirilen ögenin URI'si tarafından tanımlanabilir. Web sitesi ise ilişkili alan adlarını kapsayan bir alan adının içeriği veya bir ana makineden alınan bir içerik alt kümesi gibi teknik açıdan sınırlandırılmış entelektüel olarak ilişkili bir dizi kaynak olarak tanımlanabilir.

Web arşivleme tek seferlik bir eylem olmayıp birbirine entegre bir dizi faaliyet sürecini barındırır. Bu kapsamda seçim, sağlama, erişim, uzun süreli koruma gibi bilgi yaşam döngüsü aşamalarına ek olarak bilgi bütünlüğünün ve kalite güvencesinin sağlanması, koleksiyon yönetimi, küratöryal faaliyetler, kötü amaçlı yazılımların yönetimi, kaynak mükerrerliğinin en aza indirilmesi, sitelerin zamansal tutarlılığını sağlama gibi süreçler yer almaktadır. Web kaynaklarının korunması, diğer dijital kaynakların korunmasından farklı değildir.

Web sitelerini arşivleme nedenlerine bakıldığında bilgi nesnelerinin ömürlerinin kısa olması ve bir kültürel miras ögesi olarak orijinal bir değere sahip olması öne çıkar. Web sitelerinin yaşam sürelerinin ölçülmesine yönelik yapılan bir çalışmada bu süre ortalama 75 gün olarak ifade edilmiştir (Lawrence, Coetzee, Glover, Pennock, Flake, Nielsen, Krovetz, Kruger ve Giles, 2001:8). Bunun yanında bir sayfada bulunan haber içeriği veya sosyal medya Web siteleri gibi belirli içerikler çok daha hızlı gözden kaybolmaktadır. 404 durum koduyla ifade edilen sayfa bulunamadı hatası sık karşılaşılan durum kodlarından biridir ve bu durum kodu kütüphane kataloğunda bir kitabın bulunamamasıyla eşdeğerdir. Sayfa bulunamadı sonucunu sunmak yerine sayfanın arşivde yer alan güncel bir kopyasını sunma fikri ise erişimi güçlendirmektedir. Örnek vermek gerekirse, Birleşik Krallık Hükümet Web Arşivi (UK Government Web Archive) her ay 100 milyon civarında kullanımla çok yüksek bir kullanım oranına sahiptir (Pennock, 2013:3).

Web arşivleri metin, görüntü, film, ses ve diğer multimedya formatlarını içeren Internet kaynaklarının kopyalarını depolama yazılımıyla düzenli aralıklarla otomatik olarak toplamaktadır (ISO, 2012a:vi). Amaç özünde ilişkili kaynakları mümkün olduğu kadar kendi orijinal ortamında tekrar yayınlamaktır. Web arşivlemenin ana amacı muhtelif akademik, profesyonel ve özel amaçlar için sürekli olarak Web kayıtlarının korunması ve ilişkili kaynakların orijinaline en yakın biçimde tekrar yayınlanmasıdır (ISO, 2012a:vi).

Web kaynakları kültürel mirasın bir parçası olarak kabul edilir ve bu yüzden basılı miras yayınları gibi korunur (ISO, 2012a:vi). Web arşivlemeyi ihtiva eden birçok kurum uzun süredir bunu ulusal mirası koruma misyonu olarak görmektedir ve birçok ülkede bu derleme yasası gibi yasal mevzuatlar tarafından desteklenir (ISO, 2012a:vi). Yasal olarak Web içeriğini yakalama ve arşivleme yükümlülüğünün olması sonucunda oluşan yasal zorunluluk kurum faaliyetlere yön verme açısından çok daha güçlü ve yön verici bir amacı karşılar. Web içeriğini koruma noktasında kütüphaneleri destekleyen derleme yasalarının oluşturulması Web sitelerinin arşivlenmesini hızlandırmaktadır. Yasal yükümlülükler haricinde Internet'in evrimini ve içeriğini bütünüyle kaydetmeye yönelik toplumsal bir ilgi nedeniyle de Web arşivlenmektedir.

Bilimsel yayınların toplanması Web arşivlemenin amaçları arasında yer almaktadır. Bilimsel içeriğin Web arşivinde yer alması durumunda hem söz konusu içeriğin tanınırlığı artabilir, hemde içerik ile birlikte atıfların korunması sağlanabilir. 2009 yılında yürütülen bir araştırmada süreli yayınlarda yer alan atıfların %25'inin 5 yıl içinde ve %39'unun 10 yıl içinde erişilemez duruma geldiği belirtilmiştir. (Bhat, 2009:137). Bir Web arşivinde içerikle birlikte atıfların sürekliliğini sağlamak kaynakların gelecekte de kullanılabilmesi adına önemlidir.

Web arşivleme ayrıca kurumlara özel dijital kaynakları vurgulamak ve onların tanıtımını yapmak için bir yoldur. Bu nedenle, kendi akademisyenlerinin ve öğrencilerinin ürettiği yayınları arşivleyen kurumlar olan üniversiteler de Web arşivlemeyle ilgilenmektedirler. Diğer bir ifadeyle Web arşivleme, araştırma değeri içeren çevrimiçi kaynakları toplama adına değerli bir çaba ve zorlayıcı bir motivasyondur (ISO, 2012a:16).

Web siteleri zaman içerisinde gelişmekte ve güncellendikçe değişmektedir. Bir Web sitesinin nasıl ortaya çıktığı ve belirli noktalarda hangi bilgileri içerdiği eğer yasa nezdinde gerekli görülürse, Web sitesi bir belge olarak sunulabilir veya kanıt olarak kullanılabilir (NA, 2011:7). Telif hakkı anlaşmazlığında kanıt olarak ya da kişi, aile veya dijital soyağacı için araştırmaları için kullanılabilir (ISO, 2012a:16).

Web arşivleme uzun dönemde bilginin yeniden kullanımını ve insan ve kurumların ona erişmelerini sağlayan, bilgi alma ihtiyaçlarıyla uyumlu hayati bir süreçtir ve araştırma ve kültürel faaliyetlerle bütünleşmiştir (NA, 2011:6). Web teknolojilerinde yaşanan gelişimle birlikte araştırma topluluklarındaki eğilimler araştırmacıların tek tek Web sayfalarının entelektüel içeriğinin ötesine baktıklarını ve Web arşivlerini sorgu ve analiz için muazzam zamansal veri kümeleri olarak görmeye başladıklarını göstermektedir (Pennock, 2013:4). Araştırmacılar bu noktada gömülü bilgiyi açığa çıkarma amacıyla veri ya da bağlantı madenciliğini uygulamalarına ilgi gösterirler. Veri

madenciliği veya analizi, gelişiminin ilk aşamalarında çoğunlukla sosyal bilimciler tarafından keşfedilmiş olmasına rağmen, muhtemelen diğer akademik alanlara doğru genişleyecek ve birçok farklı disiplin için kullanışlı hale gelecektir (ISO, 2012a:16).

Web hızlı bir şekilde evrim geçirmektedir ve geçicidir. Bu nedenle Web arşivleme, kültürel miras materyalinin uzun dönem korunmasında kalıcı bir misyonla hareket eden kurumlar için asli ve kritik bir görev olarak görülmektedir (ISO, 2012a:16). Bu görevin yerine getirilmesinde kurumlar birbirinden farklı derleme stratejilerinden yararlanmaktadırlar.

3.1. Web Derleme Stratejileri

Web arşivlemede en sık kullanılan yaklaşımlar yığın derleme, alan adı derleme, seçimli derleme, tematik derleme, olay eksenli derleme ve son olarak karma derleme yaklaşımları olarak sıralanabilir. Seçilen derleme yaklaşımıyla toplanan içerik kurumun türüne ve yasal sınırlamalara bağlı olarak değişir.

Arşivleme kuruluşları tarafından stratejik önemi fark edilen derleme yaklaşımları, kişisel Web sayfalarından üst seviye alan adlarının tümünü yakalamaya kadar çeşitlilik göstermektedir (ISO, 2012a:vi). Fakat ölçek ve amacına bağlı olarak Web arşivleme stratejileri arasında iki ana yaklaşım arasında bir ayrım yapılabilir: yığın derleme ve seçimli derleme (ISO, 2012a:vi). Yığın derleme ulusal alan adlarının depolanmasında olduğu gibi büyük ölçeklidir ve Web sitelerine dair çok sayıda anlık görüntüyü yakalamak için tasarlanmıştır. Seçimli derleme ise belirli bir konu, olay, format veya içerik sahipleriyle yapılan bir anlaşma ölçütüne dayalı olarak gerçekleştirilir ve sıklıkla küçük ölçeklidir. Hangi yaklaşım seçilirse seçilsin bir Web sitesinin arşivlenme sıklığı önemine ve güncellenme sıklığına bağlıdır.

3.1.1. Yığın Derleme

Yığın derleme kavramı belirli bir zamanda ve geniş çapta bütün alan adlarına ait anlık görüntülerin yakalanmasını ifade eder. Oldukça otomatikleştirilmiş bir süreçtir ve sınırlar faaliyetin ölçeğine göre belirlenir. Uygulama sıklıkla yılda bir veya iki kez olmak üzere seyrek olarak gerçekleştirilmektedir (ISO, 2012a:9). Küresel ölçekte Web'i koruma amacıyla kurulan IA Wayback Machine yığın derlemeye örnek verilebilir.

Yığın derlemenin sık yapılması sık güncellenen Web sitelerinin arşivlenmesi noktasında bir dezavantaj oluşturur. Aynı Web sitesi henüz güncellenmeden tekrar tekrar yakalanabilir. Yığın derlemenin seyrek yapılması durumunda ise güncellenen Web sitesi yakalanamadığından bazı güncellemeler atlanabilir. Aynı durumlar alan adı derleme için de geçerlidir. Bu nedenle çok sayıda kurum alan adı veya yığın derlemeyi seçimli derlemeyle birlikte uygulamaktadır (ISO, 2012a:9). Bazı arşiv kurumlarında stratejik öneminin yüksek olduğu düşünülen Web siteleri seçimli derlemeyle sağlanır ve düşük öncelikli Web sitelerini yakalama ise yalnızca yığın veya alan adı derlemeye bırakılır. Eksik içeriğin kontrolü içinse Web sunucularından dönen HTTP durum kodları incelenebilir.

3.1.2. Alan Adı Derleme

Alan adı derleme sıklıkla ülkesel veya bölgesel üst seviye alan adlarının toplandığı bir Web derleme stratejisidir. Fakat isteğe bağlı olarak genel üst seviye alan adları da kapsama dâhil edilebilir.

Yığın derleme tüm alan adlarını toplamayı hedeflerken alan adı derleme genellikle ülkesel veya bölgesel alan adlarıyla sınırlı kaldığından görece daha küçük ölçeklidir. İçerik toplamada ülke veya bölge sınırlamaları yapılabilmesi nedeniyle bazı ulusal kütüphaneler tarafından tercih edilmektedir. Yapılan bir araştırmada ulusal kütüphanelerin %33 oranında ulusal üst seviye alan adı taraması yaptığı ancak %93 oranla büyük çoğunluğun seçimli ve tematik taramalara odaklandığı gözlemlenmiştir (Lasfargues ve diğerleri, 2012:119).

Alan adı derleme yaklaşımında yığın derlemede olduğu gibi Web siteleri periyodik olarak toplanmaktadır. İnsan gücünden asgari düzeyde yararlanılır ve materyal başına düşen toplama maliyeti seçimli ve tematik derlemeye göre azdır. Bunun yanında veri tabanlarının ve dinamik oluşturulan içeriğin derlenmesinde ve görüntü ve ses dosyalarının dizinlenmesinde çeşitli sorunlar yaşanmaktadır (Oğuz, 2006b:12).

3.1.3. Seçimli derleme

Seçimli derleme, Web sitelerinin önceden belirlenmiş ölçütlere göre manuel seçilmesi ve arşivlenmesidir (Murray ve Hsieh, 2008). Seçimli derleme yaklaşımı ile açıkça tanımlanmış seçim kriterleri ışığında genellikle kültürel veya araştırma değeri yüksek materyallerin toplanması hedeflenmektedir. Seçimli derleme sürecinde

gerçekleştirilen taramalarla kurumlar belirli alan adlarına, olaylara veya Web sitelerinin konu veya kategorilerine odaklanmaktadır (Lasfargues ve diğerleri, 2012:120).

Bugüne kadar Web arşivleme için birincil yaklaşımlar seçimli, yığın ve alan adı derleme olmuştur (Glanville, 2010:129). Seçimli derleme yığın ve alan adı derlemeye oranla daha küçük bir ölçekte gerçekleştirilir. Toplama işlemi daha sık tekrarlanır ve hedef Web sitelerine daha fazla odaklanılır. Seçimli derleme yaklaşımında kaynakların seçimi bir kütüphaneci veya küratör tarafından yapılır ve kaynak seçiminde konu, kaynakların erişilebilirliğini artırmaktadır.

ISO'ya göre (2013:9) seçim süreci konu, olay, format veya içerik sağlayıcılarla yapılan anlaşma gibi önceden belirlenen kriterlere göre uygun Web sitelerini tanımlayarak gerçekleştirilir. Seçimli bir Web arşivi, kullanıcı arayüzünde gelişmiş arama fonksiyonlarını oluşturmada kullanılabilecek, seçim esnasında veya depolama işlemi sonrasında sıklıkla küratörlerce eklenen daha fazla tanımlayıcı üstveriye sahip olma eğilimi göstermektedir.

3.1.4. Diğer Derleme Stratejileri

Diğer Web derleme stratejileri tematik, olay eksenli ve karma derleme olarak sıralanabilir. Tematik derlemede önceden belirlenen URL adreslerinin arşivlenmesi ile çekirdek bir koleksiyon elde edilir. Web sitelerini toplama işlemi periyodik olarak gerçekleştirilmektedir. Olay eksenli derleme ise büyük olasılıkla popüler durumda olan bir etkinlik ile ilgili bir takım Web sitelerini toplamak için uygulanır. Olimpiyatlar veya seçimlerle ilgili Web sitelerinin toplanması olay eksenli derlemeye örnektir. Karma derleme yaklaşımında ise birden fazla yaklaşım bir arada kullanılmaktadır. Bu yaklaşım mevcut derleme yaklaşımlarının yetersiz kaldığı düşüncesiyle geliştirilmiştir (Oğuz, 2006b:35).

Derleme yaklaşımları içerisinde ayrıca ulusal derleme yasalarıyla desteklenen ve Web sitesi sahiplerinin içeriği arşive gönderdiği geleneksel derleme yaklaşımına da rastlanır. Bu yaklaşımda milli kütüphanelerin görev ve sorumluluğu fazladır. Bu yaklaşımın uygulayıcıları olarak Almanya ve Fransa Milli Kütüphaneleri örnek verilebilir.

3.1.5. Derleme Stratejilerinin Ölçeklendirilmesi

Web derleme stratejileri ölçeklendirilirken temelde iki farklı yaklaşımdan söz edilebilir. İlk yaklaşım büyük ölçekli ve derleme sıklığının güncelleme sıklığına göre yapılandırıldığı makro arşivlemedir. Milli kütüphanelerin Web derleme faaliyetleri makro arşivleme olarak karakterize edilebilir. Makro arşivleme yığın derleme, alan adı derleme ve seçimli derleme ile bunların bir kombinasyonunu içerir (NA, 2011:11). Yığın derlemede hedef geniş bir yelpazede tüm Web sitelerinin toplamakken, alan adı derlemede hedef belirli bir ülkenin veya bölgenin alan adındaki Web sitelerini toplamaktır. Seçimli derlemede ise belirli bir konudaki Web siteleri toplanır. Mikro arşivlemede ise belirli bir amaç için tek bir sitenin anlık görüntüsü periyodik olarak sağlanır (Brügger, 2005:11).

Her iki yaklaşımda da bir sitenin anında ve tamamen derlenemeyeceğini unutmamak önemlidir (Niu, 2012b). Web derlemenin amacı, Web sitesini tüm işlevleriyle birlikte yeniden inşa etmek değil, mümkün olduğunca kullanıcılar tarafından görüntülenen şekliyle Web sayfalarını yakalamaktır.

Web arşivleme yaparken hangi teknik yaklaşımın sergileneceğinin belirlenmesinde ölçeğe ek olarak seçilecek arşivleme yaklaşımı da önemli bir belirleyicidir. Büyük ölçekli Web arşivleme için üç yaklaşımdan söz edilebilir (Pennock, 2013:6):

- İstemci merkezli arşivleme (Client-side archiving)
- İşlem merkezli arşivleme (Transactional archiving)
- Sunucu merkezli arşivleme (Server-side archiving)

İstemci merkezli arşivleme görece basit, ölçeklenebilir ve uygun maliyetli olması nedeniyle en sık kullanılan yaklaşımdır ve bu yaklaşımda Heritrix veya HTTrack gibi arşiv tarayıcıları istemciler (örneğin Web tarayıcıları) gibi davranmakta ve sunucudan teslim edilen içerik yanıtlarını toplamak için HTTP protokolünü kullanmaktadır (Pennock, 2013:7). Böylelikle tarayıcı bir kaynak listedeki URL'leri izleyerek kaynaklara ilişkin tüm bağlantıları belirli bir derinlikte tarar ve mevcut dosyaların kopyalarını yakalar. İstemci merkezli arşivlemede yakalama işleminin başarısı materyalin arşiv tarayıcı için erişilebilir olmasına bağlıdır (NA, 2011:15). Bu yaklaşımda medya akışının ve gizlenmiş içeriğin yakalanması zordur.

İstemci merkezli arşivleme ile tarayıcı teknolojisini kullanarak Web içeriğini koruma fikri Web arşivleme araçlarının geliştirilmeye başlandığı Kraliyet Kütüphanesi'nin Kulturw3 Girişimi ile ilk olarak 1996 yılında İsveç'te ortaya çıkmıştır (Hakala, 2014:178). Web arşivleme yaygınlaştıkça birçok kültürel miras kurumu arasında bu yaklaşım yaygınlaşmıştır.

İstemci merkezli arşivlemede dikkat edilmesi gereken bir konu seçilen zamanlama algoritmasıdır. Seçilen zamanlama algoritmasında eğer tarayıcıdan gereğinden fazla sorgu yapması istenirse hedef sunucu işlevsiz hale gelebilir. Derleme işleminin alacağı zaman, her bir istek için zaman aralığı belirlendikten sonra dosya sayısı ve zaman aralığının çarpımıyla hesaplanabilir (Hakala, 2014:178).

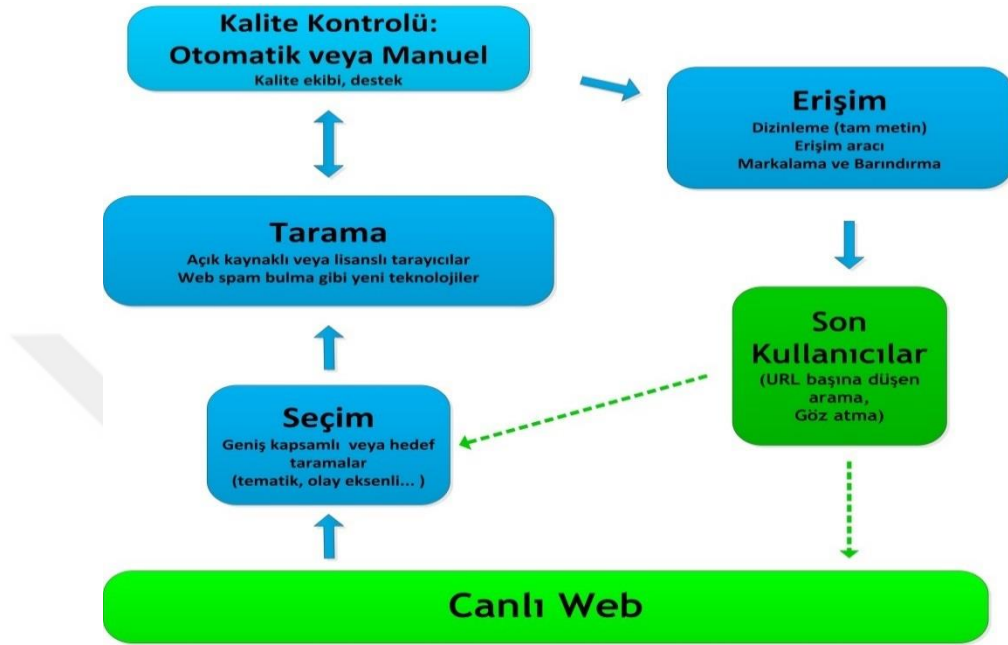
İşlem merkezli arşivleme Web içeriğini barındıran Web sunucusuna erişim gerektirir ve bir yöntem olarak daha az sıklıkta kullanılır (NA, 2011:15). Bir sitenin kullanıcıları ile sunucu arasındaki işlemlerin kaydedildiği bu yaklaşımda, görüntülenmeyen içerik arşivlenmez. İşlem merkezli arşivleme, yasal hesap verilebilirliğin önemsendiği, kullanıcının eylemlerinin ve veri girişlerinin olduğu durumlarda, örneğin finansal veya ticari işlemlerin arşivlendiği özel durumlarda kullanılır (Pennock, 2013:7).

Sunucu merkezli arşivlemede ise dosyalar HTTP protokolüne başvurmadan doğrudan sunucudan kopyalanır (Pennock, 2013:7). Sunucu sahipleriyle aktif iş birliği gerektiren bu yaklaşımdan genellikle HTTP protokolü üzerinden içeriğin toplanamadığı durumlarda yararlanılır. Tarayıcıların gözden kaçırdığı harita veya veri tabanı içeriğini yakalamada bu yöntem kullanışlı olabilir. Ayrıca bu yaklaşımla istemci merkezli Web arşivlemede tarayıcıların erişemediği site bölümlerini arşivlemek mümkündür (NA, 2011:15).

Web arşivleme projelerine uyum sağlamak için toplama kurumları kendi stratejilerini ve becerilerini geliştirmeli ve mevcut kaynaklarıyla kendi gereksinimlerini en iyi ve etkili biçimde karşılayabilecek yöntemi seçmelidirler.

3.2. Web Arşivleme Yaşam Döngüsü Aşamaları

Web arşivlemeye temel olan bazı aşamaları yansıtmaya amacıyla Şekil 1’de genel bir Web arşivleme üretim şemasına yer verilmiştir.



Şekil 1. Web Arşivleme Üretim Şeması (Lasfargues ve diğerleri, 2012:120)

Web arşivleme yaşam döngüsünde değerlendirme ve seçim, sağlama, organizasyon ve depolama, tanımlama ve üstveri, erişim ve kullanım ve son olarak koruma aşamaları yer almaktadır. Uygulamada bazı değişiklikler görülse de bu aşamalar Web arşivlemenin özü olarak kabul edilebilir. Tüm adımların tamamlanmasının ardından yakalanan içeriğin genel kalitesine ve kurumun arşivleme politikasına bağlı olarak bazı aşamalar yeniden gözden geçirilebilir.

3.2.1. Değerlendirme ve Seçim

Değerlendirme terimi arşiv topluluğunda, kayıtların kıymetini değerlendirme ve kayıtların muhafaza edilip edilmeyeceği ve ne süreyle korunması gerektiğine karar verme sürecine karşılık gelir (Niu, 2012b). Değerlendirme aşamasında tüm Web arşivleri, bir veya daha fazla kriter temelinde koruma amacıyla Web kaynaklarını seçer. Bu nedenle değerlendirme ve seçim aşamaları benzerlik gösterir.

Seim ařamasında belirlenen seim kriterleri genellikle koleksiyonu oluřturan kuruluřun temel misyonu ile uyumludur ve iřlemler mevzuat ve kuruluřun koleksiyon geliřtirme stratejisi tarafından ngrlen bir biimde gerekleřtirilir (ISO, 2013:9). Operasyonel hususlar, personel alıřtırma, kaynak ve uzmanlık gibi sınırlılıklar ise sıklıkla uygulamayı etkilemektedir (ISO, 2013:9).

Niu'ya gre (2012b) mevcut Web arřivlerinde neyin korunacađını belirlemek iin alan adı, konu, olay, medyanın veya Web sitesinin tr veya kategorisi, kaynađın telif durumu ve deđeri ile kuruluřun finansal gc gibi eřitli seim kriterleri uygulanmaktadır. Teorik olarak, nesnel kriterlere dayanan seim kolaylıkla otomatikleřtirilebilir. Teknik dzeyde medya trlerinin ve alan adlarının belirlenmesi ile dergi ve blog gibi Web sitesi trlerinin ayırt edilmesi yazım aracılıđıyla mmkndr. Ayrıca Web ieriđinin kalitesi ve poplerliđi gelen bađlantıların ve ziyaretilerin sayısına, evrimii videolar iin izleyicilerin sayısına ve kullanıcı derecelendirmelerine dayanarak tanımlanabilir.

Seim kriterinde hangi Web derleme ynteminin tercih edileceđi Web arřivinin kapsamını belirlenmede nemlidir. Kurumlar arasında bir seim kriteri olarak alan adı arřivleme olduka yaygındır. Birok Avrupa lkesi Web'i kendi lke alan adlarında arřivlemektedir. Fakat Internet gibi sınırları cođrafi ve fiziksel yapının tesine dađılmış olan kresel bir sistemde alan adları, ulusal ieriđi kesin bir biimde tespit ve tanımlama iin yeterli olmayabilir (ISO, 2013:9). lke alan adlarının tesine geerek tm Web'i arřivlemeye alıřan bir kuruluř olan IA ise yalnızca yzeyde yer alan Web sayfalarını yakalayabilmektedir (Lecher, 2006:2013).

Birer seim kriteri olarak konuya ve olaya dayalı seimde insan muhakemesine ihtiya duyulur. Bilgi profesyonellerinin Web kaynaklarını manuel seimi zaman alıcı ve pahalıdır ve bu nedenle yalnızca kk aplı Web arřivlerinde kullanılmaktadır (Niu, 2012b). Bu yntemin alternatifi kullanıcı nerilerini toplamak ve konu uzmanları ile birlikte alıřarak bir URL sicili oluřturmaktır. Buna Birleřik Krallık Hkmet Web Arřivi'nin oluřturulmasında URL sicillerinden yararlanılması rnek verilebilir.

Ktphanecilerin veya arřivcilerin, Web sitesinin arřivlenip arřivlenmeyeceđine karar vermek iin tek tek Web sayfaları yerine btn bir Web sitesine ynelik deđerlendirme yapmaları Web sitelerinin deđerlendirme ve seim srecini daha verimli hale getirebilir.

Web arřivlerine kaynak seiminde deđer bazlı seim kriterleri de uygulanabilir. Bu ynteme tarihsel, kltrel, sosyal, eđitsel veya akademik bakıř aılarından deđerli bulunan Web kaynaklarının toplanması veya spam filtrelerinin uygulanması rnek

verilebilir. Değer bazlı seçim kriterinin uygulandığı Web arşivlerine, Ulusal Tayvan Üniversite Kütüphanesi'nin oluşturduğu Web arşivi örnek gösterilebilir.

Bazı kaynak türlerinin dâhil veya hariç tutulmasına ilişkin her zaman bir kesinlikten bahsedilemez (ISO, 2013:9). Kimi kurumlar sosyal ağları, blogları ve benzeri interaktif platformları arşivlemeye karar verirken diğerleri bunları kapsam dışında bırakabilir. Bu, çevrimiçi reklamcılık, pornografi ve bilgisayar virüslerini içeren veya onlardan etkilenmiş dosyalar gibi kaynaklara aynı ölçüde uygulanacak geçerli temel bir politika kararıdır (ISO, 2013:9). Bu gibi durumlarda kaynaklar bir örneklem içerisine dâhil edilerek toplanır ve gelecekte haklarında karar verilmek üzere arşivlenebilir.

3.2.2. Sağlama

Web arşivinin ölçeğine, Web arşivi ile Web sitesi sahipleri arasındaki ilişkiye ve arşivlenmiş Web içeriğinin doğasına bağlı olarak, Web içeriğini edinmek üzere farklı sağlama yöntemleri kullanılabilir (Niu, 2012b). Bu yöntemler otomatik tarama, manuel tarama ve yasal derleme ile kurumlardan gelen Web materyallerini kabul etme olarak sıralanabilir.

Web içeriğinin sağlanmasında en sık kullanılan yöntem taramadır. Tarama işlemi sırasında Web sitelerinin kopyalarını almak ve onları toplamak için arşiv tarayıcılardan yararlanılır. Arşiv tarayıcıları ve arama motorlarında kullanılan Web tarayıcılarıyla oldukça benzer yapıdadır. Arşiv tarayıcılarına açık kaynaklı olanlara HTTrack, Heritrix, WGet ve Hanzo Arşivleri tarafından tasarlanan Miyamoto tarayıcısı örnek olarak verilebilir.

Tarama işleminde genellikle kaynak liste olarak da ifade edilen ve başlangıç noktası olarak da düşünülebilecek Web sitelerinin ana sayfalarından yararlanılır. Tarayıcı kaynak listede yer alan bağlantıları izler ve Web içeriğini indirir. Sağlama aşamasında daha önce alınan seçim kararları birer filtre olarak uygulanır ve bir kaynak listesini derlemek ve tarayıcı parametrelerini yapılandırmak için temel oluşturur (Niu, 2012b). Tarama sürecinde yalnızca önceden belirlenen filtreleri geçen bağlantılar arşivlenir. Tarayıcılar mimarilerin heterojenliğine (dinamik, sosyal Web vb.), dillere ve formatlara (video, metin, ses dosyaları vb.) uyum sağlamalı ve bağlantıları yorumlayabilmelidir (Lasfargues ve diğerleri, 2012:120).

Tarayıcılar için içeriğin keşfedilmesinde bağlantılar anahtar roldedir. Bağlantıların yapısı ve çıkarılması tarama sürecinde sorun yaşanıp yaşanmayacağını belirler. Bağlantıların çıkarılması sırasında iki yöntem yaygın olarak kullanılır. Bunlardan ilki ayrıştırma'dır. Bu yöntem, içeriği ayrıştıran, bağlantıyı çıkaran ve keşfedilen içeriği anında arşivleyen tarayıcıların oluşturulmasını kapsar (Lasfargues ve diğerleri, 2012:120). Ayrıştırma yöntemi tarayıcı teknolojisinde sık kullanılmakta ve kurum içi çözümlere uygulanabilmektedir. IA tarafından geliştirilen Heritrix [16] ve Xavier Roche tarafından geliştirilen Htrack [17] ayrıştırıcı tarayıcılara örnek verilebilir.

Bağlantıların çıkarılması sırasında kullanılan diğer yöntem yürütmedir. Ayrıştırılması mümkün olmayan siteleri arşivlemek için kullanılan yürütme yöntemi, içeriğin yakalanarak çalıştırıldığı bir Web tarayıcının taklit edilmesine dayanır (Lasfargues ve diğerleri, 2012:120). Örneğin, Hanzo Archives tarafından üretilen yürütme tabanlı tarayıcı geniş bir yelpazedeki ayrıştırılması mümkün olmayan siteleri arşivlemek için kullanılmaktadır [18]. Günümüzde Internet Bellek Vakfı (IMF) gibi bazı kurumlarda ayrıştırıcı ve yürütücü tarayıcı yaklaşımlarına birlikte yer veren yaklaşımlara rastlanmaktadır (Lasfargues ve diğerleri, 2012:121). Ayrıştırıcı ve yürütücü tarayıcı yaklaşımların birlikte kullanımı her ikisinde de en iyi şekilde yararlanmayı mümkün kılar.

Tarama işleminde bir diğer önemli konu ölçeklenebilirliktir. Bir tarayıcının kapasitesi ve bir sayfayı daha önce ziyaret edip etmediğini tespit edebilme yeteneği ölçeklenebilirlik açısından önemlidir. Mevcut birçok tarayıcı ulusal alan adlarının boyutuna göre ölçeklendirilmemiştir (Lasfargues ve diğerleri, 2012:121). Ayrıca Internetin hızla büyüyen hacmi, Web tarayıcılarının tasarımı ve uygulanması için ciddi bir zorluk meydana getirmektedir. Sağlama aşamasında çözülmesi gereken sorunlar, SPAM ve tuzak algılama ile dinamik HTML ve gizli Web analizi olarak sıralanabilir (Lasfargues ve diğerleri, 2012:121).

Birçok bilgi kurumunda Web yayınlarının edinilmesinde tarama işleminden yararlanılır. Fakat dinamik Web içeriğinin veya akış medyasının derlenemediği bazı durumlarda Web kaynaklarının tarayıcı kısıtlamaları nedeniyle manuel olarak yakalanması da gerekebilir. (Niu, 2012b). Web içeriğinin sağlanmasında ayrıca kütüphaneler ve arşivler resmî kurumlardan, bağışçılardan ve yayıncılardan yasal derleme ile transferleri kabul edebilir.

3.2.3. Organizasyon ve Depolama

Mevcut Web arşivleri, arşivlenmiş Web içeriğini organize etmek ve depolamak için üç yaklaşım kullanmaktadır: yerel dosya sistemleri, Web tabanlı arşivler ve Web tabanlı olmayan arşivler (Masanés, 2009:29). Yaklaşımların tamamı Web sayfalarının entelektüel içeriğini korumakta fakat özgünlüğünü ve bütünlüğünü koruma düzeyleri farklılık göstermektedir.

Yerel dosya sistemi biçiminde oluşturulan Web arşivlerinde tarayıcı Web'dekine benzer şekilde dosya sisteminde gezinir. Bu yaklaşımla Web sitelerinin hiyerarşik yapısı ve bağlantı ilişkileri korunur. Web kaynaklarının yerel dosya sistemlerine uyumlu hale getirilmesi için URI'lerin yeniden adlandırılması ve bağlantıların dosya sistemi içinde gezinmeye elverişli hale getirilmesi gerekebilir.

Web tabanlı arşivlerde bağlantıların canlı Web yerine arşive işaret etmesi için yönlendirilmeleri veya dönüştürülmeleri gerekmektedir. Özgünlüğü ve bütünlüğü en yüksek seviyede koruyan Web tabanlı arşivlerde Web sayfaları ve ilişkili meta veriler gruplandırılır ve konteyner dosyalarında saklanır ve orijinal URI'ler ve bağlantılar korunur (Niu, 2012b).

Web tabanlı olmayan arşiv yaklaşımı, hipermetin bağlamından Web belgelerini çıkarır ve bunları katalog tabanlı erişim moduna göre yeniden düzenler veya onları PDF dosyalarına dönüştürür (Niu, 2012b). Bu yaklaşımda özgünlüğün ve bütünlüğün korunma seviyesi düşüktür.

Web arşivlemede disk alanı kullanımı ise genellikle üç kategoride düşünülebilir: Yakalama faaliyetlerini kontrol eden bir veri tabanı, dokümanların işlendiği ve arşivleme için hazırlandığı bir çalışma alanı (örneğin üstveri çıkartma) ve arşivlenen dokümanların saklanması (Hakala, 2014:179). Bununla birlikte, bir Web taraması sırasında Web derleminin büyüklüğünü kesin olarak bilmenin bir yolu yoktur. Depolama alanı ihtiyacını etkileyen temel faktör yakalayıcının değişen mi (bir yakalama veya çoklu yakalama sırasında her dosyanın yalnızca bir kez depolandığı) yoksa değişmeyen mi (her yakalamada her dosya depolanır) yapıda olduğudur (Hakala, 2014:180). Örnek vermek gerekirse, Fin Web Arşivi'nde kullanılan NEDLIB yakalayıcısının değişen yapıda olduğu ve eğer değişen yapıda olmasaydı oluşturulan arşivin iki ila üç kat daha büyük olacağı aktarılmıştır (Hakala, 2014:180). Bu noktada dosyaların birden çok kopyasını tutmak savurganlığa neden olacağından değişen yapıda bir yakalayıcı kullanmak önerilebilir.

Depolama alanı ihtiyacını belirlemenin bir diğer yolu ise arşivdeki mükerrrer kaynakların belirlenmesidir. Bu noktada arşivleme modülünün bir kaynağa ait tüm kopyaları bulabilmek için dosyaların sağlama toplamalarını hesaplayabilmesi beklenir. Sağlama toplamalarından yararlanılmadığı takdirde depolanan kaynakların manuel olarak kontrolü gerçekleştirilebilir fakat bu işlem masraflı ve zahmetli olabilir.

3.2.4. Tanımlama ve Üstveri

Web arşivleri çok katmanlı ve hiyerarşik yapıdadır. Toplanan Web siteleri Web sayfalarını, Web sayfaları da metin, görüntü, ses ve video dosyalarını içerir. Üstveriler de bu sırayı takip ederek öncelikle Web siteleri ardından Web sayfaları ve son olarak dosyalar için oluşturulur. Ayrıca dosya seviyesinde nesneler için nadiren de olsa dosya formatı, dosya boyutu ve değiştirilme tarihi gibi üstveriler oluşturulur (Niu, 2012b).

Üstveri oluşturma yaklaşımı ve üretilen üstverilerin zenginliği, Web arşivinin ölçeğine ve arşivleme organizasyonunda bulunan kaynaklara bağlı olarak değişir (Niu, 2012b). Üstveriler, büyük ölçekli Web arşivlerinde genellikle otomatik olarak oluşturulurken küçük ölçekli Web arşivlerinde manuel olarak oluşturulabilir. Ayrıca kullanıcılar tarafından yapılan etiketleme ile birlikte yorum yapma veya derecelendirme yoluyla da üstveri oluşturulabilir.

Web arşivlerinde üstveri kullanımı korumayı ve erişimi destekleyici niteliktedir. Kaynakların içinde gömülü halde bulunan, arşivleme sürecinde otomatik olarak oluşan veya sorumlu görevli tarafından elle oluşturulan birçok farklı üstveri türü bulunmaktadır (ISO, 2013:13). Metadata Kodlama ve Aktarım Standardı (METS) tarafından üstverinin beş farklı türünü tanımlanmıştır (ISO, 2013:13-14):

- **Tanımlayıcı Üstveri:** Web arşivlerini kataloglayan ya da manuel olarak üstveri ekleyen kurumlar genellikle daha çok tanımlayıcı üstveriye sahiptirler. Tanımlayıcı üstveriler içerik bilgisinin özelliklerini yansıtmakla birlikte içeriği keşfetme ve tanımlama gayesiyle kullanılmaktadır.
- **Yapısal Üstveri:** Yapısal üstveri içeriğin fiziki yapısına yönelik özellikleri aktarır. Web kaynakları sıklıkla yapılandırılmış ve birbirine bağlanmış öğelere yoğunlaşır. Yapısal üstveri aracılığıyla yapısal ilişkiler açıkça ifade edilebilir ve METS gibi üstveri şemalarına kaydedilebilir. Bu tür üstveri, arşivin dolaşıma uygunluğunu sürdürebilmek için bağlantılara gereksinim duyulan veri göçü uygulamalarında kullanılabilir.

- **Kaynak Üstveri:** Kaynak üstveri bir kaynağın neden ve nasıl oluşturulduğunu ve yaşam döngüsü boyunca başından nelerin geçtiğini aktarır. Kaynak üstveri bilgi yaşam döngüsü boyunca toplanır ve içerikte yaşanan değişimi yansıtır. Kaynak üstveri düşük seviyelerde yapılandırma dosyaları, tarama raporları, log dosyaları gibi bir arşiv tarayıcısının kayıt aktivitelerini içeren dosyaları ve URL, tarama tarihi ve sunucunun IP adresini içeren Web sunucusu ile tarayıcı arasındaki etkileşimi aktaran enformasyonu kapsar.
- **Teknik Üstveri:** Teknik üstveri dijital bir nesneye erişimin nasıl sağlanacağını, nesne üzerinde değişikliklerin nasıl yapılacağını ve onun nasıl korunacağını ayrıntılarıyla belirterek nesnenin teknik özelliklerini tanımlar. Dijital nesneye ait yazılım ve donanım bilgisi veya özgünlüğü sağlamak için oluşturulan sağlama toplamları ve dijital imzalar teknik üstveri örneğidir.
- **Hak Üstverisi:** Hak üstverisi kaynaklara yasal olarak verilen kullanım hakkını ve kaynak mülkiyetini tanımlar. Yetkisiz erişimleri önleme adına haklara ilişkin oluşturulan üstverinin kaynakla birlikte korunması gerekir.

3.2.5. Erişim ve Kullanım

Arşivciler genellikle verilerin sağlanması depolanması, yönetilmesi ve korunması üzerine yoğunlaşmaktadır. Bununla birlikte veriler analize ve kullanıcı erişimine de açık olmalıdır. Bir Web sitesi arşivlendiğinde kalıcı bir URL ye dönüşür ve çevrimiçi olarak erişilebilir kalmaya devam eder (National Archives [NA], 2017:5). Erişim ve kullanım amacıyla Web arşivlerinde keşif ve arama izni veren kullanıcı arayüzlerinden yararlanılır. Arayüz üzerinde sunulan erişim yöntemler ise URL arama, anahtar sözcük arama, tam metin arama alfabetik arama, gelişmiş arama ve konu araması olarak sıralanabilir.

Bir Web arşivine erişimde en yaygın kullanılan yöntem URL aramadır. Yakalanan URL yakalama tarihiyle birleştirilir ve bu aynı kaynağın farklı versiyonlarını ayırt etmede kullanılır. Toplanan tüm verilerin dizinlenmesinin zorluğu URL arama özelliğinin sık kullanılan bir arama yöntemi olmasında etkili olmuştur. Avrupa'daki Web arşivleri üzerine yapılan bir ankette, Web arşivlerinin %68'inin URL aramayı desteklediği görülmüştür (Costa ve diğerleri, 2016:4). Fakat kullanıcıların URL'leri hatırlamasını zorunlu kıldığı için URL arama sınırlıdır ve URL arama güncel olmayan içeriği de getirebilir. Şekil 2'de Wayback Machine aracılığıyla Milli Kütüphane ana sayfasına yönelik yapılan bir URL araması göstermektedir:

Arama yöntemlerinden bir diğeri olan alfabetik arama içeriğın A-Z harf diziliminde sıralandığı erişim metodudur. Gelişmiş aramada ise birden fazla arama bileşeninden yararlanılır ve daha spesifik sonuçların alınması muhtemeldir.

Konu arama özellikle seçimli, tematik ve olay eksenli arşiv koleksiyonlarına erişimde yaygın olarak kullanılmaktadır. Küratöryal çaba gerektirir. Bununla birlikte geleneksel bibliyografik yönetim metotları tıpkı basılı kitapların ve dergi makalelerinin kataloglanmasında olduğu gibi konu araması yapılabilen Web arşivlerine uygulanabilir (ISO, 2013:10). Web arşiv kayıtlarını kütüphane kataloğuna aktarmak kaynakların görünürlüğünü arttıracaktır. Fakat bu yaklaşım kaynak yoğunudur ve kataloglanacak nesnelerin miktarının çok olmasının yanı sıra kaynak tanımlamadaki zorluklar nedeniyle yaklaşımın Web arşivlemede uygulanması zordur (ISO, 2013:10). Web kaynaklarının kataloglanması tüm kaynaklar yerine özel koleksiyonlar seviyesinde sağlanabilir.

Web arşivleme yoluyla yakalanacak bilgi makinece ulaşılabilir olmalıdır. Başka bir ifadeyle tarayıcı Web sitesine erişebilmelidir. Giriş ekranı, onay ekranı, seçim listesi veya arama ekranı aracılığıyla sunulan bilgiye erişim, makinece ulaşılabilir değildir ve bu nedenle bir Web tarayıcı tarafından yakalanamaz (NA, 2011:9). A-Z listesi veya site haritası gibi doğrudan indirme yapılabilecek alternatiflerin yaratılması sitenin arşivlenebilmesi adına faydalı olacaktır.

Erişim sürecinde arayüz tasarımı ve kullanılan erişim metotları kadar kaynakların telif hakkı durumu da önemlidir. Birçoğuna ücretsiz erişim olmasına rağmen Internet kaynakları genel olarak telif hakkı ile korunur. Yasal mevzuata ve izinlere bağlı olarak toplama kurumları bir dizi erişim stratejisini hayata geçirmektedirler (ISO, 2013:11):

- *Karanlık arşiv: Koleksiyonlara hiç kimse erişemez. (Bir görevli veya küratöryal amaçlar dışında)*
- *Gri arşiv: Koleksiyonlar yalnızca kendilerine yetki verilen son kullanıcılarca görülebilir (örneğin: araştırmacılar) ve/veya erişim bina bünyesinde sınırlandırılmıştır (örneğin: kütüphane okuma odası).*
- *Çevrimiçi arşiv: Erişim tüm kullanıcılar içindir ve kaynaklar genellikle toplayıcı kurumunun Web sitesinden erişilebilir durumdadır.*
- *Alternatif olarak bir arşiv, bölümlerini yukarıdaki stratejilere uyumlu hale getirerek karma bir model kullanabilir. İzinlerin sadece varsayım ve ima düzeyinde olduğu, açıkça izin verilmediği durumlarda devre dışı bırakma politikası temelinde depolanan kaynaklara erişim hak sahiplerinden izinlerin alınması yoluyla sağlanabilir.*

Web arşivlerinin erişim ayrıca üstverilerin zenginliğine ve kullanılan arama ve dizinleme araçlarına bağlıdır. Uygun bir tarayıcı ile oluşturulmuş bir Web arşivi son kullanıcılara kendiliğinden erişilebilir olmadığından, arşivdeki dokümanlar, tercihen Web içeriğini dizinleme için optimize edilmiş tam metin arama motoruyla birlikte dizine eklenmelidir (Hakala, 2014:181).

Web arşivleri büyüdükçe sunduğu arama yetenekleri de büyümelidir. Web sitelerinin oluşturulma aşamasında Web arşivinin kullanıcı dostu olarak tasarlanması erişimi arttırmaktadır. Bu amaçla Web arşivlerinde görselleştirmeyi öne çıkaran daha güçlü ve etkili araçlar üretilmiştir. Örneğin Birleşik Krallık Web Arşivi, zaman içinde Web arşivinde arama terimlerinin görülmesini göstermek için arama sonuçlarını bir grafik içinde döndüren kullanıcılar için bir NGram görüntüleyici sağlamaktadır. İsmi “N” harfi ve “Gram” kelimesinin birleşiminden alan NGram, verilen bir dizilimdeki tekrar oranını bulmaya yarayan yöntemdir [157]. Buradaki “N”, tekrarın kontrol edildiği değerdir. “Gram” ise bu tekrarın dizilim içerisindeki ağırlığını ifade etmek için kullanılmıştır. Bazı Web arşivlerinde ayrıca etiket bulutları, konu hiyerarşisi görselleştirmesi ve posta kodu tabanlı erişim de dâhil olmak üzere farklı araçlar da modellenmiştir (Hockx-Yu, Crawford, Coram ve Johnson, 2010:5).

Web arşivleme daha popüler hale geldikçe erişimi geliştirme adına kullanıcı gereksinimleri daha fazla dikkate alınmaya başlamıştır. Uygulama örneklerine bakıldığında birçok kurumun Web arşivine çevrimiçi erişim sağlamak için Internet Archive tarafından geliştirilen açık kaynak kodlu Wayback Machine’in kullandığı görülür. Arşivlenen URL, gezintinin gerçekte canlı Web’de değil de arşivde yapılmasını sağlamak için Wayback Machine tarafından anında yeniden yazılmaktadır (Lasfargues ve diğerleri, 2012:122). Internet Bellek Vakfı (IMF), dört yıl önce sunucu tarafı bağlantısını yeniden yazmayı ilk uygulayan olmuştur ve o günden beri erişimini sürekli geliştirmektedir (Lasfargues ve diğerleri, 2012:122).

Arşiv kaynaklarının yeniden gösteriminde sıklıkla “canlı Web kaçağı” olarak ifade edilen problem, arşivlenen kaynağın içerisindeki linklerin Web arşivi içerisindeki arşivlenen sürümü işaret etmek yerine canlı site üzerindeki güncel kopyayı çözümlemesinde ortaya çıkar (ISO, 2013:11). Bu genellikle URL’lerin yeniden yazımındaki yanlışlıktan kaynaklanır ve sıklıkla JavaScript’teki gömülü linklerin erişim yazılımı tarafından algılanmamasının bir sonucudur (ISO, 2013:11).

Web arşivlerinde erişilebilirlik arşivin bulunduğu ülkedeki yasal koşullara da bağlıdır. Bazı ülkelerde mevzuat ulusal kütüphaneye Web sitelerini koruma ve arşivlenmiş kopyaları erişime sunma izni verir. Yeni Zelanda’nın yasal derleme mevzuatı bunu örnekler. ABD örneğinde ise Kongre Kütüphanesi tarafından arşivlenmiş tüm Web

sitelerinin kaynakça kayıtları halka açık hale getirilmiştir ve yalnızca üreticileri tarafından izin verilen Web sayfalarına kamu tarafından erişim sağlanabilir (Grotke ve Jones, 2010:17). Web sitesi sahiplerinin izin vermediği durumlarda ise kamu tarafından Web kaynaklarına erişim sağlanamaz. Kamu tarafından erişilebilen bazı Web arşivleri, Web sitesi sahipleri ile rekabeti önlemek için daha az işlevsellik ve geciktirilmiş erişim sunar. (Masanes, 2006:9). Bazı ülkelerde ise Web arşivleri karanlık arşivdir veya kullanıcılara yalnızca yerinde erişim sunmaktadır. Geciktirilmiş erişim örneklerine bakıldığında, bir Web sitesi derlendikten 3 ay sonra Harvard Üniversitesi erişim aracı WAX'da ve 6-12 ay sonra ise IA Wayback Machine'de görüntülenebilir (Niu, 2012b).

3.2.6. Koruma

Arşiv, müze ve kütüphane gibi kültürel miras kurumları büyük ölçekli Web arşivleme yaptıklarında topladıkları materyalleri kalıcı olarak korumayı hedeflerler (NA, 2011:10). Kalıcı koruma ile hedeflenen, derlenen Web materyalinin gelecekte erişilebilir olmasını sağlamaktır. Web sitesi üreticilerinin dijital materyallerin kırılganlığından haberdar olmaması ve bu materyallerin korunması için sorumluluk almayı istememeleri kültürel miras kurumlarının bu sorumluluğu almalarında etkili olmuştur (Murray ve Hsieh, 2008).

Diğer dijital kaynaklarla karşılaştırıldığında Web arşivlerini korumanın en belirgin zorluğu veri miktarının çokluğu ve dosya formatları ile medya türlerindeki çeşitliliktir (International Organization for Standardization [ISO], 2013:12). Web'in resim, video, müzik, oyun, veri tabanı ve birçok uygulamayı içinde barındıran sayfalar arası bağlantıları içeren bir mekanizmaya sahip olması da Web arşivleme sürecinde korumayı zorlaştırır. Bu nedenle Web'de yer alan bağlantıları ve veri akışını koruma Web arşivleme sürecinde büyük önem arz etmektedir.

Web içeriğinin korunması en temel düzeyde teknoloji eskimesine karşı bitlerin güvenliğini sağlanması ve onları kullanılabilir durumda tutma ihtiyacını doğurmuştur. Veri taşıyıcılarında yaşanabilecek eskime ve bozulma riskine karşı verileri kopyalayarak yeni veri taşıyıcılarına düzenli olarak aktarmak gerekir. Kurumlar bu süreçte ayrıca kaynakların işlevselliğini, hareketliliğini ve kullanıcı deneyimini koruma amacıyla öykünüm ve göç gibi dijital koruma stratejilerinden yararlanırlar. Öykünüm ve göç, dijital arşivleme sistemleri için koruma planlamanın önemli bir bölümü olarak düşünülmelidir (ISO, 2013:13). Göçün orijinal veri akışını değiştirmesi öykünümün ise orijinal veri akışını değiştirmemesi bu iki yaklaşım arasındaki temel ayrımdır. Söz konusu stratejilere ek olarak dijital koruma sürecinde yenileme, yedekleme, dijital arkeoloji, normalleştirme ve teknoloji koruma (bilgisayar müzesi yaklaşımı) gibi stratejilerden de yararlanılabilir.

Web arşivlerinin büyük ölçeği ve çok sayıda dosya formatını içeren yapısı nedeniyle bu stratejileri uygulamak Web arşivleri için zorlayıcı süreçleri barındırır.

Öykünüm eski bir sistemin davranışını ve işlevselliğini öykünücü adı verilen uzman bir yazılım kullanarak mevcut sistem üzerinde yeniden yaratmayı gerektirir (ISO, 2013:13). Temelde eski sistemin taklit edilmesine dayanır ve eski kaynaklara erişimi mümkün kılar. Geliştirilen öykünücü yazılımları pahalıdır fakat dosya dönüştürmeleri sağlanmadığından bileşenlerle tek tek ilgilenme zorunluluğu bulunmaz. Web arşivlemede, öykünücü Web sayfalarının depolandığı tarihteki ortam oynatıcıların ve Internet tarayıcıların fonksiyonlarını yeniden oluşturmak zorundadır (ISO, 2013:13). Öykünüm süreklilik arz etmez ve her durumda uygulanabilen bir strateji değildir.

Göç bir dosyanın içinde bulunduğu teknik ortamında kullanılamaz hale gelmesiyle önceden belirlenen dosya formatlarına periyodik olarak dönüştürülmesi işlemidir (ISO, 2013:13). Yeni dosya formatına yapılan her bir dönüştürme içerik değişikliklerine ve bozulmaya yol açabilir. Öykünüm ise arşivlenen malzemeleri mümkün olduğunca orijinaline yakın bir yerde muhafaza etme avantajına sahiptir, oysaki tekrarlanan göçlerle dijital objenin bütünlüğü zamanla zarar görecektir (Glanville, 2010:130).

Web arşivleme sürecinde bir dijital koruma stratejisi olarak öykünüm ve göçün hangisinden yararlanılacağına ilişkin tartışmalar sürmektedir. Bazı araştırmalar, özellikle büyük arşivlerle uğraşırken, öykünümün uzun vadede daha az maliyetli olacağını öngörmesine rağmen, şu an için göçün daha az maliyetli olduğu olduğu kanıtlanmıştır (Glanville, 2010:129-130). Öykünüm aynı zamanda arşivlenen malzemeleri mümkün olduğunca orijinaline yakın şekilde muhafaza etme avantajına sahiptir, oysaki tekrarlanan göçlerle dijital objenin bütünlüğü zamanla hasar görebilir (Glanville, 2010:130).

Dijital koruma sürecinde Web derleme yoluyla sağlanan dokümanların sadece disk üzerinde depolanması yetersizdir. Uzun dönem koruma erişim, koruma faaliyetleri ve koleksiyon yönetiminde kritik önemi bulunan Web arşivi içerisindeki kaynaklarla ilişkilendirilmiş üstveriyi gelebilecek zararlardan korumayı da kapsar (ISO, 2013:13). Arşivlenen dokümanların nereden ve ne zaman alındığını belirlemek için derleme sürecinde ilişkili üstveriler de koleksiyona kazandırılmalıdır.

Üstveri alanında koruma zorlukları Internet içeriğinin zenginliği ve çeşitliliği ile içeriğin Web arşivleri içerisinde nitelendirilmesi noktasında yaşanmaktadır (Lasfargues ve diğerleri, 2012:124). Koruma sürecinde farklı türlerde üstveri toplanmalı ve üstveriler arşivin güvenilirliğini sağlamak için korunmalıdır. Tanımlayıcı, yapısal, teknik, kaynak

ve hak üstverileri ile birlikte koruma üstverisinin varlığı koruma faaliyetleri için destekleyicidir. Seçim sürecinden sağlanan üstveriler, seçilen derleme politikası hakkında bilgi verir. Dokümanlara ilişkin üstveri (boyut, format, sağlama toplamı) koruma politikasını (göç, öykünüm, bit seviyesinde koruma) güçlendirebilir (Lasfargues ve diğerleri, 2012:124). Buna ek olarak, süreç üstverisi, arşivleme işlemi sırasında sunucu ile iletişim hakkında bilgi sağlar.

Dijital koruma, kütüphaneler ve kültürel miras kurumları için uluslararası çapta önemli bir sorundur (Glanville, 2010:128). Ağa bağlı dijital içeriğin hassaslığı ve hızlı bir şekilde üretilmesi ve kaldırılması, bu materyallerin gelecekte de mevcut olması için korunuyor olmasını ve arşivleme eyleminin yapılmasını gerektirmektedir. Kütüphaneler, bu tür eylemleri üstlenmek için en iyi durumda bulunan kurumlardır (Glanville, 2010:128). Bununla birlikte, sürdürülebilir ve kapsamlı bir dijital koruma programına ulaşmak için başta telif hakları olmak üzere birçok engel bulunmaktadır.

Hakala'ya göre (2014:180-181), telif ihlallerini ortaya çıkarmak için bir tarayıcı belgelerin alındığı URL'lerin bir listesini oluşturmalı ve saklamalıdır. URL listelerinde bulunan ve telif hakkıyla korunan belgelerin MD5 sağlama toplamalarına bakılarak bir belgenin izinsiz kopyalarının bulunup bulunmadığını kontrol edebilir. MD5 algoritması Web arşivlerinde çift kopya kontrolü için kullanılan 128 bitlik bir değerdir ve aynı zamanda bir RFC 1321 [19] Internet standardıdır. Bununla birlikte metinler arasında göze çarpan benzerlikleri bulmak için daha gelişmiş dilsel yöntemler de uygulanabilir. Derleme sırasında aynı zamanda belgenin yakalandığı kesin zamanı gösteren bir zaman damgası üretilir. Eğer doküman aynı lokasyondan yeniden alınırsa, temelde MD5'in aynı olması için ikinci zaman damgası saklanır. Arşiv daha sonra, dokümanın ilk ve son zaman damgası tarafından tanımlanan süre boyunca Web'de değiştirilmemiş ve kullanılabilir durumda kaldığını doğrulamak için kullanılabilir. Eğer üçüncü yakalama turunda doküman değişmemiş olarak bulunursa, ikinci zaman damgası güncellenir.

Koruma ayrıca format konusuyla bağlantılıdır. IIPC, içerik ve üstveri depolama konusunda Web arşivleri için özel bir ISO standardı (ISO 28500: 2009) oluşturmayı başarmıştır (Lasfargues ve diğerleri, 2012:123). Bu standartta aktarılan “.warc” formatı birden çok sayısal kaynağın bağlantılı bilgi ile birlikte bir arşiv dosyasında birleştirilmesini sağlar. Ardından kaynaklar tarihlendirilir, URI'ler tarafından tanımlanır ve MIME varlıklarını kullanarak basit metin başlıkları eklenir (Lasfargues ve diğerleri, 2012:124). WARC dosya formatı, Web arşivcileri tarafından toplanan bilgi bloklarını depolamak için Internet Archive tarafından geliştirilen ARC'nin revize edilmiş halidir.

WARC örneğinde olduğu gibi dijital nesne içeriğinin uzun süreli erişilebilirliğini kolaylaştıran dijital materyaller için bir depolama konteynerinin formatının seçilmesi ve sayısal nesnelerle ilgili davranış ve işlevlerin sürekli olarak tanınmasının desteklenmesi bir dijital arşivin temel görevlerinden bazılarıdır (Kim ve Ross, 2012:22). Bu görevler Web günlükleri veya sosyal medya gibi multimedya içeriği barındıran komplike dijital nesneler açısından özellikle zorlayıcıdır. Depolama sürecinde genel anlamda “.tar”, “.arc”, “.aff” gibi farklı konteyner formatlara da yer verilebilir. Fakat Web arşivleme sürecinde en yaygın kullanılan formatlar “.warc” ve “.arc”’dir.

Web arşivlerinde konteyner dosyaların kullanımı dosyaların bir araya getirilmesini sağlar. Böylelikle veri işleme sürecinde sayısız küçük dosya birkaç büyük dosya haline getirilir. Depolama işlemini ve veri transferini kolaylaştıran bu uygulama ortaya konan uzun süreli koruma hedefini destekler. Bunun nedeni konteyner dosyaların genellikle üstverinin depolanan kaynaklarla birlikte saklanmasına izin vermesidir (ISO, 2013:22).

Dijital nesne formatlarının dijital bilgiyi koruma üzerine etkisi hakkında gerçekleştirilen çalışmalardan elde edilen güncel bilgiye dayalı olarak, dijital korumayı desteklemesi amacıyla depolama konteyneri formatlarında yedi ana özellik aranmaktadır (Kim ve Ross, 2012:23-24):

- 1. Veri Bütünlüğü:** *Konteyner formatı, yakalama ve depolama süreçlerinde veriyi ham veriye mümkün olan en yakın biçimde korumalıdır. Bu, bir sistem diskindeki ham verilerin çoğaltılması veya akışlı içeriğin paket bazında kaydedilmesini gerektirir. Böylelikle disk bozulmaları ve veri kaybı en aza indirgenir ve hesap verebilirliği, bütünlüğü, özgünlüğü ve kurtarılabirliği en üst düzeye çıkarmak için kullanılacak olan diskte yapılan değişiklikler hakkında ek bilgi tutma olasılığı artar.*
- 2. Verilerin kurtarılabirliği:** *Seçilen konteyner formatı, mümkünse, verilerin kurtarılmasını desteklemelidir.*
- 3. Veri doğrulamasının desteklenmesi:** *Konteyner formatı, doğrulama prosedürleri ile birlikte dijital imzaları ve yasa dışı erişimden korunmak için şifreleme algoritmalarını desteklemelidir.*
- 4. Veri yönetimi süreçlerinin ölçeklendirilebilirliği:** *Konteyner formatı, herhangi bir boyuttaki dosyaları, herhangi bir boyuttaki veri kümelerini ve eklenen hizmetleri işlemek üzere ölçeklenebilir arşiv içindeki tüm işlemleri yapabilme özelliklere sahip olmalıdır. Arşiv içerisindeki dosyalara rasgele erişim ve depolama gereksinimlerini azaltmak için sıkıştırma işlemi desteklenebilmelidir.*
- 5. Şeffaflık:** *Konteyner formatı uzun süreli korumayı desteleme ve arşivdeki materyale erişim sağlama sürecinde kamuya açık şekilde kullanılabilen araç ve özellikleri*

desteklemelidir. Formatta değişiklikler yapmak, yeni versiyonları dağıtmak ve hesap verebilirliği ve özgünlüğü izlemek mümkün olmalıdır.

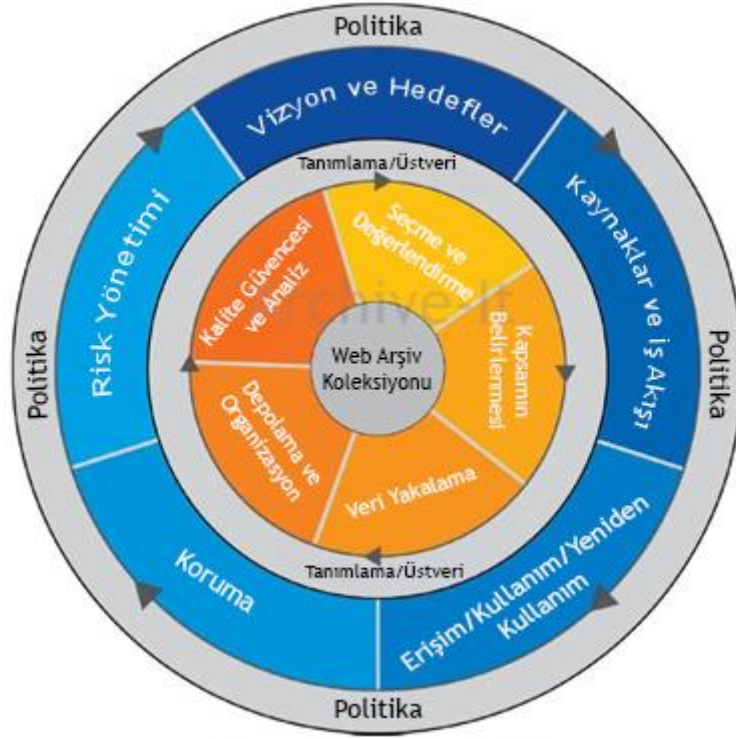
6. **Üstverileri gömme esnekliği:** Konteyner formatı, mümkünse, veri nesneleri ile birlikte kullanıcı tanımlı üstverilerinin eklenmesini desteklemelidir.
7. **Verilerin taşınmasında esneklik:** Konteyner formatı çeşitli yöntemler, ortamlar ve işletim sistemleri kullanarak erişilebilir olmalıdır. Ayrıca ister küçük ister büyük boyutta olsun veri nesneleri yer aldığı ortamda yakalanabilir, aktarılabilir ve çalıştırılabilir olmalıdır.

Ayrıca, arşiv içerisindeki kaynaklar aralarında küçük farklılıklar olsa bile ayrı ayrı tanımlanmalı ve depolanmalıdır. ISBN veya ISSN gibi geleneksel tanımlayıcı numaraların, bir elektronik kitabın birden çok versiyonunun olabilmesi ve her birinin aynı ISBN'i alabilmesi nedeniyle bir arşiv tanımlayıcısı olarak kullanılamayacağı bilinmelidir (Hakala, 2014:180).

3.2.7. Archive-It Web Arşivleme Yaşam Döngüsü Modeli

Günümüzde birçok kuruluş Web arşivleme çalışmalarının uygulama ve yöntem geliştirme aşamalarında sorunlar yaşamaktadır. Sorunlar sürekli gelişen ve belirli içerik türlerini etkili bir biçimde arşivlemeyi zorlaştıran Web teknolojilerinden kaynaklanmakta ve değişen organizasyon yapıları ile çatışan ve değişen politika kararları uygulamalardaki diğer engeller olarak karşımıza çıkmaktadır.

Web arşivlemede en iyi uygulamaların eksikliğini gidermek ve dijital korumaya temel oluşturmak gayesiyle, Archive-It ekibi Web Arşivleme Yaşam Döngüsü Modelini (WALCM) geliştirmiştir (Internet Archive [IA], 2013:1). Geliştirilen Model ile Web arşivlemede sık kullanılan iş akışlarını sunmak ve uygulanabilir bir program oluşturarak süreç yönetimini kolaylaştırmak amaçlanmıştır. Model Şekil 3'te gösterilmektedir.



Şekil 3. Archive-It Web Arşivleme Yaşam Döngüsü Modeli (IA, 2013:3)

Web Arşivleme Yaşam Döngüsü Modelinin dış yüzeyinde bir politika şeridi bulunur. Politika şeridinin içerisinde ve Model'in ilk halkasında karar verme süreçlerini yansıtan vizyon ve hedefler, kaynaklar ve iş akışı, erişim, kullanım ve yeniden kullanım, koruma ve son olarak risk yönetimi aşamaları yer alır. Bu aşamalar Model'in bir Web arşivi yaratma ve tanımlama ekseninde daha genel konuları içeren dış halkasını oluşturmaktadır (IA, 2013:21).

Modelin iç bölümünde tanımlama ve üstveri işlemlerini yansıtan ikinci şerit yer alır. İkinci şeridin içerisinde ve Model'in ikinci halkasında ise seçme ve değerlendirme, kapsamın belirlenmesi, veri yakalama, saklama ve organizasyon, kalite güvencesi ve analiz aşamaları yer alır. Bu aşamalar, yani iç halka, bir Web arşivi programının günlük aşamalarını tasvir etmektedir (IA, 2013:21). Modelin merkezinde ise tüm bu süreçlerin bir çıktısı olarak oluşturulan Web arşivi koleksiyonu yerini almaktadır.

Archive-It Web Arşivleme Yaşam Döngüsü Modeli'ni detaylı olarak inceleyecek olursak (IA, 2013:4-27):

- **Politika:** Web arşivlemenin bütün süreçlerine etki eden politika şeridinde politikanın belirlenmesi, geliştirilmesi ve uygulamalara yansıtılması yer alır. Alınan politika kararları yeni bir politika geliştirmeyi veya var olan politikayı yeni toplama aktivitelerine uyarlamayı içerebilir.
- **Vizyon ve Hedefler:** Bu aşamada ulaşılmak istenen hedefler belirlenirken yaşam döngüsünün ileri safhalarında alınacak kararlar şekillendirilir. Kuruluşlar Web arşivleme programlarının hedeflerini netleştirir.
- **Kaynaklar ve İş Akışı:** Bu aşamada kuruluşlar Web arşivleme programının üretilmesini ve sürdürülmesini sağlayacak kaynak ve iş akışlarını incelerler. Kuruluşlar Web arşivleme programının nasıl ilerleyeceğini ve değişebileceğini belirlemek için mali durumu, uzmanlığı, personeli, muhtemel ortakları ve diğer konuları içeren mevcut kaynaklarını gözden geçirirler.
- **Erişim, Kullanım ve Yeniden Kullanım:** Kuruluşlar bu aşamada koleksiyonlarına nasıl erişim sağlanacağı ve kullanıcıların içeriği nasıl görüntüleyeceği hakkında karar verirler. Başarılı bir Web arşivleme programı için erişim, kullanım ve yeniden kullanım politikaları hayattır. Kuruluşlar eğer destekleniyorsa Web materyalinin kamuya sunumunda açık erişime yönelik çalışmalarda bulunurlar. Bu süreçleri yönetmek Web arşivleme yaşam döngüsünün erişim, kullanım ve yeniden kullanım safhasının temel amacıdır. Erişim safhasında ayrıca kullanılacak teknolojinin ve erişim sağlama aracının seçimi yer alır. Modelde canlı bir Web sayfası deneyimi yaşar gibi halka arşivlenen Web sayfalarını tarama olanağı sunan açık kaynaklı görüntüleme aracı Wayback Machine'in kullanımı önerilmektedir.
- **Koruma:** Kuruluşlar Web arşivleme aktivitesinde topladıkları verilerin nasıl korunacağına ilişkin karar verirler. Verinin birden fazla kopya halinde ve birden fazla yerde muhafaza edilmesi korumayı güçlendirir. Modele göre kuruluşlar WARC dosyalarını ve ilişkili üstverileri korumak ve saklamak için Archive-It servisine güvenmekte ve koruma için yedekleme, şeffaflık ve veri bütünlüğü kontrolleri içeren en iyi uygulamaları takip etmektedirler.
- **Risk Yönetimi:** Kuruluşlar bir Web arşivleme programı yaratırken riskleri düşünür, erişimin yanı sıra telif ve izinlerini değerlendirir. Risk yönetimi kararları hangi sitelerinin kuruluşlarca arşivlenmek üzere seçileceğine bağlı olarak değişmektedir. İçeriği yakalamadan önce tüm kuruluşlar izin istemek zorunda değildir; arşiv ve kütüphane gibi birçok kuruluş Web üzerinde kamuya açık içeriği yalama konusunda yetki ve hak sahibidir. Bu durum adil kullanım ilkesiyle açıklanır. Kimi durumlarda ise hak sahiplerinden izin alınmakta veya onların talepleri doğrultusunda depolama işlemi yapılamamakta veya depolanan içerik erişimden kaldırılmaktadır. Rahatsızlık yaratabilecek ve dava konusu olabilecek yakalama işleminde ise kurumlar içeriği arşivinden kaldırma kararı alabilir.

- **Tanımlama ve Üstveri:** Modelde tanımlama ve üstveri politikaya benzer şekilde diğer aşamalarla örtüşür. Web arşivlemenin birçok yönünde olduğu gibi, en iyi uygulamalar, Web arşivleri için üstveri ve tanımlama trendlerinin üretimine ve kullanımına bağlı olarak değişmektedir. 2013 yılında toplanan verilere göre Archive-It üyesi kuruluşların %90'ı koleksiyon seviyesinde, %60'ı kaynak seviyesinde, %15'i doküman seviyesinde üstveri üretmiştir. Kaynaklar listeler URL'lerin taranması için başlangıç noktası ve ayrı ayrı arşivlenen Web sayfası dokümanlarıdır. Buna ek olarak kuruluşların %60'ı hem koleksiyon hem de kaynak üstverisi üretmektedir.
- **Seçme ve Değerlendirme:** Bu aşamada kuruluşlar hangi Web sitelerini toplayacaklarını belirler. Bu aşama yaşam döngüsünün "vizyon ve hedefler" politika aşamasına göre daha ayrıntılı ve özel karar noktalarını içerir. Yakalanacak Web siteleri ve belirli URL'lere odaklanılır.
- **Kapsamın Belirlenmesi:** Kuruluşlar Web sitesinin hangi bölümlerini, arşivleyeceğine dair tercihte bulunur. Yaşam döngüsünün bu aşaması yakalanacak Web içeriğinin formatına ve kapsam parametrelerine bağlı olarak oldukça teknik olabilir.

Taramaların kapsamını belirleme noktasında Web sitesinin tamamını listelemek yerine Web sitesinin bir bölümünü listeleyerek taramanın sınırlandırılması önerilmektedir. Sunucu sınırlama araçlarının kullanımı kurumlara belirli bir alan adı veya sunucudan yakalanan içeriği sınırlandırma izni vermektedir. Archive-It üyesi kuruluşlar arasında sunucu sınırlama araçları %73'lük kullanımla en çok kullanılanlar araçlar arasında ilk sıradadır.

Kapsamın belirlenmesi adına ayrıca tarama süresini sınırlandıran araçlardan da yararlanılır. Tarama süresini sınırlandıran araçlar ise aynı listede %64 ile en çok kullanılanlar araçlar arasında ikinci sıradadır.

Kapsamın belirlenmesinde ayrıca formatın belirlenmesi de gerekebilir. Sosyal medyanın giderek hayati bir unsur haline gelmesiyle birlikte bazı kuruluşlar flash ve javascript formatlarındaki içeriği yakalamaya çalışırken bazı kuruluşlar da PDF gibi tek biçimli formatları yakalamak isteyebilirler.

- **Veri Yakalama:** Kuruluşlar arşivlemek veya arşivlememek üzere, tarama sıklığı ve dosya türleri hakkındaki tercihleri aracılığıyla nasıl yakalama yapacakları konusunda ince ayar yaparlar. Yaşam döngüsünün kapsamın belirlenmesi ve veri yakalama aşamaları benzer eylem ve kararları içerecek şekilde örtüşür.

Kuruluşlar yakalayacakları Web sitelerine ve onların ne kadarını yakalayacaklarına karar verdiklerinde sürecin veri yakalama aşamasında planlarını eyleme geçirirler. Bu noktada tarama yazılımının bileşenleriyle muhatap olunur. Taramaların zamanlaması ve sıklığı, uzun taramaların sonlanma ve yeniden başlama zamanları belirlenir. Kuruluşlar belirli bir zaman aralığında tüm sitelerini taramamakta, farklı koleksiyonlar ve Web siteleri için farklı tarama zamanlamalarını istedikleri gibi uygulamaktadırlar. Verinin toplandığı dönemler için en popüler tarama sıklığı bir kez, haftalık ve aylık olanlardır.

Archive-It veri yakalama konusunda yaşanabilecek sürprizleri önlemek için ortaklarına veriyi gerçekten arşivlemeden taranan veriye ait raporları sunan bir test tarama özelliği

sunar. Bu seçenek kuruluşlara kaynaklarını dikkatsizce kullanmadan önce neyi arşivlemiş olacaklarını görme şansı verir. Web sitelerinin şekil ve yapı açısından çeşitliliği göz önüne alındığında, Web arşivlemenin veri yakalama aşamasında birçok sürpriz ile karşılaşılabilir. Örneğin, bir Web sitesi tahmin edilenden çok daha büyük olabilir ve depolama alanlarını tüketebilir.

- **Depolama ve Organizasyon:** Bu adım arşivlenen dosya için geçici veya uzun dönem saklama planını içerir. Bazı kuruluşlar için yaşam döngüsünün saklama ve organizasyon aşamaları koruma eylemlerini de teşkil edebilir.
- **Kalite Güvencesi ve Analiz:** Kuruluşlar arşivledikleri şeyleri ve ortaya çıkan koleksiyonun yaşam döngüsünün başında koydukları hedefleri ne ölçüde karşıladığını gözden geçirirler. Tarayıcıların ürettiği raporlar vasıtasıyla arşivlenen Web sitelerinin kalitesi ve bütünlüğü değerlendirilir. Web arşivlemenin birçok yönü gibi Web arşivleme yapan kuruluşlar arasında kalite güvencesi için ortaya çıkan tek bir en iyi uygulama bulunmamaktadır. Her bir kuruluşun kalite güvencesi için harcadığı zaman ve emek miktarı Web arşivlemedeki personel düzeyine, hedeflere göre çeşitlenmektedir. Taramalar yinelenirken, kalite güvencesi ara sıra tekrar eden bir bakım faaliyetinden ve az zaman ve dikkat harcanması gereken bir aktiviteden daha fazlasıdır. Kuruluşların büyük çoğunluğu hizmetin bir parçası olarak üretilen geçmiş tarama sonuçlarını her zaman veya sıklıkla gözden geçirmektedirler. Kuruluşlar ayrıca bir Web arşivleme programına başladıklarında ne kadar ve hangi türde materyali topladıklarıyla ilgilenme eğilimindedirler.
- **Web Arşiv Koleksiyonu:** Yaşam döngüsünün merkezi, koleksiyonun kendisi yani arşivlenen Web içeriğidir. Bu içerik önceki tüm adımların sonucudur ve korunacak olandır. Verinin yakalanması ve koleksiyonun korunması tüm Web arşivleme aktivitelerinin kalbidir ve bu nedenle modelin merkezidir.

3.3. Yasal Derleme ve Telif Hakları

Bu bölümde Web arşivleme için yasal dayanak, yasal derleme ve telif hakları konularına yer verilecektir.

3.3.1. Web Arşivleme için Yasal Dayanak

Tüm Internet kaynaklarının toplanması konusunda yasal zorunluluğun bulunduğu ülkelerde, bu durum genellikle tek bir kuruluşun sorumluluğundadır (ISO, 2013:15). Bu sorumluluk ulusal arşivler, kütüphaneler ve/veya bir konsorsiyum tarafından da paylaşılabılır.

Web arşivleme kuruluşları telif hakları, kişisel veriyi koruma ve onun gizliliği gibi fikri mülkiyete ilişkin konularda yasa önünde sorumludurlar. Arşivleme kuruluşları ayrıca hakaret içeren içeriği yayınlamak, yasadışı içeriğe sahip olmak ve onu dağıtmaktan da sorumlu tutulabilir (ISO, 2013:14). Ancak yasal derlemeye ilişkin yürürlükteki mevzuat Web arşivlemeyi etkinleştirmek ve ona destek sağlamak için derleme kurumlarına koruma sağlayabilir.

Web arşivlemenin, derleme kuruluşlarının görev ve statülerini özel olarak tanımlayan telif, yasal derleme veya herhangi bir yasayla başlatılması kurumlar adına faydalıdır (ISO, 2013:14). Söz konusu yasa ile milli kütüphaneler, milli arşivler, özel derleme kurumları, kamu arşivleri ve müzelerin statüsü ve eylemleri Web arşivlemeyi kapsayacak şekilde tanımlanabilir.

Günümüzde bazı ülkelerin Web arşivleme için henüz net bir yasal dayanağa sahip olmadığı bazılarının ise sadece uygulamalarını düzenlemek ve yorumlamak için ikincil mevzuat oluşturulduğu gözlemlenmiştir. Ulusal bir yasanın yürürlükte olduğu ülkelerde dahi yasaların Web arşivlemeyi kapsayıcılığı yorumlamaya açıktır (ISO, 2013:14). Web arşivlemeye ilişkin yasal derleme mevzuatının oluşturulmadığı durumlarda ise kurumlar yayıncılarla anlaşarak Web arşivleme faaliyetleri yürütmektedir. Bu durumda kurumlar hak sahipleriyle özel anlaşmalar imzalayabilir veya yasal riski yönetmek için arşivlenen kaynaklara erişimi sınırlayabilirler (ISO, 2013:14). Ayrıca kuruluşların genellikle Web kaynaklarına erişim ve onları yakalama konusunda açık ve belirgin olmayan izinlerle işlem yaptıkları görülmektedir.

Mevzuat için belirleyici unsur kaynakların depolanmasından önce hak sahiplerinden gereken izinlerin alınmasının gerekli olup olmadığıdır. Alınan ve alınamayan izinler kuruluşun derleme stratejisini belirler. Bu nedenle Web materyalinin sağlanmasında kullanılacak yöntem konusunda mevzuat belirleyici roldedir. Örneğin, derleme sürecinde hak sahiplerinden izin almak gerekiyorsa ve seçimli derleme düşünülüyorsa yalnızca yığın derleme tercih edilebilir. Alınan izinler doğrultusunda ayrıca Web derlemenin derinliği ve sıklığına karar verilir. Yayıncıdan alınan ve birden çok yayını kapsayan genel kapsamlı bir izin, yayıncılarla yapılan anlaşma maliyetlerini düşürebilir. Arşivlenen içeriğin doğası da izin alma veya almama konusunda karar vermeyi etkiler. Örneğin Kongre Kütüphanesi, haber kuruluşlarının bloglarını ve Web sitelerini arşivleme noktasında ilgili kuruluşun iznini ister, ancak diğer birçok Web sitesine yalnızca kütüphanenin Web sitelerini arşivlediğini bildirir (Grotke ve Jones, 2010:17).

İzin yönetimine alternatif yaklaşımlar varsayılan veya uygulanan izinler temelinde depolanan hazır haldeki materyalin kaldırılmasını gerektiren yayından kaldırma veya devre dışı bırakma politikalarıdır. Bir bilgi merkezi arşivlenen materyale yönelik telif haklarıyla ilgili sorun yaşamamak adına bir Web sitesini silmeye karar verirse oluşturacağı bir yayından kaldırma politikası kapsamında silme işlemini gerçekleştirebilir. Yayından kaldırma veya devre dışı bırakma politikaları Web arşivlerinin barındırdığı içeriği kontrol etmelerinde onlara yardımcı olabilir (Patel, 2007:27-28).

Yasal derlemenin, toprak ve teknoloji gibi sınırlamalara sahip olması nedeniyle mutlak bir çözüm olamayacağı söylenebilir. Örneğin, Yeni Zelanda'da Ulusal Kütüphane şifrelerle korunan denizaşırı içeriği ve yerel içeriği derlemek için izin almakla yükümlüdür ve telif hakkı iznine sahip olmayan içeriği koleksiyonuna dâhil etmemektedir (Cadavid, 2014:10; Paynter and Mason, 2006). Singapur Web Arşivi ise telif hakkı iznine sahip olmayan materyalleri koleksiyona dâhil eder fakat telif hakkı sahiplerinin arşivlenmiş materyalin kaldırılmasını talep edebilecekleri bir yayından kaldırma politikasını kullanır (Cadavid, 2014:10; Glanville, 2010). Hollanda'da da benzer bir yayından kaldırma politikası izlenmektedir.

Mevzuatta yer verilmesi gereken bir başka unsur da robot.txt dosyalarını dışlama uygulamasının kabul veya reddedilmesidir. Zira, hedef Web sitelerinin yakalanması veya atlanması Web arşiv koleksiyonlarında farklılığa yol açacaktır. Yığın ve alan adı derleme stratejilerinin devreye alınması halinde bu farklılık daha belirgin hale gelecektir. Örnek vermek gerekirse, 2006 yılı telif yasasına göre, Fransa Milli Kütüphanesi, Fransız Web'ini taranırken robot.txt dosyalarını dışlamayı reddedebilir (Lasfargues, Oury ve Wendland, 2008:4). Robot.txt konusu ilerleyen bölümlerde daha ayrıntılı olarak ele alınacaktır.

Mevzuat derlemenin kalitesini ve bütünlüğünü arttırmaya yardım edecek şekilde bir toplama kurumuna korunan bir bilgiyi elde etme hakkı tanıyabilir (ISO, 2013:14). Bu tip enformasyon örnekleri ulusal alan adlarını ve ücretli yayınların dijital hak yönetimine ilişkin listeleri içerebilir (ISO, 2013:14). Mevzuat ayrıca yayıncıları Web materyallerini sağlamaya mecbur edebilir ve karşı gelme durumunda cezai yaptırımlar getirebilir.

3.3.2. Yasal Derleme ve Telif Hakları

Derleme yasaları, bir ülkenin kayıtlı bilgi ürünlerinin erişilebilir kılınmasını ve bu bağlamda kültürel mirasın oluşturularak gelecek kuşaklara aktarılmasını sağlayan en önemli araçlardan biridir (Zan, 2006:7). Web arşivleme uygulamaları ise birçok ülkede yasal derleme mevzuatı ile desteklenmektedir. Yasal derleme bir yayıncının yayınlarının kopyalarını, telif hakkı mevzuatı ile güçlü bir tarihsel ilişki yaşayan ulusal kütüphaneye verme yükümlülüğünü ifade eder ve telif hakları ile birlikte entelektüel üretimin kontrolü, korunması ve yaygınlaştırılması ile ilgili yasal ürünleri kapsar (Cadavid, 2014:1).

Bir ülkenin yasal mevzuatı içerisinde derleme olgusunun varlık göstermesinin amacı; ulusun kültürel varlığını oluşturan tüm yayınların, gelecek nesillere ulaştırılması için korunmasını ve saklanmasını güvence altına almaktır ve derleme uygulamalarının gerçekleştirilmesinde asli görev genellikle derleme kanunları aracılığı ile ulusal kütüphanelere verilmiştir (Zan, 2009:300). Bu noktada yasal derlemenin temel amacı insanlığın gelecek nesilleri için entelektüel çalışmaların toplanması, muhafazası, organizasyonu ve korunması olarak ifade edilebilir. Paris Deklarasyonu'nda ise yasal derleme mevzuatının amaçları ulusal yayınların toplanması ve korunması ile ulusal yayınları oluşturan unsurların resmi bibliyografik kayıtlarının oluşturulması olarak belirlenmiştir (UNESCO ve IFLA, 1977:6).

Günümüzde kültür politikasının bir aracı ve tüm alanlardaki araştırmacıların en önemli kaynaklarından biri haline gelen yasal derlemeye ilişkin ilk mevzuat 1537 yılında Fransa'da yürürlüğe girmiştir (Cadavid, 2014:2) ve yasal derleme zamanla dünya genelinde birçok ülkede kendine yer bulmuştur. 20. yüzyıla kadar kapsamında yalnızca basılı yayınlar yer almış ve sonrasında kapsamı ses kayıt teknolojilerinin, görsel-ışitsel medyanın ve dijital bilginin gelişimi neticesinde çeşitli format ve yayınları da içine alacak şekilde genişletilmiştir. Fakat bu durum yasal derleme mevzuatı için yeni zorlukların ortaya çıkmasına neden olmuştur. Örneğin, dijital bağlamda telif hakkı, yasal derleme hedeflerinin başarılması için bir engel olabilir ve hem erişimi hem de korumayı zorlaştırabilir (Cadavid, 2014:1-2). Bu nedenle telif hakkı yasalarının güncel ve dijital alanda yeterli olması önemlidir.

Yayınların yaşam döngüsü boyunca hukuki kesinlik temin eden koleksiyon yönetiminde telif hakkı yasası önemli bir role sahiptir (Cadavid, 2014:9). Bununla birlikte, derlenen Web sitelerinin telif hakkı sahipleriyle anlaşmaya varılamaması veya yeterli yasal derleme mevzuatı olmaması durumunda Web arşivleri, koleksiyonlarını erişilebilir hale getirirken çeşitli sorunlarla karşı karşıya kalacaklardır (Cadavid, 2014:9). Yasal derlemeye ilişkin net yasal çerçeveler oluşturulmazsa Web arşivleri karanlık arşiv haline gelecektir.

Danimarka, Fransa, Almanya, Norveç, Kanada, Yeni Zelanda ve Güney Afrika'daki Ulusal Kütüphaneler de dâhil olmak üzere birçok ulusal kütüphane, yasal derleme yasalarını ağ üzerinde sunulan dijital materyali kapsayacak şekilde genişletmeden yana hareket ettiler (Beunen ve Schiphof, 2006:17). Bu durum kütüphaneye, telif hakkı ihlalinin dolaylı dava açma tehdidinden kaçınırken, ulus için kültürel önem taşıyan materyalleri arşivlemek ve korumak için açık bir yetki vermektedir (Glanville, 2010:130).

Dijital kayıtların ve Web verisinin yönetimi, koleksiyonların toplanması, depolanması, erişimi ve kullanımı ile ilgili telif hakkı sorunlarını içerir. Bunun nedeni Web arşivleme sürecinde içeriklerin kopyalanarak yeniden üretilmesidir. Bu kopyalar, telif hakları ile korunan eserlere atıfta bulunulduğunda, kanunen veya telif hakkı sahibi tarafından yetkilendirilmiş olmalıdır (Cadavid, 2014:4).

Web kaynaklarını edinirken, telif hakkı sahiplerinden izin almak isteyip isteme kararı, Web arşivinin yasal ortamına, Web arşivinin ölçeğine ve arşivlenen içeriğin ve arşivleme organizasyonunun niteliğine bağlıdır. Yasal derlemenin Web kaynaklarını kapsadığı bir ülkede, yasal derleme kütüphanesinin o ülkede üretilen Web yayınlarını arşivlemek için izin almasına gerek yoktur. Ayrıca kamu kayıtlarını koruma noktasında yasal yetkiye sahip olan devlet arşivleri de kayıt üreticilerinden izin almaya ihtiyaç duymaz. Fakat aynı yasal çevrede Web arşivleme kurumları açıkça yetkilendirilmediği takdirde, büyük ölçekli Web arşivleme yerine küçük ölçekli Web arşivleme için izin istemek daha olasıdır, çünkü nispeten az sayıda telif hakkı sahibinden izin almak daha kolay idare edilebilir.

Ulusal Web arşivlerinde bilginin toplanmasında ve korunmasında olduğu kadar ona erişim sağlamada da telif hakkı sorunları ortaya çıkmaktadır. Telif hakkı noktasında Web arşivlemeye yönelik erişim kısıtlamalarının temel nedeni, telif hakkı sahibinden önceden izin almadan çalışmaların yapılmasına izin verilmemesi ve kullanımın, telif hakkı sahiplerinin ekonomik çıkarlarını etkileyebilmesidir (Cadavid, 2014:11). Bununla birlikte, Web arşivlerine erişilebilirliğin olumsuz bir ekonomik etkiye sahip olduğunun ampirik bir kanıtı yoktur (Cadavid, 2014:11).

Web arşivleri, beşerî sosyal ve doğa bilimleri alanlarında birçok araştırmaya kaynaklık etmektedir (Dougherty, Meyer, Madsen, Heuvel, Thomas ve Wyatt, 2010:9) Yerinde erişim sunan ulusal kütüphane Web arşivlerine sahip ülkeler, eğitim ve araştırma telif hakkı yasalarının istisnaları olduğu için onları araştırma veya çalışma durumunda yetkilendirebilir (Cadavid, 2014:12). Telif hakkı esnekliğinin olmaması, sayısal beşerî bilimlerin son yıllarda benimsediği araçların kullanılmasını engellediğinden bu gibi

durumlar da Web arşivinin kullanımı sınırlandırılabilir. Örneğin Fransa Milli Kütüphanesi'ndeki araştırmacılar, Fransız Web Arşivinin veri madenciliği analizi için kullanılmasına ilişkin kısıtlamalarla karşı karşıyadırlar (Stirling ve Illien, 2011:18).

3.4. Web Arşivleme Yazılım Araçları

Web arşivleme sürecinde Web içeriğini yakalamak, toplamak, korumak ve ona erişim sağlamak için çeşitli yazılım araçlarından yararlanılmaktadır. Sürecin farklı aşamalarında yazılım araçları birbirine entegre edilerek kullanılmakta ve kurumların kullanabilmesi amacıyla kaynak kodları genellikle ücretsiz ve açık kaynak olarak sunulmaktadır. Ancak ihtiyaçların etkin ve verimli bir şekilde karşılanmasını sağlamak için bir çözüm seçmeden önce iş ve arşivleme ihtiyaçlarının açıklığa kavuşturulması gerekmektedir.

Web arşivleme sürecinde sürdürülen araştırma ve uygulamalarla kullanılacak araçlar ve yöntemler üzerine bir fikir birliğine varıldığı söylenebilir. Yine de Web çok hızlı bir tempoda değişmekte olduğundan, yazılım alanında kalıcı uyarlamalar gereklidir (Lasfargues ve diğerleri, 2012:120).

Çalışmamızda Web arşivlemede kullanılan yazılım araçları sağlama, küratör, yeniden oynatma, depolama ve muhafaza, arama ve keşif ve analiz araçları ile yardımcı uygulamalar olarak gruplandırılmıştır. Söz konusu Web arşivleme yazılım araçları dâhil olduğu grup ve açıklamalarıyla birlikte Tablo 1’de sunulmaktadır:

Tablo 1. Web Arşivleme Yazılım Araçları

Sağlama Araçları	ArşivFacebook	Bireylerin Facebook hesaplarını arşivlemeleri için oluşturulmuş bir Mozilla Firefox eklentisi [20].
	Brozzler	Gömülü URL’leri ve sayfaları ayrıştırmak ve bağlantıları çıkarmak için kullanılan dağıtık bir Web tarayıcısı [21].
	F(b)arc	Grafik API’sini kullanarak Facebook’tan veri arşivlemek için kullanılan bir komut satırı aracı ve Python kütüphanesi [22].
	grab-site	WARC çıktısı, tüm taramalar için gösterge paneli ve dinamik yok sayma örüntüleri sunan Web tarayıcısı [23].
	Heritrix	IA’nın açık kaynaklı, genişletilebilir ve Web ölçekli arşiv tarayıcısı projesidir. Tarayıcı, gelecekteki araştırmacı ve nesiller adına kültürümüzün sayısal eserlerini toplamayı ve korumayı hedeflediğinden ona “kadın varis” anlamına gelen bu antik isim uygun görülmüştür. Heritrix [24], robots.txt dışlama direktiflerine ve META robot etiketlerine saygı

		göstermek ve normal Web etkinliğini bozma olasılığı bulunmayan ölçülü, uyarlanabilir bir hızda materyal toplamak için tasarlanmıştır. Avrupa genelinde kültürel miras kurumlarında kullanım oranı %80'dir (Lasfargues ve diğerleri, 2012:120).
	html2warc	Çevrimdışı verileri tek bir “.warc” dosyasına dönüştürmek için kullanılan basit bir program [25].
	HTTrack	Açık kaynak kodlu bir Web kopyalama yardımcı programı [26].
	Lentil	Instagram'dan görüntülerin toplanmasını destekleyen ve çeşitli görsellere göz atma, favori görselleri seçme, onları paylaşma ve denetleme fonksiyonu sağlayan bir Web uygulaması [27].
	SiteStory	Bir Web istemcisi (tarayıcı) ile bir Web sunucusu arasında gerçekleşen işlemleri seçimli olarak yakalayan ve saklayan işlem merkezli bir arşiv [28].
	twarc	Twitter JSON verilerini arşivlemek için bir komut satırı aracı ve Python kitaplığı [29].
	WARCreate	Bir kullanıcının göz atılabilir herhangi bir Web sayfasından bir “.warc” dosyası oluşturmaya izin veren bir Google Chrome uzantısıdır [30]. Ortaya çıkan dosyalar daha sonra Internet Archive'ın açık kaynak Wayback Machine gibi diğer araçlarda kullanılabilir. Sonuçlarını standartlaştırılmış bir şekilde güvenli bir şekilde arşivlemek isteyenler için kişisel bir Web arşivleme çözümü olma yolunda ilerleyen bu araç Digital Preservation 2012'de National Digital Stewardship Alliance (NDSA) tarafından bir yenilik ödülü almıştır.
	WAIL	Web sayfalarını korumak ve tekrar oynatmak için herkes tarafından kolaylıkla kullanılabilmesi amaçlanan çoklu Web arşivleme araçlarının üstüne eklenen grafik kullanıcı arayüzü [31].
	Web2Warc	Herkesin kendi küçük Web arşivlerini oluşturmaya olanak tanıyan kullanımı kolay ve yüksek düzeyde özelleştirilebilir bir tarayıcı [32].
	Webrecorder	Göz attığınız herhangi bir Web sitesinin yüksek doğruluklu ve etkileşimli kayıtlarını oluşturmaktadır [33].
	Wget	Açık kaynak kodlu bir dosya alma programı [34].
	Wpull	Wget uyumlu bir Web indiricisi ve tarayıcısı [35].
	Warrick	Web arşivlerinden Web sitelerini yeniden oluşturmak için kullanılan indirilebilir açık kaynaklı bir araç ve Web servisi [36].
Küratör Araçları	CINCH	Internet üzerinde erişilebilir dokümanların yığın halinde alınması ve bir koruma sistemine aktarılması için açık kaynaklı bir araç [37].
	NetarchiveSuite	Kütüphanecilerin Web materyalini derlenmelerine, tanımlamalarına ve kontrol etmelerine izin veren bir kürasyon aracıdır. Küçük seçimli derlemelerden, tüm ulusal alan adlarının derlenmesine kadar ölçeklendirilebilir. Sistem, herhangi bir sayıdaki makine üzerinde tamamen

		dağıtılabılır ve derlenen materyalin birden fazla kopyasını işleyen güvenli bir depolama modülünü ve kalite güvence sürecini otomatikleştiren bir kalite güvence aracını içerir [38].
	Web Curator Tool (WCT)	Seçimli Web derleme sürecini yönetmek için kullanılan bir araçtır. Kütüphanelerde ve diğer toplama organizasyonlarında kullanılmak üzere tasarlanmıştır [39].
	DeepArc	Kullanıcıların var olan ilişkisel veri modelleri ile XML Şemaları olarak belirtilen bir veya birkaç hedef veri modeli arasında bir eşleme oluşturmalarını sağlayan grafiksel bir düzenleyicidir [40].
Yeniden Oynatma Araçları	PyWb	Web arşiv yeniden oynatma araçlarının Python uygulaması [41].
	OpenWayback	Wayback Machine'i geliştirmeyi amaçlayan açık kaynak projesi. Arşivlenen Web sitelerini kullanıcının tarayıcısında yürütmek için kullanılan anahtar yazılım [42].
Depolama ve Muhafaza Araçları	HTTrack2ARC	HTTrack çıktısını ARC formatına dönüştüren bir araç [43].
	Java Web Archive Toolkit (JWAT)	ARC ve WARC dosyalarını okumak ve doğrulamak için kullanılan bir araç [44].
	JHOVE2	Açık kaynak kodlu bir format tanımlama aracı. Yeni format modülleri ARC, WARC ve GZIP formatlarını içerir [45].
	WarcManager	WARC dosyalarının içeriğini keşfetmek için kullanılan bir araç [46].
Arama ve Keşif Araçları	warc-discovery	WARC ve ARC dizinleme ve bulma aracı [47].
	NutchWAX (Nutch with Web Archive eXtensions)	Web arşivlerinde arama yapmak için Nutch arama motorunu ve uzantılarını kullanan Web arşivi arama ve dizinleme aracı [48].
	WERA (WEB aRchive Access)	Bir Web arşivi arama ve navigasyon uygulaması. WERA, NWA Toolset'den geliştirildi. Web arşivlerine Wayback Machine benzeri bir Internet arayüzü erişimi sağlamakta ve tam metin araması yapmaktadır [49].
	Wayback Machine	ARC veya WARC dosya formatında depolama yapan Web arşivleri için gösterim aracı olan Wayback Machine, arşivlenen Web kaynaklarının geçici olarak gezinilmesine olanak tanır [50].
	Xinq (XML INquire)	Bir XML veri tabanına erişme amacıyla kullanılan bir arama ve göz atma aracı [51].

Analiz Araçları	ArchiveSpark (S table)	Web arşivleri için kolay bir şekilde veri işleme, çıkartma ve türetme sağlayan araç [52].
	warcbase	Web arşivlerini analiz etmek için kullanılan açık kaynaklı bir platformdur [53].
	har2warc	HTTP Arşivi (HAR) formatını Web Arşivi (WARC) formatına dönüştürür [54].
	Jwat	WARC, ARC ve GZIP dosyalarını okuma, yazma ve doğrulama için kullanılan araç [55].
	The Archive Browser	Arşiv içeriğine göz atmanıza ve bunları ayıklamanıza izin veren bir programdır. Dosyaları arşivlerin içinden açmanıza ve ön izleme yapmanıza izin verir [56].
Yardımcı Uygulamalar	The Unarchiver	WARC dâhil birçok arşiv formatının içeriğini ayıklamak için kullanılan bir program [57].
	Warcats	Web ARChive (WARC) dosyalarını işleme aracı [58].
	WarcPartitioner	Bölme (W) ARC Dosyaları, MIME tür ve yılına Göre parçalara ayırma aracı [59].
	webarchive-indexing	WARC ve ARC dosyalarının Hadoop veya yerel dosya sistemi üzerinde toplu olarak dizinlemek için kullanılan araç.

3.5. Web Arşivleme ile İlişkili Standartlar

Bu bölümde Web ve Web arşivlemeye ilişkin bazı standartlara yer verilecektir.

3.5.1. Robot.txt Protokolü

Web robotları, tarayıcı veya örümcek olarak adlandırılan Web üzerinde hareket eden programlardır. Google gibi arama motorları Web içeriğini dizinlemek için; spam gönderenler e-posta adreslerini taramak için bunları kullanırlar.

Bir Web sitesinin robotlarla taranıp taranmayacağına yönelik izin verme noktasında sıklıkla robot.txt protokolünden yararlanılır. Robots.txt protokolü ile tarayıcının bir Web sitesinin tamamını veya bir bölümünü yakalaması engellenebilir. Web sitesinin içine gömülen /robots.txt dosyası, herkese açık bir dosyadır. Herkes, sunucunuzun hangi bölümlerinin robotların kullanmasını istemediğinizi görebilir.

Robots.txt protokolü kullanmamak arşivleme için izin vermek olarak yorumlanabilirken, bunların kullanımı da her zaman arşivlemeyi reddetme anlamına gelmez. Robots.txt önlemleri genellikle Web sitelerinin arama motorları tarafından dizinlenmesini ve önbelleğe alınmasını önlemek için kullanılır (Glanville, 2010:132). Bu noktada bilinmesi gereken bu materyallerin derlenememesi durumunda, bu materyali referans alan sitelerin eksik veya kullanılamaz hale geleceğidir (Glanville, 2010:132).

Herhangi bir standart organıyla bağı bulunmayan /robots.txt, fiili bir standarttır. Robot.txt protokolüyle ilgili iki önemli doküman 1994 tarihli Robot Hariç Tutma Standardı (A Standard for Robot Exclusion) [60] ve 1997 tarihli Web Robotlarının Kontrolü için Bir Yöntem (A Method for Web Robots Control) 'dır [61].

3.5.2. WARC Dosya Formatı

Kültürel miras kurumları, Web tarayıcıları gibi Web ölçekli araçları kullanarak Web materyallerini toplamak ve takip etmek için en uygun yolları bulmaya çalışmaktadırlar. Bir Web tarayıcısı, bir dizi ilkeye göre otomatik bir şekilde Web'e göz atan bir programdır; bir URL listesi ile başlayarak her sayfayı bir URL ile tanımlar ve sayfadaki tüm köprüleri (örneğin diğer sayfalara, resimlere, videolara, komut dizilerine veya stil talimatlarına bağlantılar) bulur ve onları tekrar tekrar ziyaret etmek için URL listesine ekler (International Organization for Standardization [ISO], 2009:vii). Bu noktada milyarlarca Web sayfası nesnesinin derlenmesi ve yönetimi oldukça büyük bir zorluk olarak karşımıza çıkmaktadır.

Miras kurumlarının temel hedefi Web üzerinde bulunan dijital nesnelerin yakalanması, toplanması ve basit ve güvenli bir biçimde depolanmasıdır. Bu süreçte dijital nesneler üzerinde işlem yapabilmek ve çok sayıda veri nesnesini bir bütün haline getirerek yönetmek için bir konteyner formatına ihtiyaç duyulur.

WARC (Web Archive) dosya formatı, taranan Web sitelerinden toplanan dosyaları aralarındaki ilişkileri koruyarak bir paket haline getirmek, onları tanımlamak ve üstverileri ile birlikte arşivlemek için tasarlanan bir konteyner dosya formatıdır. WARC Web'den derlenen içerikleri depolama amacıyla IA'nın oluşturduğu ARC dosya formatının [ARC_IA] bir uyarlamasıdır. ARC formatının WARC formatına dönüştürülmesi sürecinde IIPC üyeleri Avustralya, Kanada, Danimarka, Finlandiya, Fransa, İzlanda, İtalya, Norveç, İsveç, İngiliz Ulusal Kütüphaneleri ve Kongre Kütüphanesi (ABD) ile Internet Archive'ın deneyimleri ve çalışmaları etkili olmuştur.

WARC formatı ARC formatında kaydedilen birincil içeriğin yanında atanmış üstverileri, yapılan dönüştürmeleri ve kaynak bölümlendirmelerini de barındırır. 2009 yılında bir ISO standardı olarak onaylanan WARC ile amaçlananlar (ISO, 2009:1):

- *HTTP, DNS ve FTP gibi genel Internet uygulama katmanı protokollerinden elde edilen içeriği ve kontrol bilgilerini saklamak;*
- *Diğer depolanmış verilere (örneğin, konu sınıflandırıcısı, keşfedilen dil, kodlama vd.) ilişkin üstverileri saklamak;*
- *Veri sıkıştırmasını desteklemek ve veri kayıt bütünlüğünü korumak;*
- *Derleme protokolünden gelen yanıt bilgilerini ve tüm kontrol bilgilerini saklamak,*
- *Depolanan diğer verilerle bağlantılı veri dönüştürmelerinin sonuçlarını saklamak;*
- *Diğer depolanmış verilere (aynı veya büyük ölçüde benzer kaynakların var olması durumunda depolamayı azaltmak için) bağlı olarak mükerrerlik algılama faaliyetlerini saklamak;*
- *Mevcut işlevselliği aksatmadan genişletme sağlamak;*
- *İstenildiği takdirde, kısaltma veya bölümleme yoluyla aşırı uzun kayıtların taşınmasını desteklemek şeklinde sıralanabilir.*

Bir Web derlemesi bağlamında, Web sitelerini oluşturan dosyalar WARC kayıtları olarak WARC dosyalarında saklanır (ISO, 2009:19). Web derleme işlemi yapılandırmasına bağlı olarak, bir Web sitesinin farklı parçaları bir veya daha çok WARC dosyasında yer alabilir. Dolayısıyla, bir Web sitesinin arşivini kullanıcılara sunmak için, erişim yazılımının farklı WARC kayıtlarını WARC dosyalarından ayıklaması gerekebilir.

Kayıt boyutları uygun olduğu zaman, her bir WARC dosyası için ortalama 1 GB alan ayrılması önerilir (ISO, 2009:26). Uygulama gerekliliklerine bağlı olarak büyük boyutlu kayıtlar kesilebilir, parçalara bölünebilir veya daha büyük boyuttaki WARC dosyalarına yerleştirilebilir. Ayrıca, dosya mükerrerliğinin önüne geçme adına WARC dosya adlarının mükerrerliğini olasılıksız hale getiren uygulamalardan yararlanılabilir.

WARC dosyalarını adlandırma aşamasında ön ek, zaman damgası, seri, tarama makinesi adı ve format uzantıları kullanılır. Önek genellikle bu dosyayı oluşturan projeyi veya taramayı yansıtan bir kısaltma; zaman damgası dosyanın başlangıçta başladığı zamanı gösteren 14 basamaklı bir GMT zaman damgası; seri dosyayı oluşturan süreçte kullanılan benzersiz seri numarası ve tarama makinesi, dosyayı oluşturan makinenin alan adı veya IP adresidir (ISO, 2009:26). Ardından “.warc” dosya uzantısı ve depolama alanında tasarruf sağlama amacıyla kullanılan bir sıkıştırma formatı olan “.gzip” gelmektedir. WARC dosyalarının adlandırılmasına yönelik kullanılan şablon aşağıda sunulmaktadır (ISO, 2009:26):

Önek-ZamanDamgası-Seri-Taramamakinesi.warc.gz

3.5.3. HTML-XML-CSS-JavaScript ve HTTP

HTML ve XML, W3C tarafından muhafaza edilen ve çok iyi düzeyde dokümanite edilmiş temel Web standartlarıdır. HTML'nin en son sürümü, HTML5 olup, bu standart multimedya materyalinin derlenmesini kolaylaştırma potansiyeline sahiptir (Pennock, 2013:5). XML ise, W3C tarafından tanımlanmış bir standart, veri saklamanın yanında farklı sistemler arasında veri alışverişi yapmaya yarayan bir ara format ve bir işaretleme dilidir. İnsanlar ve işletim sistemleri tarafından okunabilecek belgeler oluşturmaya yarar.

CSS bir Web sitesinin görselleştirilmesinde kullanılan kodlama teknolojisidir. JavaScript ile birlikte Web tarayıcılarının karşılaştıkları diğer standartlar arasında yer alır. JavaScript, yaygın olarak Web tarayıcılarında kullanılmakta olan dinamik bir programlama dilidir. Etkileşimli ve dinamik Web sayfaları oluşturmaya yarar. HTTP ise hipermetin alışverişinde kullanılan bir protokoldür ve verileri Web üzerinden iletmeye yarar.

3.5.4. ISLI

Teknoloji alanındaki gelişmeler ister dijital isterse geleneksel biçimde depolansın farklı türdeki kaynakların birbirine bağlanma yeteneğini güçlendirmiştir (International Organization for Standardization [ISO], 2015:v). Kaynaklar arasında bağlantılar oluşturmak için bir bağlantı tanımlayıcısı kullanmak, kaynakların birlikteliği için bir temel oluşturur ve sistemler arasında iş birliğine dayalı içerik oluşturmaya ve veri birlikte işlerliği oluşturmaya destekler. Tanımlayıcılar ayrıca hem multimedya hem de diğer alanlarda bağlantılı kaynakların değerini arttırarak yeni kaynakların ortaya çıkmasını sağlayacaktır.

Uluslararası Standart Bağlantı Tanımlayıcı (International Standard Link Identifier-ISLI) bilgi ve dokümantasyon alanındaki varlıklar arasındaki bağlantıları tanımlamak için kullanılan evrensel bir standarttır [62]. ISLI, Uluslararası Standartlaştırma Örgütü (ISO) tarafından geliştirilmiş ve 2015 yılında yayınlanmıştır. Bağlantılı bir varlık fiziksel (örneğin, basılı kitap), elektronik (örneğin, metin, ses ve video) veya soyut (örneğin, referans çerçevesinde fiziksel konum veya günün saati) olabilir [62].

ISLI'nın temel fonksiyonu varlıkların özelliklerini değiştirmeden yenilikçi ve güvenilir bir şekilde bağlantı vermek olarak ifade edilebilir. ISLI ile kaynakların erişilebilirliğini arttırmak amaçlanır ve böylece kaynak kullanımı yoluyla daha fazla değer yaratılır. ISLI, varlıkların içeriğini, mülkiyetini, erişim hakkını veya mevcut tanımlamasını değiştirmez ve onlara özelliklerini değiştirmeden yenilikçi ve güvenilir bir şekilde bağlantı vermeyi sağlar.

ISLI kullanımı ayrıca kullanıcıların ses dosyaları veya video dosyaları gibi zengin medya içeriğini daha kolay bir şekilde deneyimlemelerine olanak tanır (ISO, 2015:v). Dahası, sadece geleneksel kullanıcılara zenginleştirilmiş bilgi sunmakla kalmayıp, aynı zamanda görsel olarak engelli bireylere metin materyalinin erişilebilirliğini geliştirir ve eğitim materyallerinin yeteneklerini genişletirler (ISO, 2015:v).

ISLI'nın bağlantı modeli üç unsuru içerir: bir kaynak, bir hedef ve aralarındaki bağlantı. Bir ISLI kodu ile tanımlanan bir bağlantı, kaynaktan hedefe doğrudur ve tek yönlüdür. Bir ISLI kodu, Şekil 4'te gösterildiği gibi bir kaynak ile hedef arasındaki bağlantıyı belirtir (ISO, 2015:2).



Şekil 4. ISLI Bağlantı Modeli (ISO, 2015:2).

ISLI, makine okumasına yönelik bir koddur ve kod üç bölümden oluşur: servis kodu, bağlantı kodu ve kontrol basamağı (ISO, 2015:2). Servis kodu altı basamaktan oluşur. Bağlantı kodu değişkendir ve basamak sayısının uzunluğu her hizmet için ayrı ayrı tanımlanır. Kontrol hanesi ise servis kodu ve bağlantı kodundan hesaplanan bir basamaktan oluşur ve hata tespiti için kullanılır.

Bir ekranda bir ISLI kodu gösterilmesi veya yazdırılması gerekiyorsa, alanları ayırmak için tire simgesi "-" eklenmeli ve sayısal koddan önce "ISLI" ifadesi eklenmelidir. Fakat "-" ve "ISLI" sembolleri tanımlayıcının bir parçası olarak görülmez (ISO, 2015:2). Aşağıda örnek bir ISLI koduna yer verilmiştir.

ISLI 116063-4520086293791473426443001-9

3.5.5. Üstveri Standartları

Web arşivleme yaşam döngüsünde derleme kurumlarının gereksinimlerine bağlı olarak doğrudan veya dolaylı olarak kullanılan standartlar MARC 21, ISAD(G), Dublin Core, LCSH, DDC, METS ve PREMIS olarak sıralanabilir.

Bugüne kadar MARC 21, ISAD(G) ve Dublin Core tümüyle Web arşivlerinde tanımlayıcı üstverinin kaydedilmesi için kullanılmış ve LCSH ve DDC kullanımıyla kaynak keşfi daha da gelişmiştir (Pennock, 2013:18). Bu standartlara Kongre Kütüphanesi'nin bir çalışması olan BIBFRAME de eklenebilir.

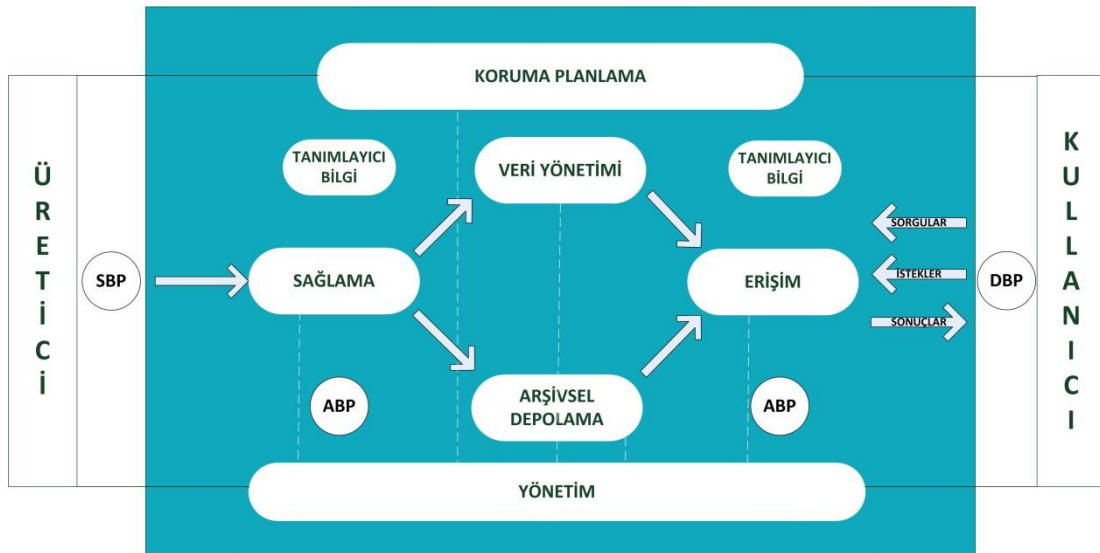
Tanımlama standartlarıyla benzer şekilde hem METS hem de PREMIS Web arşivleri hakkında ek bilgi kaydetmek için WARC konteynerleri içinde kullanılabilir. Bu standartları tek tek ele alacak olursak:

- MARC 21 Format for Bibliographic Data: MARC 21 bibliyografik veri giriş standardı metin, ses, görüntü vb. materyaller için bibliyografik verileri, otorite bilgisini, materyal ve sınıflandırma bilgilerini makine tarafından okunabilir biçimde sunma ve değiştirme için yaygın olarak kullanılan standarttır. Bibliyografik veriler genellikle başlıkları, isimleri, konuları, notları, yayın verilerini ve bir maddenin fiziksel açıklamasıyla ilgili bilgileri içerir [63].
- ISAD(G) General International Standard Archival Description: Uluslararası bir arşivsel tanımlama standardı olan ISAD(G), arşiv tanımlamaları için genel rehberlik sağlar. Mevcut ulusal standartlar ile birlikte veya ulusal standartların geliştirilmesi için kullanılabilir [64].

- Dublin Core (DC) Metadata Element Set: DC Metadata veri seti geniş bir yelpazede yer alan kaynakları tanımlamak için kullanılan ve on beş bileşenden oluşan bir standarttır [65].
- Library of Congress Subject Headings (LCSH): Kongre Kütüphanesi Konu Başlıkları Listesi, Kongre Kütüphanesi'nde bulunan kaynakları kataloglama için 1898'den beri kullanılan bir standarttır. Buna ek olarak LCSH, çoğunlukla çeviri amaçlı olarak uluslararası düzeyde kullanılmaktadır [66].
- Dewey Decimal Classification (DDC): Dewey Onlu Sınıflama Sistemi, farklı formatlarda bulunan bilgileri organize etmek için kullanılan bir sınıflama standardıdır. Başta Amerika Birleşik Devletleri olmak üzere birçok ülkede yaygın bir biçimde kullanılmaktadır [67].
- BIBFRAME: Kongre Kütüphanesi tarafından oluşturulmuştur. Standart hem Web'de hem de daha geniş ağ dünyasında bağlı veri tekniklerine dayanan bibliyografik tanımlamanın geleceği için bir temel oluşturmaktadır [68].
- Metadata Encoding and Transmission Standard (METS): Metadata Kodlama ve Aktarım Standardı, dijital kütüphane içindeki nesnelerle ilgili tanımlayıcı, yönetsel ve yapısal üstveriyi kodlamak için bir standarttır [69]. Gelecekte farklı biçimlerdeki üstveriyi bir araya getirmek için nesneleri METS dokümanları olarak kodlama düşüncesi, gelecekteki Web sayfalarının korunması ve onlara erişim için gereklidir (Guenther ve Myrick, 2007:152). Standart, uzun süreli korumayı destekler.
- Preservation Metadata: Implementation Strategies (PREMIS): Dijital nesnelerin korunmasını desteklemek ve uzun süreli kullanılabilirliklerini sağlamak için üstverilere yönelik uluslararası standarttır. PREMIS Veri Sözlüğü ve onun destekleyici belgeleri, dijital arşivleme sistemlerinde koruma meta verilerini uygulamak için kapsamlı, pratik bir kaynaktır [70].

3.5.6. Açık Arşiv Bilgi Sistemi Referans Modeli

Açık Arşiv Bilgi Sistemi Referans Modeli (Open Archival Information System Reference Model) dijital arşivlere ve dijital korumaya yönelik en sık kullanılan başvuru kaynaklarından biridir. Model dünya genelinde birçok araştırma projesinde kullanılmıştır ve bir ISO standardı [71] olarak kabul edilmiştir. Şekil 5'te Açık Arşiv Bilgi Sistemi Referans Modeli gösterilmektedir.



Şekil 5. Açık Arşiv Bilgi Sistemi Referans Modeli (ISO, 2012b:45)

Model’de arşivleme ve koruma faaliyetleri bilgi paketleri yardımıyla gerçekleştirilir. Sağlama, arşivleme ve dağıtım bilgi paketlerine dönüştürülen içerik ve koruma tanımlama bilgileri konteyner dosyalar haline getirilir ve dolaşıma sokulur. Sağlama, arşivsel depolama, veri yönetimi, erişim, yönetim ve koruma planlama aşamalarından oluşan Model, özünde dijital bilgiyi korumak ve ona daimî erişim sağlamak amacıyla oluşturulmuş kavramsal bir çerçevedir. Model, Web arşivlerine özgü olmasa da karmaşık dijital içeriğin belirli bir teknik çözüm veya araçtan bağımsız olarak ölçekli olarak depolanması ve yönetilmesi için ideal bir iş akışı sunar (ISO, 2012b:1).

3.6. Web arşivlemede Mevcut Durum ve Zorluklar

Web siteleri yakalanması, tanımlanması, yönetilmesi ve korunması en karmaşık ve zorlayıcı dijital nesnelerdir (Guenther ve Myrick, 2007:143). Hızla değişen doğaları ve kısa yaşam süreleriyle hareket eden bir hedefe benzetilebilirler. Materyallerin değişmeden veya yok olmadan derlenmesi, derlenen materyalin tanımlanması, Web sitesinin sınırlarını belirleme, derleme sonrasında toplanan versiyonların yönetimi ve ona erişim Web arşivcilerin üzerinde çalıştığı başlıca alanlardır.

Web arşivleme sürecinde yaşanan zorluklar bu bölümde yasal zorluklar, seçim zorlukları, Web mimarisi ve tarayıcı sınırlılıkları, sosyal medya arşivleme, virüsler ve zararlı yazılımlar ve son olarak veri tekilleştirme başlıkları altında incelenecektir.

3.6.1. Yasal zorluklar

Web arşivlemede yaşanan zorluklardan ilki yasal derleme mevzuatın toplama kurumuna yetki vermemesidir. Başka bir ifadeyle yasal zorlukların aşılmasında ülkedeki mevzuat ve toplama kurumunun sahip olduğu yetki belirleyici olmaktadır ve yasal derleme mevzuatının Web arşivlemeyi icra eden toplayıcı kurumu desteklemesi önemlidir.

Derleme mevzuatına ek olarak fikir ve mülkiyet hakları alanlarında da zorluklar yaşanmaktadır. Telif hakkı mevzuatı bazı eserlerin kopyalanmasına açıkça izin vermediğinden, uzun vadeli bir koruma stratejisinin parçası olarak Web sitelerinin kopyalarının üretimi soruna neden olabilir (Pennock, 2013:9). Bu sorunun aşılması adına içerik sahipleriyle bir araya gelerek lisans anlaşmalarının yapılması gerekecektir. İçeriğin kopyalarını alma ve ona erişim noktasında bazı Web sitelerinde ise Creative Commons veya Crown Copyright gibi açık lisans ve telif hakkı bilgileri gösterilmektedir. Bu süreçte yayın yoluyla hakaret, intihal suçlamaları ve veri koruması ya da gizlilik de göz önüne alınmalı ve her biri için toplama yaklaşımları, bir kuruluşun Web arşivleme politikasında uygun hukuki önerilere dayanarak açıkça belirtilmelidir (Pennock, 2013:9). Ayrıca Web arşivleme yapan kurumun çeşitli politika ve standartlardan yararlanması yaşanan sorunların aşılmasında yarar sağlar.

3.6.2. Seçim Zorlukları

Web derleme yaklaşımları bölümünde detaylı olarak verildiği gibi Web arşivlemeye yönelik yaklaşımlar alan adı, yığın, seçimli, tematik, olay eksenli ve karma derleme olarak gruplanabilir. Hangi yaklaşım seçilirse seçilsin koleksiyona dâhil edilecek veya dışlanacak kaynaklar kurumsal derleme politikalarıyla uyumlu olacak şekilde oluşturulan seçim politikalarıyla belirlenir. Neyin derleneceğini belirlemek, başka bir ifadeyle seçim aşaması, Web arşivlemede yaşanan sorunlar arasında üst sıralarda yer almaktadır. Seçim aşaması ayrıca kürasyonla birlikte, kütüphanecilerin ve arşivcilerin yaptıkları işin önemli bir parçasıdır.

Seçim politikası en başta alan adı derleme yaklaşımlarını etkiler. Alan adı koleksiyonları genellikle belirli bir ülkeyle ilişkili tüm Web sitelerini toplama girişimleridir. Alan adı derlemede kütüphaneler, belirli bir ülkenin alan adını taramaya odaklanır. Bu, yalnızca ulusal alan son ekiyle (örneğin .fr veya .pt) biten Web sitelerini değil, o ülkede farklı bir alan adı sonekiyle barındırılan Web sitelerini veya o ulusa yönelik içerik toplamaya odaklanan Web sitelerini de içerebilir (Pennock, 2013:10).

Günümüzde birçok ülkenin alan adı uzantısı kolaylıkla tespit edilebilmektedir. Fakat Amerika Birleşik Devletleri örneğinde bu durum zorlayıcıdır. Bunun nedeni bu ülkede içerik oluşturanlar tarafından üretilen tüm Web sitelerini kapsayıcı bir ana listenin oluşturulmamış olmasıdır. Ayrıca ABD için herhangi bir ülke alan adı uzantısı da bulunmadığından bu ülkede barındırılan tüm Web sitelerinin derlenmesi mümkün görünmemektedir (Grotke, 2011:17).

Yığın derlemede ise halka açık olan Web'in büyük bir bölümü arşivlenmeye çalışılır. Yığın derleme ile toplanan koleksiyonlar büyük boyutta ve kapsamlıdır. Koleksiyon ne kadar büyük ve karmaşık olursa, eksik olma ihtimali de o kadar yüksek olur.

Yığın derleme ve alan adı derlemede ayrıca büyük miktarda depolama alanına gereksinim duyulur. Ortaya çıkan çift kopyaların ortadan kaldırılması ve veri mükerrerliğinin önüne geçilmesi adına veri tekilleştirme uygulamalarından yararlanılmalıdır. Buna ek olarak, seçme ve depolama süreçlerinde yapılan tercihler, organizasyonlara değerli ve iyi kaynakları veya otomatik üretilen önemsiz ve istenmeyen içerikleri ayıklama konusunda izin verir (ISO, 2013:8).

Seçimli derlemede Web siteleri derleme kurumlarının toplama politikasıyla alakalı olarak belirli bir koleksiyona dâhil edilmek üzere seçilir. Seçim sonucunda oluşan koleksiyonlar belirli bir tema veya konuyla ilgili Web sitelerini bir araya getiren özel koleksiyonlar olarak sunulabilir. Tematik derleme ve olay eksenli derlemede ise belirli bir olaya karşılık olarak özel koleksiyonların üretildiği seçimli bir arşivleme türü olarak görülür.

Seçimli derleme, özellikle değerli olduğu düşünülen sitelerdeki kaynaklara odaklanır ve belirli bir toplama alanı içinde yakalamayı sağlar. Bu noktada bütünlüğü sağlama adına koleksiyona bozuk linkler de dâhil edildiğinden tüm dosyaların mevcut olmama ihtimali yüksektir. Bu nedenle alan adı derleme ile karşılaştırıldığında daha iyi sonuçların alınacağını söylemek zordur. Ayrıca seçimli derleme sürecinde kasıtlı veya kasıtsız olarak küratörler ve kütüphaneciler yanlılık gösterebilir. Bu durum seçimli koleksiyonun araştırma değerini düşürebilir.

3.6.3. Web Mimarisi ve Tarayıcı Sınırlılıkları

Büyük miktardaki bilgi şaşırtıcı bir hızda Web'e eklenmektedir (ISO, 2013:8). Arşivleme sürecinde Web tarayıcıları URL'lere ulaşır ve statik durumdaki içeriği yakalarlar. Mevcut URL tabanlı tarama metodu, Web formları ve sorgu arayüzlerini de içeren, içerik iletimini HTML protokolleri dışından akışlı medya ve sosyal medya üzerinden sunan, yaygın ifadeyle derin veya gizli Web olarak bilinen Web'in hızla artan büyük bir bölümünü toplamının gerisindedir (ISO, 2012a:8-9).

HTML'deki URL'lere açıkça başvurulmadığında, örneğin Flash veya Javascript sunumlarında veya kullanıcılarla oluşturulan bazı dinamik etkileşim sonuçlarında, arşivleme tarayıcıları bazen başvuru alan içeriği yakalamada başarısızlık gösterebilir (ISO, 2013:8). Bununla birlikte, topluluğun karşılaştığı en temel sorun, arşivleme organizasyonlarının sürekli gelişimine yanıt vermelerini ve ortaya çıktıklarında yeni içerik türlerini arşivleme yeteneğini geliştirmelerini gerektiren yeni formatlar, protokoller ve platformlar ile hızla değişen Web'tir (ISO, 2012a:9).

Web'in büyüklüğü ve dinamik yapısı düşünüldüğünde Web'in tamamının arşivlenmesi mümkün görünmemektedir. Bununla birlikte mevcut tarayıcı teknolojisi, günümüz Web teknolojisinin birkaç adım gerisindedir (Grotke, 2011:18). Web tarayıcılarının büyük bir bölümü şu anda akışlı medyayı veya kullanıcı girişi gerektiren veri tabanı içeriğini arşivleyememektedir. Veri tabanlarının arşivlenmelerindeki temel zorluk birçoğunun patentli olmasından kaynaklanır. Önerilen yaklaşım ise DeepArc [152] gibi bir araç kullanarak, patentli veri tabanlarındaki verileri XML gibi açık bir standart formata dönüştürmektir (Niu, 2012b). Ayrıca derin Web olarak da anılan, tarayıcıların görebilmelerinin ve erişebilmelerinin zor olduğu dinamik ve şifre korumalı içeriğin yakalanmasında da zorluklar yaşanmaktadır.

Tarayıcının ilerlemesini durduran veya önleyebilen diğer sorunlar, tarama boyutunun işlemsel sınırları (tarama kapsamı, keşfedilen ana bilgisayarları veya zamanlanmış URL'leri depolamak için kullanılabilen tarayıcı belleği miktarını aştığı zaman) ve tarayıcı tuzaklarıdır (örneğin başlangıç ve bitiş tarihleri sabit olmayan dinamik sayfa yapılı alışveriş sepetleri veya çevrimiçi takvimler) (Pennock, 2013:11). Gelecek nesil tarayıcılarda bu sınırlamaları gidermek için bu alanda çalışmaların geliştirilmesi beklenmektedir.

Taramayla ilgili bir başka yaygın sorun zamansal uyumsuzluktur. Zamansal uyumsuzluk, bir tarayıcının bir Web sitesini taramayı bitirmesi için gereken süre içinde, sitenin bazı bölümleri güncellendiğinde ve kaynak URL ile ona bağlı alt URL'ler zamansal olarak eşleşmediğinde ortaya çıkar (Pennock, 2013:13). Bu durum, orta ve

büyük boyutlu Web siteleri için olasıdır. Yığın veya alan adı düzeyinde bir derleme için, alanın boyutuna bağlı olarak tam bir taramanın tamamlanması birkaç gün veya birkaç hafta sürebileceğinden zamansal uyumsuzluk sorununun yaşanması olasıdır. Bu nedenle tarama sonucunda ortaya çıkan koleksiyon belirli bir günü değil periyodu yansıtır.

3.6.4. Sosyal Medya Arşivleme

Toplumumuz, bilgiyi oluşturma ve iletme biçimimizdeki radikal algı değişiklikleriyle karakterize edilmekte ve insanlar ve kurumlar gün geçtikçe sosyal ağlara daha fazla dâhil olmaktadır (Lasfargues ve diğerleri, 2012:121). Sosyal medya arşivleme konusuna bireyler açısından bakıldığında sosyal medyanın bireylerin Web varlığının önemli bir parçası olduğu söylenebilir. Sosyal medya konusuna kurumlar açısından bakıldığında ise, gelişen rekabet koşullarında ve hizmet yarışında kurumlarda sosyal medyanın sıklıkla kullanıldığı, bir pazarlama ve yayın aracı olarak görüldüğü ve sosyal medyadan elde edilen büyük verinin kaynak veri olarak görülerek ve analiz şirketleri aracılığıyla işlenerek kurum politikalarına yön verdiği aktarılmaktadır (Gökkurt ve Demirtel, 2014:226). Bununla birlikte sosyal medya içeriğinin kanıt değeri taşıması neticesinde arşivlenmesi gerektiği de söylenebilir. Fakat ister bireysel ister kurumsal amaçla kullanılsın sosyal medya hayatın her alanında yaygınlaştıkça kültürel miras kurumları için bu içeriği arşivleme konusu giderek zorlaşmaktadır.

Kullanıcılara anlık ve görsel açıdan çekici içerik sunmak için geliştirilen teknolojiler olan Twitter, Facebook, Flickr ve YouTube gibi sosyal medya sitelerinde yayınlanan içerik, genellikle Web arşivleme teknolojisi ile uyumsuz olduğundan, Web arşivleme için bir zorluk teşkil etmektedir (NA, 2015:12). Geçici ve dinamik doğası nedeniyle neyin korunacağına karar vermek zorlaşmakta ve durağan olmayan içeriği yayınlamak için özel teknolojilere ihtiyaç duyulmaktadır. Web arşivleme kurumları tarama parametreleri üzerine yönergeler yayınlayıp dolaştırarak bu sorunları ele almaktadır (Lasfargues ve diğerleri, 2012:121).

Sosyal medya arşivleme kuruluşlar arasında endişe duyulan alanların başında gelmektedir. ABD'deki Web arşivleme faaliyetlerine yönelik hazırlanan NDSA 2016 raporunda, ankete katılan kuruluşların %70'i büyük depolama alanı maliyetleri doğurabileceği gerekçesiyle sosyal medya arşivlemeden endişe duyduklarını belirtmiştir (Bailey, Grotke, McCain, Moffatt ve Taylor, 2017:15). Aynı ankette endişe duyulan diğer alanlar video arşivleme, interaktif medya ve veri tabanı arşivleme olarak sıralanmıştır. İçeriğe ilişkin kayıpların vurgulandığı bir başka çalışmada ise, sosyal medya içeriğinin %11-20 'sinin bir yıl içerisinde ve %27-41'inin iki veya üç yıl içerisinde ortadan kaybolduğu ifade edilmiştir (SalahEldeen ve Nelson, 2012:125).

Bir sosyal paylaşım sitesinin sınırlarını belirlemek ve onu arşivlemek oldukça zordur. Zor olmasının nedeni içerik için hem site sahibinden hem de katkıda bulunanların tamamından izin alınması gerekliliğidir. Bu noktada izinlere yönelik yeni problemler açığa çıkar ve arşivleme zaman alıcı ve olanaksız hale gelebilir. Fakat içerisinde genellikle değerli ve kimi durumlarda benzersiz içeriği barındırması nedeniyle sitenin arşivlenmesi önerilir. Ayrıca sosyal medya sitelerinde yer alan içeriğin henüz yönetilebilir bir büyüklükte olduğu ve arşivlemeye yönelik çözümlerin sağlıklı bir şekilde hayata geçirilmesinin olanaklı olduğu aktarılmaktadır (Gökkurt ve Demirtel, 2014:233). Wikiler ve bloglar gibi sosyal paylaşım siteleri ve video içeriği barındıran zengin medya paylaşım siteleri genellikle çok sayıda kullanıcı tarafından oluşturulan içeriği barındırırlar ve Web 2.0 Web siteleri olarak anılırlar. Geleneksel Web sitelerinin aksine kullanıcıların katkı ve yorumlarının yer aldığı Web 2.0 Web sitelerinin arşivlenmesindeki temel zorluk bu Web sitelerinin tarayıcılar için sorunlara neden olabilecek JavaScript uygulamaları bakımından oldukça zengin olmalarıdır.

Web 2.0 sitelerine ek olarak video içerikten oluşan zengin medyayı yakalama ayrı bir problem olarak karşımıza çıkar. Web videolarını yakalamak oldukça karmaşıktır ve gizlenen bağlantıların keşfi için farklı teknikler kullanılmalıdır (Lasfargues ve diğerleri, 2012:121). YouTube bu sorunun temsili bir örneğidir. Bu gibi durumlarda içeriğin dökümünü indiren ve bir dosyaya kaydeden harici bir araç kullanmak çözüm olabilir. Kapsam dâhilinde bazen kaynakların bir bölümü teknik sınırlamalardan dolayı Web arşivine eklenemeyebilir (ISO, 2013:7). Örneğin, multimedya ve interaktif kaynakları yakalama ve yeniden oynatma Web arşivleme topluluğu için önemli sorunlara neden olabilir ve sıklıkla pahalı, özelleştirilmiş çözümler gerektirebilir (ISO, 2013:7).

Sosyal medya içeriğini arşivleyen kurumlara bakıldığında ilk olarak Ulusal Arşivler (NA) göze çarpar. Ulusal Arşivler (NA) blog içeriğini birkaç yıldır arşivlemektedir ve 2013'ten itibaren bu, Twitter ve YouTube hesaplarını içerecek şekilde genişletilmiştir (NA, 2015:12). Her iki durumda da belirli bir hesap hedeflenmekte ve yalnızca hesabın üzerinde barındırılan içerik arşivlenmektedir (NA, 2015:12). Kurum arşivleme faaliyetlerine ek olarak kullanıcı gereksinimlerini karşılamak üzere özelleştirilmiş erişim arayüzleri de geliştirilmiştir. Sosyal medya arşivleme yapan durumlardan bir diğeri ise Hanzo Archives'dir. Geliştirdiği özel bir tarayıcı ile başarılı sonuçlar almaktadır.

3.6.5. Virüsler ve Zararlı Yazılımlar

Kötü amaçlı yazılım terimi, genellikle bir tarayıcı veya kelime işlemci gibi bilgisayar, işletim sistemi veya uygulama için tasarlanmış istenmeyen yazılımları açıklamada kullanılır (Pennock, 2013:13). Kapsamında virüsler, solucanlar, Truva atları, işletim sisteminde arka planda çalışan gizli program veya programlar (rootkit) [153],

aldatıcı reklam yazılımları, casus yazılımlar, tuş kaydediciler, suç yazılımları, çerez izleyicileri ve diğer kötü amaçlı yazılımlar yer alır.

Her birinin etkisi farklı olsa da kötü amaçlı yazılımların tamamı, antivirüs yazılımları tarafından virüs olarak nitelendirilmektedir (Pennock, 2013:14). Antivirüs yazılımlarının yanılabılme ve yanlış sonuçlar ortaya koyabilme ihtimali nedeniyle, bir antivirüs yazılımı taramasına dayanarak materyalin koleksiyondan çıkarılması önerilmez. Bu nedenle Web arşivlerinde kötü amaçlı yazılımları dışlama veya silme işlemleri sık görülmez. Bu noktada yapılması gereken arşiv kullanıcılarının kullandığı bilgisayar ortamında gereken koruma önlemlerinin alınmasıdır.

Kötü amaçlı yazılımlar konusunda söz edilmesi gereken bir diğer unsur Web spam içeriğidir. Arama motorlarını aldatmak için otomatik olarak üretilen Web içeriği olarak tanımlanan Web spam, günümüzde Web'in yaklaşık %20'sini oluşturan gerçek bir endüstri haline gelmiştir (Lasfargues ve diğerleri, 2012:121). Web spam Web arşivleme sürecinde kaynak israfına neden olmakta ve tarama işlemini uzatmakta ve içerik uyumluluğu açısından büyük bir zorluk oluşturmaktadır (Lasfargues ve diğerleri, 2012:121).

3.6.6. Veri Tekilleştirme

Tarayıcılar, Web derleme esnasında başvurulmuş URL'lerdeki dokümanları tekrarlı bir şekilde indirirler (Gomes, Santos ve Silva, 2006:1). Tarayıcının aynı dokümana sevk eden URL'leri bularak indirmesi veya aynı Web sitesinin tamamını veya bir bölümünü toplamasıyla çift kopyalar oluşur ve mükerrerlik meydana gelir. Bu durum kademeli tarama yapılarak oluşturulan Web arşivlerinde atık alanlar meydana getirir. Uygulama sonuçları, mükerrerliği ortadan kaldırılmanın depolama verimini arttırdığını göstermiştir (Gomes ve diğerleri, 2006:1).

Veri tekilleştirme kavramı, aynı içeriğin diğer tüm kopyalarını ortadan kaldırmayı ifade eder. Zaman içerisinde gerçekleştirilen farklı taramaların bir sonucu olarak oluşan içeriğin mükerrer kopyalarının ortadan kaldırılması iki açıdan önemlidir (Pennock, 2013:14):

- *Mükerrer içeriğin hacmi, tarama sıklığına bağlı olsa da koleksiyonlarda ortalama %25 civarındadır.*
- *Mükerrer içeriği desteklemek büyük ölçekli bir Web arşivi ve büyük miktarda depolama alanı gerektirir.*

Veri tekilleştirme bir dosyanın sadece bir kopyası kalana dek uygulanabilir. Bu durum depolama maliyetlerini azaltmaktadır. Bazı durumlarda ise daha az veri tekilleştirme yapılarak arşivde bir dosyanın birden fazla kopyasını bırakmak, ancak genel olarak tutulan kopya sayısını azaltmak yararlı olabilir (Pennock, 2013:14). Asgari düzeyde de olsa bir dosyanın birden fazla kopyasını bulundurmamak dosya bozulmaları karşısında iyi bir uygulama ve önlem olarak düşünülebilir. Veri tekilleştirmenin koruma gayesiyle ve işletme gerekçeleri nedeniyle istenmediği durumlar da vardır. Örneğin yasal süreçte kayıtları korumak ve kanıtları yok etme suçlamasından kaçınmak üzere veri tekilleştirmeden kaçınılabilir (Pennock, 2013:14).

Geniş bir belge kümesinin taranması sırasında mükerrer kopyaların yüklenmesinden kaçınılması zordur. Mükerrerlik bilgilerin erişilebilirliğini sağlamak için belgelerin kopyalanarak çoğaltılmasından, diğer sitelerin şablon olarak kullanılmasından veya erişilen belirli bir kaynak kullanılmadığında, aynı içeriğe sahip hata sayfalarını sunan Web sunucularından kaynaklanır (Gomes ve diğerleri, 2006:1). Dolayısıyla, belge mükerrerliği Web koleksiyonlarında yaygındır ve sorun depolama düzeyinde ortadan kaldırılmalıdır (Gomes ve diğerleri, 2006:1). Bu noktada, WARC arşivsel depolama formatının veri tekilleştirme uygulamalarını desteklediği söylenebilir (Pennock, 2013:14).

Artımlı olarak oluşturulmuş bir Web belge koleksiyonu, dokümanların değişmeden kalması nedeniyle mükerrerlik oluşturabilir. Periyodik taramaları sayfaların değişim sıklığını tahmin ederek gerçekleştirmek çözüm olmayabilir. Bunun nedeni birçok Web kaynağının geçmiş verilerinin mevcut olmaması ve URL'lerin kullanım ömrünün kısa oluşudur (Gomes, Miranda ve Costa, 2011:408).

Çift kopyaların yönetiminde URL kalıcılığı önemli bir faktördür. Araştırma makalelerinin temel alındığı URL'lerin kalıcılığı üzerine yapılan bir araştırma, dört yıl sonunda URL'lerin %40-50'sine erişilemeyeceğini göstermiştir (Gomes ve diğerleri, 2006:2). Yüksek Link Popülerliği Puanı (pagerank) [72] değerine sahip 151 milyon Web sayfasının örneklem olarak alındığı bir başka araştırmada ise geçen 11 haftada URL'lerin yaklaşık %10'u kullanılamaz hale gelmiştir (Fetterly, Manasse, Najork ve Wiener, 2003:4).

Arşivlenmiş belgelere referans vermede yaşanan URL kalıcılığı sorununun, tam nitelikli alan adı kaydı (FQDN) [73] sahipleri tarafından kaydedilen ve sürdürülen tekbiçimli kaynak adları (URNs) kullanılarak çözülebileceği düşünülmüştür (Daigle, van Gulik, Iannella ve Faltstrom, 2002; Gomes ve diğerleri, 2006:2). Devamında ulusal kütüphaneler, seçilen yayınlar için URN'leri korumaya çalışmış fakat insan müdahalesi gerekmesi nedeniyle kayda değer bir başarı sağlanamamıştır (Noronha, Campos, Gomes,

Silva ve Borbinha, 2001; aktaran Gomes ve diğerleri, 2006:2). Bu nedenle günümüzde veri tekilleştirmede URN varlığı önemli görülmemektedir.

Web'den toplanan kopya belgelerin görünümünden hangi URL'nin bir kopya olduğunu ve hangisinin orijinal olduğunu otomatik olarak tanımlamak mümkün değildir (Gomes ve diğerleri, 2006:2). Dolayısıyla, birden fazla URL aynı dokumana işaret ettiğinde, URL'lerden birini orijinal belge olarak değerlendirmek ve geri kalanlar kopyalar olarak görmek gerekir.

Web arşivlerinde dosya mükerrerliği tam ve kısmi kopyalar olarak iki grupta ele alınmaktadır. Eğer iki belge byte değeri olarak eşitse belgeler birbirinin tam kopyasıdır fakat eğer içeriğin bir bölümü aynı ise belgeler kısmi kopyalardır. Web arşivleri tam ve kısmi kopyalarla birlikte sahte kopyaları da barındırabilir. Teorik olarak, eğer iki belge aynı doküman parmak izine sahipse bunlar birbirinin kopyasıdır. Bununla birlikte, parmak izi algoritmaları, iki farklı belge için aynı imza üretilmesine neden olan küçük bir çarpışma olasılığı sunar (Gomes ve diğerleri, 2006:4). Bu durumda sahte kopyalar ortaya çıkmaktadır. Büyük bir belge koleksiyonunda mükerrerlikleri tespit etmek için yalnızca parmak izi algoritmalarını karşılaştırılmak belgelerin yanlışlıkla birbirinin kopyası olarak tanımlanmasına ve depolanmamasına neden olabilir (Gomes ve diğerleri, 2006:4). Sahte kopyalar çok büyük Web arşivlerinde görülmektedir ve kullanılan parmak izi algoritmasına bağlı olarak karşılaşılma olasılığı milyonda birden daha azdır (Gomes ve diğerleri, 2006:4). Disk hatası veya yazılımdaki hatalar nedeniyle bir belgenin kaybedilmesi çok daha olasıdır. Sahte kopyaların oluşmaması için ilk olarak byte cinsinden karşılaştırma yapılmalı ve ardından sonra doküman parmak izinden yararlanılmalıdır.

Dosya mükerrerliğini önlemek amacıyla ayrıca farklı tarayıcı teknolojilerinden yararlanılmaktadır. Mercator Tarayıcısı, aynı belgenin birden fazla kez işlenmesini önlemek ve indirilecek sayfaların sayısını azaltmak için, yinelenen sayfaları tanımlayan bir içerik izleme testi yürütür (Gomes ve diğerleri, 2006:7). Bu mekanizma, hafızada tutulan küçük bir karma tablo ve tek bir disk dosyasında tutulan büyük sıralı bir listeden oluşan iki bağımsız parmak izi kümesine dayanır (Heydon ve Najork, 1999:200).

3.7. Web Arşivlemede İstatistik ve Kalite Göstergeleri

İstatistikler ileri seviye analiz ve yorumlama için temel oluşturan nesnel verilerdir (ISO, 2013:16). Bu nedenle istatistik ve kalite göstergeleri Web arşivlerinin işletiminde oldukça önemli bir role sahiptir. Kütüphaneciler seçim ve kürasyon işlemlerinde, yöneticiler kaynak planlamasında ve faaliyetlerin başarı durumunu değerlendirmede, teknik personel ise sistem sürdürülebilirliğine ve planlamaya ilişkin aşamalarda istatistik ve kalite göstergelerinden yararlanmaktadır. İstatistik ve kalite göstergelerinin kullanımı ile ayrıca ileri seviye analiz ve yorumlama için tarafsız verilerin sağlanması amaçlanır. Böylelikle genel bir performans ölçümü ve hizmet karşılaştırması sağlanabilir.

Çalışmamızda Web arşivlemeye yönelik istatistikler Web arşiv koleksiyonun geliştirilmesine, nitelenmesine, kullanımına, korunmasına ve maliyetine yönelik olarak gruplandırılmış ve çeşitli kalite göstergeleri sunulmuştur. Bu bölümün hazırlanmasına ISO'nun 2013 yılında yayımladığı Web Arşivleme için İstatistik ve Kalite Konuları Teknik Raporu kaynaklık etmiştir [74].

Koleksiyon geliştirmeye yönelik istatistikler planlama ve izleme gibi adımlara katkı sağlamaktadır. Bunun için ilk olarak koleksiyonun boyutu hesaplanmalıdır. Koleksiyonun boyutu hesaplanırken yapılan taramalar sonucunda yakalamaların, URL'lerin, alan adlarının, sunucuların ve byte'ların sayımı gerçekleştirilir.

Koleksiyon nitelemeye yönelik istatistikler alınan küratöryal kararları destekler ve Web arşivinin niteliğini ve kapsamını belirleme adına önemlidir. Bir Web arşivi koleksiyonunu niteleme için kullanılan temel istatistikler üst seviye veya ikinci seviye alan adlarına göre dağılım, format türlerine göre dağılım, dil nitelendirmesi ve kronolojik kapsam olarak sıralanabilir.

Web arşivlerinde koleksiyon kullanımı genellikle ziyaretçilerin (fiziksel, sanal, robot ve istenmeyen ziyaretçiler) koleksiyon kullanım bilgilerinin toplanmasıyla ölçülmektedir. Arşivlenen Web kaynaklarını tanımlama ve onlara erişim koşulları ulusal mevzuat ve kuruluş politikalarına göre farklılaşmakta, ayırım beyaz, gri ve karanlık arşivler arasındaki en uygun olan kullanımı ölçme yöntemini seçmekle ortaya çıkmaktadır (ISO, 2013:28). Ayrıca bazı durumlarda Web arşivlerinin yasal koşullar veya kurum politikaları nedeniyle içeriği kontrollü coğrafi mekânlarda sunmaları gerekebilir.

Bir Web arşivinde yer alan dijital kaynakların korunması fiziksel ve mantıksal koruma olarak iki kategoride gerçekleştirilir. Fiziksel koruma veri güvenliğine yönelik

uygulamaları, mantıksal koruma ise göç ve öykünüm gibi stratejileri içermektedir. Fiziksel korumaya yönelik istatistikler, kaybolan veya bozulan kaynakların boyutu ile çoğaltılan ve aktarılan kaynakların boyutunu ölçmeyi gerektirir. Mantıksal korumada ise dosya formatlarına yönelik hesaplamalara yer verilir. Bu noktada, mevcut dosya formatları listelenmeli ve format başına düşen byte veya URL sayıları hesaplanmalıdır. Çoğaltılan, aktarılan ve dönüştürülen dosya formatlarına yönelik yüzdesel grafiklerin oluşturulması uzun süreli koruma ve depolama süreç planlamaları adına yararlı olacaktır. Web arşivlerinin bütünlüğünü göstermesi açısından kayıp veya bozuk veri miktarı da izlenmeli ve öykünüm ve göç gibi aktif koruma stratejilerine ilişkin uygulama verileri de tutulmalıdır.

Web arşivlemede maliyet kalemlerine bakıldığında donanım, yazılım ve işgücü maliyetleri öne çıkmaktadır. Donanım maliyetleri depolama, sistem alt yapısı ve bakım maliyetlerinden oluşur. Yazılım maliyetleri genellikle kullanılan yazılımların lisans ücretlerini ifade eder. Günümüzde birçok Web arşivinin uluslararası veya kar amacı gütmeyen kuruluşlar tarafından geliştirilen ve devamlılığı sağlanan ücretsiz, açık kaynaklı yazılımları kullandığı bilinmektedir. Açık kaynaklı yazılımların kullanımı lisans ücretlerini azaltsa da yazılımların kurumsal sistemlere entegrasyonu teknik uzmanlık (yazılım özelleştirmeleri veya güncellemeler) gerektirdiğinden dolayı olarak iş gücü maliyetleri artar. Web arşivlemede maliyet ayrıca, hizmet alımı olarak da ifade edilen dış kaynak kullanımı ve kurum kaynaklarının kullanımı olarak iki başlıkta ele alınabilir. Eğer hizmet alımı yapılıyorsa sağlayıcıya ödenen bedel doğrudan dış kaynak maliyeti olarak görülür. Dış kaynak kullanımı yoluyla alınan hizmetler genellikle veri derleme, dizinleme, erişim ve depolamayı kapsar fakat içeriği seçme, sözleşmeleri yürütme gibi ek maliyetler genellikle kurumun iç kaynaklarından karşılanır (ISO, 2013:35).

Kalite kendinde var olan bir dizi özellikle ihtiyaçları karşılama derecesi olarak tanımlanabilir (ISO, 2013:37). Kalite göstergeleri yönetim ve paydaşlar tarafından belirlenen hedeflerin bir Web arşivleme programı tarafından ne ölçüde yerine getirildiğini anlamaya yarayan göstergeleri içerir. Derleme kurumları kalite göstergeleri yardımıyla bir politika öncülünde hangi materyallerin arşiv kapsamına alındığını, istenilen materyallerin ne kadarının sağlandığını, kaynak kullanımının etkinliğini ve verimliliğini ve kaynakların erişilebilirliğini değerlendirebilmelidir. Web arşivlemede kullanılan kalite göstergeleri ve söz konusu göstergelerin hesaplanması için gerekli görülen tüm bileşenler Ek 2’de detaylı bir biçimde aktarılmaktadır. Bu göstergeler teknik ve küratöryal perspektiften Web arşivlemenin bugünkü durumunu anlamamıza yardımcı olacaktır. Web arşivlemede yararlanılan istatistik ve kalite göstergelerinin Web arşivleme iş akışlarıyla eşleştirilmesi amacıyla ISO tarafından oluşturulan İlişkili Performans Göstergeleriyle Birlikte Web Arşivleme Süreci Tablo 2’te sunulmuştur.

Tablo 2. İlişkili Performans Göstergeleriyle Birlikte Web Arşivleme Süreci (ISO, 2013:51)

İş Akışı	Aşama	İlgili İstatistik	Performans Göstergeleri
	Genel	-personel sayısı -maliyetler	-toplanan URL başına maliyet -Web arşivlemede çalışan personel (%)
	Seçim	Seçimli toplama: -izinler -seçilenler/aday gösterilenler	-sağlanan izinler (%)
	Tarama	Koleksiyon geliştirme ve niteleme: -hedefler -hedef örnekleri -URL -durum kodları -sunucular/alan adları -konteyner dosyalar -üst veya ikinci seviye domainler -alan adı başına düşen hacim -MIME türleri -tarihler	-belirli bir süreçte Web üzerinde kaybedilen veri (%) -zorunlu kapsamda arşivleme (%)
	Koruma	Koleksiyon koruma: -kopyalama -üstveri -tanımlı dosya formatı -koruma stratejisiyle tanımlı format	-kopya kaynaklar (%) -hasarlı, kayıp, erişilemez kaynaklar (%) -tanımlı dosya formatlı kaynaklar (%) - koruma stratejisiyle tanımlı format (%) -virüs taraması yapılmış kaynaklar (%)
	Kullanım	Koleksiyon erişilebilirliği ve kullanımı: -sayfa görüntüleme -ziyaret (oturum) -benzersiz ziyaretçiler -ziyaret süresi -ziyaret başı sayfa görüntüleme -site içerisinde kullanılan arama terimleri	-son kullanıcının erişebildiği kaynaklar (%) -tam metin indekslenen kaynaklar (%) -kataloglanan kaynaklar (%) -yıllık erişilen kaynak (%) -Web arşivini kullanan kütüphane kullanıcıları (%)

3.8. Web Arşivlerinde İşlevsellik

Web arşivi kullanıcıları için önemli olan işlevler temel arama ve taramadan gelişmiş, kişiselleştirilmiş ve özelleştirilmiş servislere, veri madenciliği ve Web sitesinin yeniden yapılandırılmasına kadar uzanır (Niu, 2012a). Fakat Web arşivleri tarafından bu işlevselliklerin desteklenmesine yönelik yapılan araştırmalar sınırlıdır. Bu noktada Web arşivlerinin hangi işlevleri desteklediğini belirlemek onları karşılaştırma adına yararlı olabilir. Bu noktada Web arşivleme yapacak kurumlara yönelik Uluslararası Internet Koruma Konsorsiyumu (IIPC) tarafından hazırlanan ve Ek 3’de sunulan Web Arşivleri için İşlevsellik Test Aracı’nın kullanımı yararlı olabilir. Liste aynı zamanda mevcut Web arşivleri için bir ölçüm veya kendi kendini değerlendirme aracı olarak da kullanılabilir.

Web arşivlerinde işlevsellik konusu genel olarak arşivin kullanımıyla ve ortaya çıkan ihtiyaçlarla ilgilenir. Arama parametreleri ve arama sonuçlarının listelenmesi, politika ile ilgili işlevler, kişiselleştirilmiş hizmetler, veri madenciliği ve kayıp Web sitelerinin yeniden yapılandırılması gibi konular işlevsellik ile ilgili çeşitli örneklerdir ve bu bölüm kapsamında aktarılacaktır.

3.8.1. Arama Parametreleri ve Sonuçların Listelenmesi

Arama parametreleri içerisinde en yaygın kullanıma sahip olan URL ile arama parametresidir. URL ile aramada yaygın olan uygulama, yakalanan URL ile yakalanma zamanı ve tarihinin birlikte sunulmasıdır. URL’nin tamamının veya bir bölümünün kullanımı durumunda karşılaşılan sonuçların farklı olması muhtemeldir.

Web arşivlerinde arama seçeneklerinin artırılması işlevselliği destekler. URL ile arama haricinde Web arşivlerinin birçoğunda tam metin arama, ulusal veya bölgesel alan adı arama ve anahtar sözcük ile arama parametreleri yaygın olarak kullanılmaktadır. Bu seçeneklere ek olarak konu araması, çoklu alan araması veya formata ve tarihe göre arama gibi işlevsellikler de Web arşivlerinde kullanıcılara sunulmaktadır.

Arşivlenen içeriğin gösterilmesi noktasında Web arşivleri tarafından, Web sayfasının orijinal URL’sine dayanarak arşivlenen Web içeriği için yeni URL’ler oluşturulmaktadır. Dikkatli bir şekilde bakıldığında canlı bir Web sayfasının URL’si ile bir Web arşiv sayfası URL’si birbirinden kolaylıkla ayırt edilebilir. Bununla birlikte bazı Web arşiv sayfalarının üst kısımlarında banner kullanıldığı görülmektedir. Kimi durumlarda sayfada canlı Web sayfası görünümü sağlama adına banner gizleme işlevselliği de sunulur. Kongre Kütüphanesi ve Birleşik Krallık Web Arşivleri’nin

kullanıcılarına sayfada yer alan banner'ı gizleme ve gerektiğinde yeniden aktif hale getirme işlevselliği sunması buna örnektir. Kongre Kütüphanesi Web Arşivleri'nde Banner Kullanımı Şekil 6'da gösterilmektedir.



Şekil 6. Kongre Kütüphanesi Web Arşivleri'nde Banner Kullanımı (Niu, 2012a).

Bir arama parametresi olarak yetişkinlere uygun içeriği filtreleme bir arama parametresi olarak düşünülebilir. Fakat güvenli arama parametresinin Web arşivlerinde kullanılmasına yönelik herhangi bir örnekle çalışmamızda henüz karşılaşılmamıştır. Seçimli derleme ile oluşturulan arşivlerde bu özelliğin desteklenmesi gereksiz olabilir fakat yığın derleme yapılan arşivlerde bu tarz içerik koleksiyon içerisinde barındırılmaktadır.

Arama parametreleri içerisinde erişim engellerini ortadan kaldırma da önemli bir parametredir ve bazı Web arşivlerinde diğerlerinde görülemeyen uygulamalara rastlanabilir. Örneğin, Birleşik Krallık Hükümet Web Arşivi, canlı Web siteleri ile entegredir ve istenen bir URL, Birleşik Krallık hükümetinin canlı Web sitesinde artık mevcut olmadığında, talep otomatik olarak Birleşik Krallık Hükümet Web Arşivi'ne yönlendirilir (Niu, 2012a).

3.8.2. Politikayla ilgili işlevsellikler

Politikayla ilgili işlevsellikler genelde kullanıcıların bir arşivde bulunan içeriğe erişimini engelleme, içeriği bir arşivden kaldırma veya bir arşive içerik ekleme ile ilgilidir. Faaliyetlerini sürdüren Web arşivlerinin tamamının tutarlı birer politika dokümanına sahip olma ihtimali zayıftır. Politikaların oluşturulmasını etkileyen temel

unsurlar telif ve mülkiyet hakları, yayıncılardan alınan izinler ve kanunun verdiği derleme ve erişime sunma yetkileridir.

3.8.3. Veri Madenciliği ve Kişiselleştirilmiş Hizmetler

Veri madenciliği geçmişin hikâyelerini açıklamaya ve verilerin gelişiminin analizi ve modellenmesi yoluyla gelecekteki olayları öngörmeye yardımcı olur (Costa ve diğerleri, 2016:6). Web arşivleri, bu gelişimin ortaya çıkarılması ve güçlendirilmesi için istisnai bir veri kaynağı olarak düşünülebilir. Fakat Web arşivleri büyük veri setlerini barındırmakla birlikte veri madenciliği uygulamalarının sayısının oldukça az olduğu görülür. Bunun nedeni veri madenciliği ve büyük veri analizi gibi işlevselliklerde gelişmiş bilgi keşfi ve analizi araçlarına ihtiyaç duyulmasıdır (Niu, 2012a). Yine de Web arşiv dosyalarının gelecekteki araştırmalara kaynaklık etme potansiyeli bugün için onları korunmaya değer kılmaktadır.

Uygulama sayısı pek fazla olmasa da gelecekte yaşanabilecek olayları tahmin etmek ve öngörü kazanabilmek adına Web arşivlerinde yer alan verilerin ayıklanması iyi bir yöntem olabilir. Örneğin, Radinsky ve Horvitz, gelecekte yaşanabilecek olayları tahmin edebilmek için veri madenciliğinden yararlanmış ve haberleri ve Web'i araştırdıkları çalışmada Angola'daki kuraklıklar ve fırtınalar arasında kolera salgınını katalize eden bir ilişki bulmuşlardır (Costa ve diğerleri, 2016:6).

Veri madenciliği işlevlerini sağlamasa da birçok geleneksel kütüphane tarafından katalog arama ekranında ve dijital kütüphane arayüzünde bazı kişiselleştirilmiş hizmetler sunulmaktadır. Nadir görülen uygulama örneklerine bakıldığında Yeni Zelanda Ulusal Kütüphanesi Kataloğu'nun kullanıcılar için arama ve kayıtları kaydetme gibi kişiselleştirilmiş kullanıcı hizmetlerini sağladığı ve Web arşiv kayıtlarına bu Katalog aracılığı ile erişilebildiği görülmüştür. Web arşivlerinin pek azında bir işlevsellik olarak arama ve sonuçları kaydetme gibi kişiselleştirilmiş hizmetlerin görülmesi nedeniyle Yeni Zelanda Ulusal Kütüphanesi Web arşivi arayüzü üzerinden bu hizmetin sunulması kayda değerdir. Avustralya Web Arşivi ise Web arşivine eklenen yeni başlıklarla ilgili aylık raporlar yayınlayan tek Web arşividir fakat bu hizmetin kişiselleştirilmiş bir hizmet olarak ele alınması zordur (Niu, 2012a). Son olarak, Birleşik Krallık Web Arşivinin, arşivlenen içeriğin incelenmesine dayanan iki görselleştirme arayüzü sunduğu söylenebilir: etiket bulutları ve arama teriminin nasıl oluştuğunu gösteren bir grafik olan N-gram arama [76].

3.8.4. Kayıp Web Sitelerinin Yeniden Yapılandırılması

Bazı Web arşivlerinde kaybedilen Web sitelerinin en azından bir bölümünün kurtarılması mümkündür. Bu aşamada kayıp Web sitelerinin kullanılması amacıyla üretilen WARRICK vb. yazılım araçlarından yararlanılabilir [151]. WARRICK aracının, IA Wayback Machine ve Google, MSN ve Yahoo önbelleklerinden yararlandığı bilinmektedir (Niu, 2012a).

Kayıp Web sitelerinin yeniden yapılandırılması noktasında herhangi bir adım atmadan önce tarama sürecindeki gecikmenin hesaplanması önerilmektedir. Örneğin, IA Wayback Machine için bir Web sitesinin yakalanmasından arşivde sunumuna kadar geçen süre ortalama 6-12 aydır [154].

3.9. Web Arşivlemenin Evrimi

World Wide Web'in wikiler, ansiklopediler, fotoğraflar, bloglar ve vbloglar gibi çeşitli medya türleri aracılığıyla her türlü bilgiyi yayınlayabilen dinamik ve özgür bir doğası vardır. Web arşivleri ise bir bölümü benzersiz ve kültürel açıdan değerli olan bu bilgileri korur. Web arşivlerinin kapsamı ise ulusal, bölgesel veya uluslararası olmak üzere barındırdığı içeriğe göre değişmektedir. Örneğin IA küresel Web'in arşivlenmesini amaçlarken, Cacak Web Arşivi, bu Sırp şehri ile ilgili siteleri korumayı amaçlar (Costa ve diğerleri, 2016:8).

Web arşivleme alanında yaşanan bilgi eksikliği, Web'in kısa ömürlü doğası ve canlı Web'deki verinin büyük bir bölümünün henüz arşivlenememiş olması yerel, ulusal veya evrensel nitelikteki Web mirasının günden güne kaybolmasına sebep olmaktadır. Web arama teknolojisinin yeterli olmaması, mevcut arama motorlarının Web sitelerinin tamamına erişememesi ve benzer şekilde arşiv tarayıcılarının da istenilen performansa ulaşamaması da bu kayıplara gerekçe olarak gösterilebilir. Ayrıca birçok resmî Web sitesi yalnızca kendisini barındıran ülkenin ana dilindedir.

Web mirasını koruma amacıyla konsorsiyumlar, ulusal kütüphaneler, ulusal arşivler, üniversiteler ve özel şirketler onlarca hatta yüzlerce yıllık bilgiyi kapsayan koleksiyonlara ev sahipliği yapmaktadır. 2014 yılı verilerine göre dünya genelinde 33 ülkede en az 68 Web arşivleme girişimi sürdürülmekte ve yaklaşık 17 petabyte boyutunda 534 milyardan fazla dosya Web arşivlerinde saklanmaktadır (Costa ve diğerleri, 2016:3). Bununla birlikte, 1996 yılından beri sürdürülen ortak çabalara rağmen Web arşivleri ve sundukları hizmetler hakkında bilgi eksikliği bulunmaktadır. Web arşivlemenin tarihini, mevcut durumunu, güçlü yanlarını, zayıf yönlerini, zorluklarını, yönelimlerini ve

ihtiyaçlarını bilmeden bu büyük ölçekli bilgi yığınlarını yönetmek ve muhafaza etmek mümkün görünmemektedir.

Şekil 7 'de sunulan tarihteki ilk Web sitesi, 1991 yılında Tim Berners-Lee tarafından Avrupa Nükleer Araştırmalar Örgütü (CERN) organizasyonunda oluşturulmuştur [155]. CERN'de bulunan bu en eski sürüm, dünya çapında Web projesini tanımlamaktadır.

World Wide Web

The WorldWideWeb (W3) is a wide-area [hypermedia](#) information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an [executive summary](#) of the project, [Mailing lists](#) , [Policy](#) , November's [W3 news](#) , [Frequently Asked Questions](#) .

[What's out there?](#)

Pointers to the world's online information, [subjects](#) , [W3 servers](#), etc.

[Help](#)

on the browser you are using

[Software Products](#)

A list of W3 project components and their current state. (e.g. [Line Mode](#) ,X11 [Viola](#) , [NeXTStep](#) , [Servers](#) , [Tools](#) , [Mail robot](#) , [Library](#))

[Technical](#)

Details of protocols, formats, program internals etc

[Bibliography](#)

Paper documentation on W3 and references.

[People](#)

A list of some people involved in the project.

[History](#)

A summary of the history of the project.

[How can I help ?](#)

If you would like to support the web..

[Getting code](#)

Getting the code by [anonymous FTP](#) , etc.

Şekil 7. Tarihteki İlk Web Sitesi

İlk Web arşivleme girişimi 1996 yılında faaliyete başlayan, ABD merkezli ve kar amacı gütmeyen bir kuruluş olan Internet Archive'dır. Internet Archive'ın ardından Web arşivleme faaliyetlerine başlayan Avustralya'daki Pandora ve Tasmanian Web Arşivleri ve İsveç'ten Kulturarw3 Web arşivi de öncü oluşumlar olarak sıralanabilir. O yıllardan itibaren pek çok araştırma kuruluşu ve bağımsız girişim tarafından Web arşivleme için açık kaynaklı araçlar, standartlar ve en iyi uygulamaların geliştirilmesine yönelik yoğun bir çaba sarf edilmiştir.

Web arşivleme alanında araştırma kuruluşları ve bağımsız girişimlerin ortaya çıkmasındaki nedenler, bir Web arşivinin oluşturulmasının ve işletilmesinin karmaşık ve masraflı olması ve dijital içerik koleksiyonlarının toplanması, oluşturulması ve korunması amacıyla arşivleme hizmetlerine gereksinim duyulmasıdır. Kurumlar adına Web arşiv hizmetleri sağlayan ve hatta bilgi keşfi yapan bu girişimlerin bir bölümü ticaridir ve Web

Arşiv Servisleri (WAS) olarak adlandırılabilir. 2014 yılı verilerine göre dünya genelinde Web arşivleme girişimlerinin en az %19'u WAS kullanmaktadır (Costa ve diğerleri, 2016:9). Tarihteki ilk WAS örneklerine Archive-It [77], ArchiveTheNet [78] ve Web Archiving Service [79] ve güncel WAS örneklerine Aleph Archives [80], Hanzo Archives [81] ve Reed Archives [82] örnek verilebilir.

Web arşivlerinin coğrafi bölgelere göre dağılımına bakıldığında genellikle gelişmiş ülkelerde Web arşivleme faaliyetlerinde bulunulduğu söylenebilir. 2014 yılı verilerine göre dünya genelinde Web arşivleme yapan 68 kurum 33 ülkeye dağılmıştır ve Web arşivleme yapılan ülkelerin coğrafi dağılımı Şekil 8'de göstermektedir (Costa ve diğerleri, 2016:10):



Şekil 8. Web Arşivleme Yapılan Ülkelerin Coğrafi Dağılımı
(Costa ve diğerleri, 2016:10)

Web arşivlerinin evriminde uygulanan seçim politikası, derleme stratejisi, arama yöntemleri, yasal dayanak, yazılımların belirleyici olduğu söylenebilir. Web arşivlemede seçim politikaları kurumdan kuruma farklılık göstermekte ve tercih edilen tarama stratejisini etkilemektedir. Kaynakların sınırlı ve Web'in tamamının arşivlenemeyecek olması kurumların çoğunu seçimli ve alan adı derleme stratejilerini uygulamaya yönlendirmiştir. 2014 yılı verilerine göre dünya genelindeki Web arşivlerinde %66 oranında seçimli ve %29 oranında alan adı veya yığın derleme yapılmaktadır ve bu oran Avrupa Web arşivlerinde %71'i oranında tematik veya seçimli ve %23 oranında alan adı derleme ve şeklindedir (Costa ve diğerleri, 2016:11). Seçim politikasını etkileyen diğer

unsurlar ise yazarlardan alınan izinler ve arşivlenecek içeriğin kurumun hedefleriyle örtüşmesidir.

Bir Web arşivinde kaynakların sunumunda kullanılan arama yönetimi kadar kurumun Web arşivlemeye yönelik yasal dayanağa sahip olup olmayışı da belirleyicidir. Arşivlenen içeriğin telif haklarına sahip olma durumu veya arşivleme öncesinde hak sahiplerinin izinlerinin alınması yaşanabilecek sorunların önüne geçilmesini engeller.

Web arşivlerinin evriminde değinilecek son konu yazılımlar ve dosya formatlarıdır. 2014 verilerine göre Avrupa genelinde en sık kullanılan yazılım araçları bir sağlama aracı olan Heritrix ve birer arama acı olan NutchWAX ve Wayback Machine'dir (Costa ve diğerleri, 2016:15). Arşivlenen verilerin depolandığı dosya formatlarına bakıldığında ise WARC, ARC ve HTTPRack en yaygın kullanılan dosya formatları olarak sıralanabilir.

3.10. Web Arşivleme Girişimleri

Bu bölümde Web arşivleme alanında gerçekleştirilen araştırma projelerine ve girişimde bulunan öncü kuruluş örneklerine yer verilecektir.

3.10.1. Web Arşivlemeye Dayalı Araştırma Projeleri

Web arşivleme konusunda geçmişte birçok araştırma projeleri yürütülmüştür. Bu çalışmada ele alınan araştırma projeleri ise LIWA, LAWA, WebART, ARCOMEM, Memento, NEDLIB ve SCAPE olarak sıralanabilir.

The Living Web Archives (Yaşayan Web Arşivleri -LiWA) Projesi'nde geniş bir zaman periyodu içerisinde taranan Web arşiv verilerini toplamak, sorgulamak ve analiz etmek için araçlar geliştirmek amaçlanmıştır. LIWA ayrıca Web spam algılama, terminoloji gelişimi, akış videolarını yakalama ve arşivlenmiş içeriğin zamansal tutarlılığını sağlama gibi arşivlenmiş bilgileri erişilebilir hale getirmeye ve bilgi erişim zorluklarını çözmeye katkı sağlamayı amaçlamıştır (Costa ve diğerleri, 2016:7). 2008-2011 yılları arasında sürdürülen proje vasıtasıyla Web içeriğini arşivlemek için yeni nesil araçların keşfedilmesi amaçlanmıştır.

Longitudinal Analytics of Web Archives (Web Arşivlerinin Boylmasına Analizi-LAWA) ile özellikle heterojen Internet verilerini büyük ölçekte analiz etmek için araçlar geliřtirmek amaçlanmıřtır (Pennock, 2013:6). Proje 2010– 2013 yıl aralıęında sürdürölmüřtür [83].

WebART Projesi özellikle beřerî bilimler ve sosyal bilimler alanlarında arařtırmayı kolaylařtıracak řekilde tasarlanmıř Web arřiv eriřim araçlarının geliřtirilmesine odaklanmıřtır (Costa ve dięerleri, 2016:7; Huurdeman, Ben David ve Samar, 2013).

The Collect all ARchives to COmmunity MEMories (Toplumsal Hafıza için Tüm Arřivleri Toplama- ARCOMEM) Projesi, sosyal Web'in korunması ve geliřtirilmesine yardımcı olmak için yenilikçi araçlar ve yöntemler geliřtirmekle ilgilidir. 2011-2014 yılları arasında sürdürölen Proje özellikle sosyal aęlara odaklanarak sosyal Web'in arřivlenmesi ve korunmasına yönelmiřtir (Pennock, 2013:6). Proje, hedeflenen Web uygulamasının türüne baęlı olarak taramanın davranıřını deęiřtirerek ve içerik analizi modölleri ile etkileřim kurarak karmařık nesnelerin Web'den çıkarılmasına olanak saęlayacak uygulama farkındalıęı olan tarayıcıların geliřtirilmesini hedeflemiřtir (Lasfargues ve dięerleri, 2012:122).

Memento Projesi HTTP protokolüne zamansal bir boyut katmaktadır, böylece bir belgenin arřivlenmiř versiyonları, Web sunucusu istenen versiyonları içermiyorsa o dokümanı tutan Web sunucusu veya mevcut Web arřivleri tarafından sunulabilir (Nelson, Sanderson, Balakireva ve Shankar, 2009). Buna adapte olabilmek için basit bir çözüm olarak kullanıcıların yalnızca bir tarayıcı eklentisi yüklemesi gerekmektedir (Costa ve dięerleri, 2016:7). Time Travel Portal [84] eklentisi sayesinde kullanıcılar birden çok Web arřivinde arama yapabilmektedir. Time Travel Portal, Web arřivleri için bir meta arama motorudur.

Aę Tabanlı Avrupa Derleme Kütüphanesi Projesi (Networked European Deposit Library Project- NEDLIB), Avrupa Komisyonu Telematik Uygulamalar Programı desteęinde ve Holanda, Fransa, Norveç, Finlandiya, Almanya, Portekiz, İsveç, İtalya öncölüęünde bařlatılmıřtır (Oęuz, 2006b:42). Proje elektronik yayınların uzun süreli kullanımını destekler ve yönetiminde ortak hareket etmeyi amaçlar.

SCAPE Projesi ise büyük ölçekli dijital koruma çözümelerini ele almıřtır (Costa ve dięerleri, 2016:7). 2011-2014 yıl aralıęında sürdürölen Proje, Web arřivleri de dâhil olmak üzere karmařık heterojen ve geniř ölçekli dijital içerik için ölçeklenebilir uzun vadeli koruma hizmetleri geliřtirmeyi amaçlanmıřtır (Pennock, 2013:6). SCAPE ile ayrıca,

Web arşivlerine özgü koruma araçları ve sorunlarını test etmek için bir Web içeriği sinama ortamı hazırlanmıştır.

Yukarıda sözü edilen projelere ek olarak bazı kurumlar semantik arama, akıllı keşif aracı geliştirme ve konusal analiz gibi alanlara ilgi göstermiştir. Örnek olarak, Birleşik Krallık Ulusal Arşivleri, arşivlenmiş Birleşik Krallık Hükümeti Web sitelerine yönelik aramaları iyileştirmek için "akıllı bir keşif aracına" yönelik yatırım yapmıştır (Lasfargues ve diğerleri, 2012:123). Bu yatırımın amacı, semantik arama yoluyla içeriğe ilişkin çok daha zengin ve daha sofistike bilgi sağlayarak Web arşivinde bulunan bilgilere erişimi geliştirmektir. Bununla birlikte, Internet Bellek Vakfı (IMF) da istatistiksel tanımlayıcıların çok büyük veri kümelerinden türetilmesi ile ilgilenen kurumlar için genişletilmiş analiz ve çıkarma mekanizmaları sağlamak üzere çalışmalar yürütmüştür (Lasfargues ve diğerleri, 2012:123). Araştırılan analiz ve çıkarma yöntemleri arasında varlık tanıma, dil algılama ve belgeler arasındaki bağlantılar yer almaktadır.

3.10.2. Kuruluşlar

Uluslararası Internet Koruma Konsorsiyumu (IIPC)

Uluslararası Internet Koruma Konsorsiyumu (IIPC) kültürel miras ve araştırmalar için Web arşivlerinin kullanımını ve onlara erişimi sağlama adına uluslararası iş birliğini teşvik eden ve kendini Web arşivleme araç, standart ve uygulamalarını geliştirmeye adanmış bir üyelik örgütüdür [85]. Örgüt, 45'in üzerinde ülkeden ulusal, akademik ve üniversite kütüphaneleri ve arşivlerden oluşur ve üyelik koşullarını karşılayan tüm kütüphanelere, arşivlere, müzelere ve kültürel miras kurumlarına açıktır. Uluslararası Internet Koruma Konsorsiyumu'nun (IIPC) kurucu üyeleri İngiltere Milli Kütüphanesi, Fransa Milli Kütüphanesi, Kongre Kütüphanesi, Avustralya Milli Kütüphanesi ve İsveç Milli Kütüphanesi'dir (Pennock, 2013:5).

Yıllık olarak üyelik başvurularını kabul eden IIPC Internet'teki bilginin sağlanması, korunması ve erişilebilir kılınmasını amaçlar. Bu amaçla IIPC çalışma grupları Web arşivleme araç, teknik ve standartlarının geliştirilmesi, üstveri üretiminin ve kullanımın teşvik edilmesi, AR-GE faaliyetlerinin ve kalite güvencelerinin sürdürülmesi gibi alanlarda faaliyetlerini sürdürmektedir [85]. IIPC ayrıca arşivleme ve koruma konularında kütüphane, arşiv ve araştırma kurumlarını destekler ve cesaretlendirir.

Birleşik Krallık Web Arşivleme Konsorsiyumu (UKWAC)

Birleşik Krallık Web Arşivleme Konsorsiyumu (UKWAC) altı kuruluşun ortak çabasıyla kurulmuştur. Bunlar İngiltere Milli Kütüphanesi (British Library), Galler Milli Kütüphanesi, İskoçya Milli Kütüphanesi, Ulusal Arşivler (National Archives), JISC (Joint Information Systems Committee of the higher education Councils) ve Wellcome Trust adlı bağımsız yardım kuruluşudur. Girişim seçimli derleme yaklaşımını benimsemiştir ve ilk yıllarda binlerce Web sitesi toplanarak arşivlenmiştir. Ayrıca Web arşivlemeye yönelik kurulum, yazılım, donanım ve yaşıatım masrafları hak sahipleri arasında paylaştırılmıştır (Oğuz, 2006b:42). Girişimin 2008 yılında dağılmasının ardından İngiltere'de ortak Web arşivleme çalışmaları gündeme gelmiştir ve 2009'dan bu yana çalışmalar Dijital Koruma Koalisyonu (DPC) tarafından koordine edilmektedir (Pennock, 2013:6).

Kuzey Ülkeleri Web Arşivi (NWA)

Kuzey Ülkeleri Web Arşivi (NWA) Norveç, İsveç, İzlanda, Finlandiya, Danimarka tarafından kurulmuştur. Norveç Milli Kütüphanesinin yönetici olduğu girişim artık devam etmemekte ve üye ülkeler faaliyetlerini IIPC çatısı altında sürdürmektedir.

Internet Archive (IA)

Internet Archive küresel Web'i korumayı amaçlayan ve üye kuruluşların dijital koleksiyonlarını depolama, kurma ve yönetme konusunda yardım sağlayan aboneliğe dayalı bir Web arşivleme servisi. 1996'da açılan IA koleksiyonunu ilk kez 2001 yılında kullanıcılarla paylaşmıştır. Herhangi bir koleksiyon geliştirme politikası olmayan girişim URL ve tam metin aramayı destekler (Pennock, 2013:21). WARC ve Heritrix gibi teknolojilerin üretiminde başrol oynayan girişim IIPC çatısı altında faaliyetlerine devam etmektedir. Archive-It kuruluşun Web arşivleme servisi.

Internet Bellek Vakfı (IMF)

Internet Bellek Vakfı (IMF), 2004 yılında, Web arşivleme girişimlerini desteklemek ve Avrupa'da Web koruması için destek sağlamak amacıyla kar amacı gütmeyen bir kuruluş olarak kuruldu (Pennock, 2013:28). WAS statüsünde olan Girişim müşterilerine Web arşivleme çözümleri sunmayı hedeflemektedir.

ArchiveNet, Internet Bellek Vakfı (IMF) tarafından sađlanan Web tabanlı bir Web arşivleme hizmetidir. Tarama esnasında dinamik ve zengin medya da dâhil olmak üzere farklı medya türlerini yakalamayı destekler. Hizmet tam metin aramayla birlikte canlı Web siteleri için otomatik yeniden yönlendirme işlevselliğine de sahiptir. Bu işlevsellikle talep edilen kaynak artık çevrimiçi değilse, kullanıcı otomatik olarak Web arşivinde yer alan kaynağın son sürümüne yönlendirilir. Hizmet ayrıca İngilizce ve Fransızca dil desteğı sunar. Hizmetin diğere öne çıkan özellikleri bağlantı ömrünü uzatma, kullanıcı deneyimini artırma, çevrimiçi görünürlüğü ve sıralamayı geliştirme olarak sıralanabilir.

Hanzo Archives (Hanzo Arşivleri)

Hanzo Arşivleri elektronik keşif, bulma ve belge yönetimi gibi ticari işletme ihtiyaçlarını karşılamak için Web sitesi arşivleme hizmeti sunmaktadır [81]. WAS statüsünde olan kuruluş, geliştirdiğı Web tarayıcısını kullanarak dinamik içerik ve sosyal ve zengin medya gibi standart bir tarayıcı için toplanması zor olabilecek içerikleri yakalamaktadır.

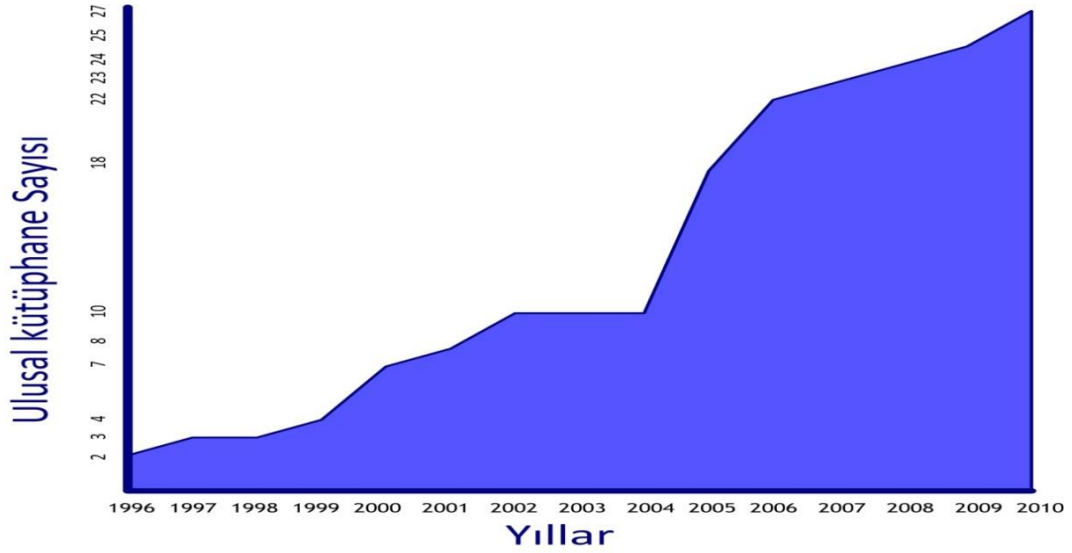
3.11. Ulusal Kütüphanelerin Web Arşivleme Uygulamaları

Bu bölümde Kuzey Amerika, Avrupa, Asya ve Okyanusya Kıta'larından toplam 27 Ulusal Kütüphane'nin Web arşivleme uygulamaları değerlendirilecektir. Güney Amerika ve Afrika Kıta'larında Web arşivleme yapan ulusal kütüphane bulunmadığı için bu coğrafyalarda yer alan ulusal kütüphanelere bölüm kapsamında yer verilmemiştir. Söz konusu ulusal kütüphanelerin birçoğı IIPC üyesi olduğundan bölümün hazırlanmasında diğere kaynaklarla birlikte büyük ölçüde bu kuruluşun Web sitesinden [86] istifade edilmiştir. Çalışma kapsamında ele alınan ve web arşivleme uygulamalarına yer veren 27 ulusal kütüphanenin coğrafi dağılımı Şekil 9'da sunulmaktadır.



Şekil 9. Web Arşivleme Uygulamalarına Yer Veren Ulusal Kütüphanelerin Coğrafi Dağılımı

Web arşivleme faaliyetinde bulunan ulusal kütüphanelerin uygulamalarını değerlendirilmeye başlamadan önce, söz konusu ulusal kütüphanelerin Web arşivleme programlarına başlama yıllarının ve yıllar içindeki sayısal artışlarının sunulması yararlı olacaktır. Şekil 10'da Web arşivleme uygulamalarında bulunan ulusal kütüphanelerin programlarına başlama yılları ve sayısal artışları sunulmaktadır.



Şekil 10. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Programlarına Başlama Yılları ve Sayısal Artışları

3.11.1. Kuzey Amerika Kıtası Ulusal Kütüphane Web Arşivleri

Kanada Milli Kütüphanesi Web Arşivleme Programı

Kanada Milli Kütüphanesi Web arşivleme programı (Government of Canada Web Archive/ Archives du Web du gouvernement du Canada) [87] 2005 yılında faaliyete girmiştir. 2004 yılında yürürlüğe giren Kanada Kütüphane ve Arşivler Yasası'ndan alınan yetki ile Kanada Milli Kütüphanesi Web sitelerini toplamaktadır. Toplayıcı Kanada Kütüphane ve Arşivleri'dir (Library and Archives Canada / Bibliothèque et Archives Canada). Web arşivleme amacıyla taramalar 6 ayda 1 kez gerçekleştirilir. Web arşivinin arayüz dili İngilizce ve Fransızca'dır. Erişim yöntemi URL arama, anahtar kelime arama, alfabetik arama ve tam metin aramadır. Derleme stratejisi ise bölgesel alan adı derlemedir. Kurum IIPC üyesidir.

Kongre Kütüphanesi Web Arşivleme Programı (ABD)

Kongre Kütüphanesi Web Arşivleme Programı [88] 2000 yılında başlatılmıştır. Koleksiyon belirli bir konuda Web temelli bilgileri sunmak üzere konu uzmanları tarafından seçilen ve arşivlenen Web sitelerinden oluşur. Web arşivleme faaliyeti gelecek nesil araştırmacılar için dijital materyalleri değerlendirmek, seçmek, toplamak, kataloglamak, erişmek ve bunları korumak için kütüphane tarafından sürdürülen çabanın bir parçasıdır. Kongre Kütüphanesi, URL atayan kişiler tarafından oluşturulan üstveriler temelinde otomatik olarak Metadata Nesne Tanımlama Şeması (MODS) kayıtları üretir ve daha sonra kayıtlar kataloglayıcılar tarafından geliştirir (Grotke ve Jones, 2010:22). Web arşivinin arayüz dili İngilizce'dir. Erişim yöntemi URL arama, alfabetik arama ve konu aramadır. Derleme stratejisi seçimli, olay eksenli ve tematik derlemedir. Olay eksenli ve tematik derleme sonucunda oluşturulan koleksiyon sayısı 40'ın üzerindedir (Grotke, 2011:17). Kurum IIPC üyesidir.

3.11.2. Avrupa Kıtası Ulusal Kütüphane Web Arşivleri

İspanya Milli Kütüphanesi Web Arşivleme Programı

İspanya Milli Kütüphanesi Web arşivleme programı (Archivo de la Web Española) [89] 2009 yılında İspanya Milli Kütüphanesi [90] ve bölgesel kütüphanelerin iş birliğiyle başlatılmıştır. Toplayıcı kurum İspanya Milli Kütüphanesidir (Biblioteca Nacional de España). Web arşivinin arayüz dili İspanyolca ve İngilizcedir. Erişim yöntemi URL aramadır. Derleme stratejileri genel ve ulusal alan adı, seçimli ve olay eksenli derlemedir. Kurum IIPC üyesidir.

Fransa Milli Kütüphanesi Web Arşivleme Programı

Fransa Milli Kütüphanesi Web Arşivleme Programı (Bibliothèque nationale de France- Archives de l'Internet) [91] 2002 yılında başlatılmıştır. Toplayıcı kurum Fransa Milli Kütüphanesi'dir (Bibliothèque Nationale de France) [92]. Program dâhilinde ilk olarak ulusal seçimlere yönelik Web siteleri toplanmış ve ardından IA ile iş birliği yapılarak tarihi koleksiyonlar için taramalar gerçekleştirilmiştir. Koleksiyon hacmi 2015 rakamlarına göre 26 milyar dosyaya ve 668 TB veri depolama alanına ulaşmıştır. Erişim yöntemleri URL arama, anahtar kelime arama, tam metin arama ve konu arama olarak sıralanabilir. Yasal sınırlamalar nedeniyle Fransa Milli Kütüphanesi Web Arşivi'ne yalnızca araştırmacılar tarafından ve Paris'te bulunan kütüphane binasında göz atılabilir. Web arşivinin arayüz dili Fransızca'dır. Derleme stratejileri ulusal alan adı, yığın, seçimli ve olay eksenli derlemedir. Kurum IIPC üyesidir.

Estonya Milli Kütüphanesi Web Arşivleme Programı

Estonya Milli Kütüphanesi Web Arşivleme Programı [93] 2010 yılında başlatılmıştır. Toplayıcı kurum Estonya Milli Kütüphanesidir [94] (Eesti Rahvusraamatukogu). Kurum 2006 yılında yayımlanan Derleme Kanunu'ndan aldığı izinle Eston Web sitelerini toplamakta ve arşivlemektedir. Web siteleri 2013 yılından beri kamuya açıktır. Ülkede Web sitesi sahiplerinin başvuru yaparak erişim sınırlama hakkı bulunsa da erişime kapatılan site araştırmacılar için hala Estonya Milli Kütüphanesinde yerinde erişime açık kalmaktadır. Erişim yöntemleri URL arama, anahtar kelime arama ve konu aramadır. Web arşivinin arayüz dili Estonca'dır. Derleme stratejileri seçimli, olay eksenli ve tematik derlemedir. Kurum IIPC üyesidir.

Hırvatistan Milli Kütüphanesi Web Arşivleme Programı

Hırvatistan Milli Kütüphanesi Web Arşivleme Programı (Hrvatski arhiv weba) [95] 2010 yılında başlatılmıştır. Toplayıcı kurum Hırvatistan Ulusal ve Üniversite Kütüphanesidir [96] (National and University Library of Croatia). Arşivi'nin görevi Hırvatistan'ın ulusal mirasının bir parçası olarak Hırvat Web kaynaklarını toplamak, korumak ve kalıcı olarak erişilebilir hale getirmektir. Arşivlenen kaynaklar 2011 yılından itibaren erişime sunulmaktadır. Arşiv'e kamusal erişim sağlanabilmektedir. Yayıncı kısıtlamaları olan Web kaynaklarına ise aynı anda yalnızca tek bir kullanıcı tarafından yerinde erişim sağlanabilmektedir Erişim yöntemleri URL arama, anahtar kelime arama, gelişmiş arama, alfabetik arama ve konu aramadır. Web arşivinin arayüz dili Hırvatça ve İngilizce'dir. Derleme stratejileri ulusal alan adı, seçimli ve tematik derlemedir. Kurum IIPC üyesidir.

İzlanda Milli Kütüphanesi Web Arşivleme Programı

İzlanda Milli Kütüphanesi Web Arşivleme Programı (Íslenska vefsafnið) [97] 1996 yılında başlatılmıştır. Toplayıcı kurum İzlanda Milli ve Üniversite Kütüphanesidir [98] (Landsbókasafn Íslands Háskólabókasafn). Arşiv tüm dünyaya açıktır. Fakat hak sahipleri erişim ücreti talep ettiğinde veya engelleme istediğinde Web siteleri erişime kapatılabilmektedir. Web arşivinin arayüz dili İzlandaca ve sınırlı İngilizce'dir. Erişim yöntemi URL aramadır. Derleme stratejileri ulusal alan adı, seçimli ve olay eksenli derlemedir. “.is” alan adı için tarama sıklığı yılda 3 kez olarak belirlenmiştir. Kurum IIPC üyesidir.

Finlandiya Milli Kütüphanesi Web Arşivleme Programı

Finlandiya Milli Kütüphanesi Web Arşivleme Programı (Kansalliskirjaston verkkoarkisto) [99] 2006 yılında başlatılmıştır. Toplayıcı kurum Finlandiya Milli Kütüphanesi'dir [100] (Kansalliskirjasto). Koleksiyon hacmi 2015 yılı verilerine göre 80 TB civarındadır. Arşivi herkes kullanabilir fakat arşive yalnızca seçilen kütüphanelerden erişim sağlanabilir ve dijital dokümanların kopyalanması yasaktır. Web arşivinin arayüz dilleri Fince, İsveççe ve İngilizce'dir. Erişim yöntemleri URL arama ve tam metin aramadır. Derleme stratejileri ulusal alan adı, bölgesel alan adı ve tematik derlemedir. Toplayıcı kurum tarafından “.fin” uzantılı gazeteler ve haber siteleri günlük olarak derlenmektedir. Kurum IIPC üyesidir.

İsveç Milli Kütüphanesi Web Arşivleme Programı

İsveç Milli Kütüphanesi Web Arşivleme Programı (Kungliga bibliotekets webbarkiv) [101] 1997 yılında başlatılmıştır. Toplayıcı kurum İsveç Milli Kütüphanesi'dir (Kungliga biblioteket [102]. Web taramaları yılda 2 kez yapılmaktadır. Arşiv herkese açıktır fakat yalnızca kütüphane içinde kullanıma sunulur. Web arşivinin arayüz dili İsveççe'dir. Erişim yöntemi URL aramadır. Derleme stratejileri ulusal alan adı, yığın ve seçimli derlemedir. Bununla birlikte “.se” alan adındaki yaklaşık 140 gazete günlük olarak derlenmektedir. Kurum IIPC üyesidir.

Danimarka Milli Kütüphanesi Web Arşivleme Programı

Danimarka Milli Kütüphanesi Web Arşivleme Programı (Netarkivet.dk) [103] 2005 yılında başlatılmıştır. Toplayıcı kurum Danimarka Kraliyet Kütüphanesi'dir [104] (Royal Danish Library). 2004' yılında yayımlanan Derleme Kanunu ile Web arşivleme çalışmaları yasal statü kazanmıştır. Web arşivinin arayüz dili Danca ve İngilizce'dir. Erişim yöntemi URL aramadır. Derleme stratejileri ulusal alan adı, yığın, seçimli ve olay eksenli derlemedir. Arşive erişim araştırma amaçlı olarak sınırlandırılmıştır. Kurum IIPC üyesidir.

Norveç Milli Kütüphanesi Web Arşivleme Programı

Norveç Milli Kütüphanesi Web Arşivleme Programı (Nettarkivet Norge) [105] 2001 yılında başlatılmıştır. Toplayıcı kurum Norveç Milli Kütüphanesi'dir (Nasjonal biblioteket). Web arşivinin arayüz dili Norveççe'dir. Erişim yöntemi anahtar kelime aramadır. Derleme stratejileri ulusal alan adı ve olay eksenli derlemedir. Kurum IIPC üyesidir.

Katalonya Milli Kütüphanesi Web Arşivleme Programı

Katalonya Milli Kütüphanesi Web Arşivleme Programı (Patrimoni Digital de Catalunya- PADICAT) [106] 2005 yılında başlatılmıştır. PADICAT, Web sitelerini korumak ve onlara açık ve kalıcı erişimi sağlayarak özerk bir devlet statüsünde olan Katalonya'nın kültürel, bilimsel ve genel yapısına ilişkin çıktılarını dijital formatta toplama ve koruma hedefi ile faaliyete alınmıştır. Toplayıcı kurum Katalonya Kütüphanesi'dir [107] (Biblioteca de Catalunya). Web arşivinin arayüz dili Katalanca,

İspanyolca ve İngilizce'dir. Erişim yöntemleri URL arama, anahtar kelime arama, alfabetik arama ve konu aramadır. Derleme stratejileri bölgesel alan adı, yığın, olay eksenli ve tematik derlemedir. Kurum IIPC üyesidir.

Slovenya Milli Kütüphanesi Web Arşivleme Programı

Slovenya Milli Kütüphanesi Web Arşivleme Programı (Spletni arhiv narodne in univerzitetne knjižnice) [108] 2007 yılında başlatılmıştır. 2006 yılında yayımlanan Derleme Kanunu toplayıcı kurum olan Slovenya Milli ve Üniversite Kütüphanesi'ne [109] (Narodna in univerzitetna knjižnica) Sloven Web mirasını toplama ve koruma için yetki vermiştir. Web arşivinin arayüz dili Slovence'dir. Erişim yöntemleri URL arama, alfabetik arama ve konu aramadır. Derleme stratejileri ulusal alan adı, seçimli ve tematik derlemedir. 2014 yılına kadar Web siteleri yalnızca seçimli olarak taranmıştır ve halen yaklaşık 1300 site bu şekilde taranmaya devam etmektedir. Seçimli taramalar günlük olarak devam ettirilen tematik taramalarla desteklenmektedir. 2014 yılında başlayan ulusal alan adı taramalarında ise tarama işlemi yılda 1 kez yapılmaktadır. Seçimli ve tematik taramalarla elde edilen koleksiyonlar herkes tarafından erişilebilir durumdadır. “.si” alan adı koleksiyonlarına ise şu anda yalnızca kütüphane çalışanları tarafından erişilebilir. Kurum IIPC üyesidir.

Birleşik Krallık Web Arşivleme Programı

Birleşik Krallık Web Arşivleme Programı (UK Web Archive) [110] 2005 yılında başlatılmıştır. Birleşik Krallık Web Arşivi, önde gelen İngiliz kurumlarının tarihsel, sosyal ve kültürel önemi nedeniyle araştırmacıların yararına seçtiği Web sitelerinin bir derlemesidir. Arşiv kamuya açıktır ve şu anda faaliyette olmayan UKWAC tarafından bir süre desteklenmiştir. Toplayıcı Kurum İngiltere Milli Kütüphanesi'dir (British Library) [111]. Web arşivinin arayüz dili İngilizce'dir. Erişim yöntemleri URL arama, alfabetik arama, tam metin arama ve konu aramadır. Kullanılan derleme stratejileri ise seçimli, olay eksenli ve tematik derlemedir. Birleşik Krallık Web arşivinde 3000'in üzerinde Web sitesi bulunmakta ve her Web sitesi belirlenen bir program dâhilinde yılda 2 kez yakalanmaktadır (NA, 2015:4). Kurum IIPC üyesidir.

Hollanda Milli Kütüphanesi Web Arşivleme Programı

Hollanda Milli Kütüphanesi Web Arşivleme Programı (Webarchief van Nederland) 2005 yılında başlatılmıştır. Toplayıcı kurum Hollanda Milli Kütüphanesi'dir [112] (Koninklijke Bibliotheek). Hollanda'ya ilişkin dil, tarih ve kültür nesnelerinin

toplanması, kataloglanması, arşivlenmesi ve erişime sunulması Program'ın temel amacıdır. Web arşivinin arayüz dili Flemenkçe'dir. Erişim yöntemleri URL arama, alfabetik arama ve tam metin aramadır. Seçilen derleme stratejisi ise seçimli derlemedir. Seçimli derleme yaklaşımının tercih edilmesindeki nedenler “.nl” alan adının çok büyük olması ve Kütüphane'nin hak sahiplerinden bir Web sitesini tarama, arşivleme ve sunma iznini arşivleme işleminden önce almak istemesidir. Kurum IIPC üyesidir.

Çek Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı

Çek Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı (Webarchiv český webový archiv) [113] 2000 yılında faaliyete başlamıştır. Toplayıcı kurum Çek Cumhuriyeti Milli Kütüphanesi'dir [114] (Národní knihovna České republiky). 2017 yılı verilerine göre koleksiyonda yaklaşık 245 TB depolanmış veri bulunmaktadır ve 5000'in üzerinde yayıncıyla anlaşma yoluna gidilmiştir. Web arşivinin arayüz dili Çekçe ve İngilizce'dir. Erişim yöntemleri URL arama ve konu aramadır. Derleme stratejisi ulusal alan adı, olay eksenli ve seçimli derlemedir. Seçimli derleme yapılması ve hak sahiplerinden izinlerin alınmış olması sayesinde koleksiyona çevrimiçi olarak erişilebilir. Alan adı koleksiyonlarına ise Çek Cumhuriyeti Milli Kütüphanesi'nde yerinde erişim sağlanabilir.

Çek Cumhuriyeti Milli Kütüphanesi'nde ulusal alan adı dışındaki “.net”, “.com”, “.org” ve “.edu” gibi genel üst seviye alan adındaki Çek Web sitelerinin toplanabilmesi için bir Web analiz aracından yararlanılmaktadır. Tarama sırasında kullanılan araç olan WebAnalyzer, Web sayfalarını analiz eder ve Çek Web'ini karakterize eden önceden tanımlanmış bazı özellikleri arar (Niu, 2012b). Önceden tanımlanmış her bir özellik bulunduğunda, URL'ye belirli bir miktar puan eklenir. Belirli bir eşik değere ulaşıldığında ise Web sayfası Çek Web'inin bir parçası olarak kabul edilir ve arşivlenir (Niu, 2012b). Kurum IIPC üyesidir.

Avusturya Milli Kütüphanesi Web Arşivleme Programı

Avusturya Milli Kütüphanesi Web Arşivleme Programı (Webarchiv Österreich) [115] 2000 yılında başlatılmıştır. Toplayıcı kurum Avusturya Ulusal Kütüphanesi'dir [116] (Österreichische Nationalbibliothek). 2009 yılında yayımlanan Derleme Yasası Kütüphane'ye Web arşivleme için yasal dayanak sağlamıştır. Web arşivinin arayüz dili Almanca'dır. Erişim yöntemleri URL arama ve konu aramadır. Erişim, Avusturya Milli Kütüphanesi'nde ve önceden belirlenen 20 civarında kütüphanede, yerinde ve kamuya açık olacak şekilde sürdürülmektedir. Derleme stratejisi ulusal alan adı, yığın, seçimli, olay eksenli ve tematik derlemedir. Kurum IIPC üyesidir.

İsviçre Milli Kütüphanesi Web Arşivleme Programı

İsviçre Milli Kütüphanesi Web Arşivleme Programı (Webarchiv Schweiz) [117] 2008 yılında başlatılmıştır. Toplayıcı kurum İsviçre Milli Kütüphanesi'dir [118] (Schweizerische Nationalbibliothek). Web arşivinin arayüz dilleri Almanca, Fransızca, İtalyanca ve İngilizce'dir. Erişim yöntemleri URL arama, anahtar kelime arama, tam metin arama ve konu aramadır. Web arşiv koleksiyonu Kütüphane'ye ait tüm dijital koleksiyona erişim sağlanan İsviçre Ulusal Kütüphanesi'nin erişim sistemi e-Helvetica'ya entegre edilmiştir. Ancak, Web sitelerinin arşivlenmiş versiyonları yalnızca İsviçre Ulusal Kütüphanesi'nin ve koleksiyonu oluşturmaya yardımcı olan partner kütüphanelerin okuma salonlarında izlenebilmektedir. Derleme stratejisi seçimli, olay eksenli ve tematiktir. Kurum IIPC üyesidir.

Letonya Milli Kütüphanesi Web Arşivleme Programı

Letonya Milli Kütüphanesi Web Arşivleme Programı (WWW Archive of the National Library of Latvia) [119] faaliyetlerine 2005 yılında bir pilot proje ile başlamıştır. 2008 yılı itibariyle de yasal derleme kapsamında Web siteleri toplanmaktadır. Ancak Yasal kısıtlamalar nedeniyle arşivlenmiş Web siteleri yalnızca kütüphane binalarında mevcuttur. Toplayıcı kurum Letonya Milli Kütüphanesi'dir [120] (Latvijas Nacionālā Bibliotēka). Web arşivinin arayüz dilleri Letonca, Rusça ve İngilizce'dir. Erişim yöntemleri URL arama, anahtar kelime arama ve konu aramadır. Arşivlenen sitelere yönelik arama yeteneklerini iyileştirmek için, her Web sitesine başlık, açıklama, anahtar kelimeler, konu başlıkları vb. gibi tanımlayıcı üstveriler atanmaktadır. Derleme stratejisi seçimli, olay eksenli ve tematiktir. Kurum IIPC üyesidir.

Almanya Milli Kütüphanesi Web Arşivleme Programı

Almanya Milli Kütüphanesi Web Arşivleme Programı (Deutsche Nationalbibliothek Webarchiv) [121] 2006 yılında faaliyete başlamıştır. Toplayıcı kurum olan Alman Milli Kütüphanesi [122] (Deutsche Nationalbibliothek), Web sitelerini toplamak, kataloglamak, dizinlemek ve arşivlemek için yasa nezdinde sorumludur. Tarama sıklığı kaynak türlerine göre değişmekle birlikte ortalama yılda 4 kezdir. Teknik olarak mümkün olan yerlerde, özel etkinliklerle ilgili sosyal ağlardan seçilen sayfalar da toplanabilmektedir. Fakat sosyal ağlardan içeriklerin düzenli olarak toplanması için bir eylem planı yoktur. Web arşivinin arayüz dili Almanca'dır. Erişim yöntemleri URL arama ve tam metin aramadır. Derleme stratejisi seçimli, yığın, olay eksenli ve tematiktir.

3.11.3. Asya Kıtası Ulusal Kütüphane Web Arşivleri

Kore Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı

Kore Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı (Çevrimiçi Arşivleme ve Internet Kaynaklarını Arama (Online Archiving & Searching Internet Sources- OASIS) 2005 yılında başlatılmıştır [123]. Toplayıcı kurum Kore Cumhuriyeti Milli Kütüphanesi'dir (National Library of Korea) [124]. Web arşivinin arayüz dili Korece'dir. Erişim yöntemleri URL arama, anahtar kelime arama ve konu aramadır. Derleme stratejisi seçimli derlemedir. Kurum IIPC üyesidir.

Japonya Milli Kütüphanesi Web Arşivleme Programı

Japonya Milli Kütüphanesi Web Arşivleme Programı (WARP) [125] 2002 yılında başlatılmıştır. Toplayıcı kurum Japonya Ulusal Meclis Kütüphanesi'dir [126] (Japan National Diet Library). Japonya Ulusal Meclis Kütüphanesi 2010'da yürürlüğe giren Ulusal Meclis Kütüphanesi Kanunu ile Web sitelerini arşivleme yetkisi kazanmıştır. Bunlar arasında ağırlıklı olarak resmî kurumların Web siteleri yer alır. Ayrıca Japonya'da düzenlenen kültürel ve uluslararası etkinliklere ve elektronik dergilere ait Web siteleri, hak sahiplerinin iznine dayanarak arşivlenir. Arşivlenmiş Web siteleri, Japonya Ulusal Meclis Kütüphanesi yerleşkesinde kullanıma sunulur. Hak sahiplerinin izinlerinin alındığı Web siteleri ise Internet üzerinde de erişime sunulmaktadır. Web arşivinin arayüz dili Japonca'dır. Erişim yöntemleri anahtar kelime arama, tam metin arama ve konu aramadır. Derleme stratejisi yığın ve seçimli derlemedir. Kurum IIPC üyesidir.

İsrail Milli Kütüphanesi Web Arşivleme Programı

İsrail Milli Kütüphanesi Web Arşivleme Programı (Israel Internet Archive) [127] 2006 yılında başlatılmıştır. Toplayıcı kurum İsrail Milli Kütüphanesi'dir [128] (The National Library of Israel). Web arşivinin arayüz dili İbranice ve İngilizce'dir. Erişim yöntemleri URL arama ve konu aramadır. Telif hakkı yasasıyla ilgili nedenlerle Arşiv sadece Milli Kütüphane binasında görüntülenebilir. Arşivdeki materyaller, İsrail Milli Kütüphanesi'nin birleşik arama arayüzü olan Merhav Sistemi'nde aktarılır ve orada sunulur. Derleme stratejisi ulusal alan adı, olay eksenli ve seçimli derlemedir. Web sitelerinin derlenme sıklığı yılda 2 kezdir.

Çin Halk Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı

Çin Halk Cumhuriyeti Milli Kütüphanesi Web Arşivleme Programı (WICP) 2005 yılında başlatılmıştır. Programın ana hedefi Çin'in politikalarını, ekonomisini ve kültürünü yansıtan Web kaynak yığınlarının toplanması ve korunmasıdır (The National Library Of China, 2009:10-11). Toplayıcı kurum Çin Milli Kütüphanesi'dir [129] (The National Library of China). Web arşivinin arayüz dili Çince ve İngilizce'dir. Erişim yöntemleri tam metin arama, anahtar kelime arama ve konu aramadır. Derleme stratejisi olay eksenli, seçimli ve tematik derlemedir.

Singapur Milli Kütüphanesi Web Arşivleme Programı

Singapur Milli Kütüphanesi Web Arşivleme Programı (Web Archive Singapore) [130] 2006 yılında başlatılmıştır. Toplayıcı kurum Singapur Milli Kütüphanesi'dir [131] (The National Library of Singapore). Singapur'da Web derleme işlemini yapmasına izin veren yasal derleme mevzuatı bulunmamaktadır (Chellapandi and Lian San, 2009) (Cadavid, 2014:9). Buna karşın Singapur Milli Kütüphanesi bir yayından kaldırma politikası kullanmakta ve bazı durumlarda, arşivleme ve materyallere erişim konusunda Web sitesi sahiplerine bildirim göndermektedir. Web arşivinin arayüz dili İngilizce'dir. Erişim yöntemleri anahtar kelime arama ve konu aramadır. Derleme stratejisi ulusal alan adı, seçimli ve tematik derlemedir. Seçimli ve tematik bir derleme kapsamında başta devlet kurumları ve resmi kuruluşlar olmak üzere yaklaşık 1000 Web sitesi seçilmiştir (Cadavid, 2014:8).

3.11.4. Okyanusya Kıtası Ulusal Kütüphane Web Arşivleri

Yeni Zelanda Milli Kütüphanesi Web Arşivleme Programı

Yeni Zelanda Milli Kütüphanesi Web Arşivleme Programı (New Zealand Web Archive) [132] 1999 yılında başlatılmıştır. Toplayıcı kurum Yeni Zelanda Milli Kütüphanesi'dir [133] (National Library of New Zealand). Web arşivinin arayüz dili İngilizce'dir. Erişim yöntemleri URL arama, anahtar kelime arama, alfabetik arama ve konu aramadır. Web sitelerine erişim, Milli Kütüphane'nin çevrimiçi kataloğunu arayarak ve ardından arşivlenmiş kopyanın bağlantısını tıklatarak sağlanabilir. Derleme stratejisi seçimli derlemedir. Kurum IIPC üyesidir.

Avustralya Milli Kütüphanesi Web Arşivleme Programı

Avustralya Milli Kütüphanesi Web Arşivleme Programı (PANDORA) [134] 1996 yılında başlatılmıştır. PANDORA, Avustralya ve Avustralyalı'ların sosyal, kültürel, politik ve entelektüel yaşamlarıyla ilgili Web materyallerini içeren oldukça kapsamlı bir Web arşividir. Kapsamında elektronik dergiler, hükümet yayınları ve Web siteleri bulunmaktadır (Oğuz, 2006b:21). Toplayıcı kurum Avustralya Milli Kütüphanesi'dir [135] (National Library of Australia). Web arşivinin arayüz dili İngilizce'dir. Erişim yöntemleri URL arama, anahtar kelime arama, alfabetik arama, tam metin arama ve konu aramadır. Derleme stratejileri seçimli ve olay eksenli derlemedir. Web Arşivleme Programı'nda seçimli derleme süreci son derece sistemlidir ve arşivleme öncesinde yayıncılardan izin alma yoluna gidilmiş ve derlenen tüm sitelerin bütünlüğü ve işlevsel olup olmadığı kontrol edilmiştir (Glanville, 2010:130). Ayrıca arşivlenen Web sitelerinin tamamına ait MARC kayıtları kataloğa aktarılmıştır. Kurum IIPC üyesidir.

3.12. Bulguların Değerlendirilmesi ve Araştırma Hipotezlerin Test Edilmesi

Bu bölümde çalışma kapsamında belirlenen araştırma problemlerini yanıtlama amacıyla elde edilen veriler ve ortaya konulan bulgular değerlendirilecek ve belirlenen araştırma hipotezleri test edilecektir.

Araştırma Temel Hipotezi: “Türkiye Milli Kütüphanesi tarafından Web arşivleme uygulamaları gerçekleştirilmezse ulusal Web mirası korunamaz.”

Temel Araştırma Problemi: “Türkiye Milli Kütüphanesi ulusal Web mirasını koruma amacıyla Web arşivleme uygulamasına yer vermeli midir? Eğer yer verecekse diğer ulusal kütüphanelerle karşılaştırıldığında seçeceği Web derleme ve erişim stratejileri neler olmalıdır?” şeklinde belirlenmiştir.

Araştırma temel hipotezinin test edilmesi ve temel araştırma probleminin yanıtlanması amacıyla elde edilen bulgular aşağıda sunulmaktadır:

- Ülkemizde 2012 yılında yürürlüğe giren 6279 sayılı Çoğaltılmış Fikir ve Sanat Eserlerini Derleme Kanunu [136] ve 28388 sayılı Resmî Gazetede yayımlanan Çoğaltılmış Fikir ve Sanat Eserlerini Derleme Yönetmeliği [137], elektronik ortamda üretilerek kullanıma sunulmuş elektronik yayınları derleme konusunda Türkiye Milli Kütüphanesi'ni yetkili ve sorumlu kılmıştır. Söz konusu mevzuatta yer alan “çevrimiçi elektronik yayın” ifadesinin kapsamına Web materyalleri de girdiğinden, Türkiye Milli Kütüphanesi faaliyet süreçlerinde Web arşivleme uygulamalarına yer vermelidir.

- 4848 sayılı K lt r ve Turizm Bakanlıđı Te kil t ve G revleri Hakkında Kanun’da ge en “Mill  K t phane Ba kanlıđı, T rkiye’nin mill  k lt r ve bilgi birikimi ile bilgi akımını sađlar.” (2013:madde 15) h km , k lt rel miras  r nlerinin sađlanması, korunması ve aktarımı noktasında T rkiye Milli k t phanesi’ne g rev vermektedir. Bu bađlamda, ulusal Web mirasının korunması amacıyla T rkiye Milli K t phanesi’nin faaliyetleri kapsamında Web ar ivleme uygulamalarına yer vermesi ka ınılmazdır.
- UNESCO ve IFLA yayınladıkları bildirgeler aracılıđıyla ulusal hafızanın ve k lt rel kimliđin ta ıyıcısı olarak da ifade edilebilecek miras varlıklarının k lt rel miras kurumlarınca koruması gerektiđi vurgusunda bulunmaktadır. Web mirasını olu turan materyallerin derlenmesine y nelik yasal zorunluluđun bulunduđu  lkelerde ise, Web ar ivleme faaliyetleri genellikle tek bir kurulu un sorumluluđundadır (ISO, 2013:15). Hem bir k lt rel miras kurumu olmasından hem de ulusal bilgi politikalarını y r tmedeki stratejik konumundan hareketle, ulusal k lt r mirası kapsamında deđerlendirilen dijital formdaki bilgi kaynaklarını koruma ve ar ivleme sorumluluđu bir ok  lkede ulusal k t phaneler tarafından yerine getirilmektedir. Bu bađlamda, dijital k lt r mirası kapsamında deđerlendirilen ve ulusal nitelik ta ıyan Web materyallerinin, T rkiye Milli K t phanesi tarafından ar ivlenmesi ve korunarak gelecek nesillere aktarılması gerekli g r lmektedir.
- Bilgi ve k lt r varlıklarının yalnızca iktisadi kalkınma ama lı deđil, daha tatmin edici entelekt el, duygusal, ahlaki ve ruhsal varlık elde etmeye y nelik bir geli imin de temel unsurlarından biri olduđu (UNESCO, 2001:2) d   ld đ nde, T rkiye Milli K t phanesi’nin bilimsel, ekonomik ve toplumsal geli imi destekleme adına ulusal Web mirasını ar ivlemesi  nemlidir.
- Ulusal k t phanelerin Web ar ivlemeye y nelik kullanılan ara ların, standartların, dosya formatlarının, istatistik ve kalite g stergelerinin olu turulma  abalarına ve uluslararası Web ar ivleme giri imlerinin ortaya  ıkmalarına  nc l k etmi  olmaları, onların Web ar ivleme faaliyetlerinin s rd r lmesindeki ve yaygınla masındaki  nemini ortaya koymu tur. T rkiye Milli K t phanesi’nin de yerel, ulusal ve evrensel deđeri bulunan Web materyallerine koleksiyonunda yer vermesi ve uluslararası Web ar ivleme giri imlerinin faaliyetlerine katılım g stererek katkı sađlaması ulusal Web mirasının korunması ve uluslararası g r n rl đ n arttırılması adına  nemlidir.

Çalışma kapsamında belirlenen araştırma temel hipotezinin test edilmesi ve temel araştırma probleminin yanıtlanması amacıyla dünya genelinde Web arşivleme uygulamalarına yer veren ulusal kütüphanelerin derleme stratejileri, Web arşivleme programlarına başlama yılları, erişim yöntemleri ve arayüz dillerine ilişkin bilgiler bir araya getirilmiş ve Tablo 3'te sunulmuştur.

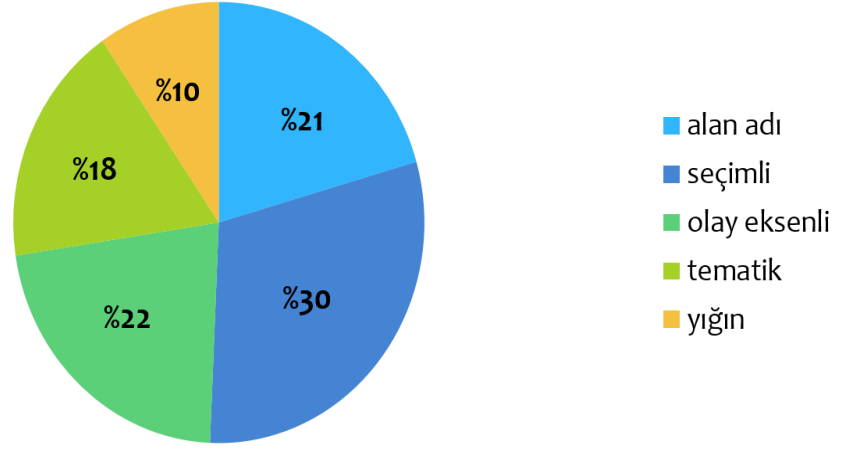
Tablo 3. Dünya Genelinde Web Arşivleme Uygulamalarına Yer Veren Ulusal Kütüphanelerin Derleme Stratejileri, Web Arşivleme Programlarına Başlama Yılları, Erişim Yöntemleri ve Arayüz Dilleri

Kıta	Toplayıcı Kurum	Web Derleme Stratejisi	Yıl	Erişim Yöntemi	Arayüz Dili
Kuzey Amerika	Kanada Milli Kütüphanesi	Alan adı derleme (bölgesel)	2005	URL arama Anahtar kelime arama Alfabetik arama Tam metin arama	İngilizce Fransızca
	Kongre Kütüphanesi (ABD)	Seçimli derleme Olay eksenli derleme Tematik derleme	2000	URL arama Alfabetik arama Konu arama	İngilizce
Avrupa	İspanya Milli Kütüphanesi	Alan adı derleme (genel ve ulusal)	2009	URL arama	İspanyolca İngilizce
	Fransa Milli Kütüphanesi	Alan adı derleme (ulusal) Yığın derleme Seçimli derleme Olay eksenli derleme	2002	URL arama Anahtar kelime arama Tam metin arama Konu arama	Fransızca
	Estonya Milli Kütüphanesi	Seçimli derleme Olay eksenli derleme Tematik derleme	2010	URL arama Anahtar kelime arama Konu arama	Estonca
	Hırvatistan Milli Kütüphanesi	Alan adı derleme (ulusal) Seçimli derleme Tematik derleme	2010	URL arama Anahtar kelime arama Alfabetik arama Konu arama	Hırvatça İngilizce
	İzlanda Milli Kütüphanesi	Alan adı derleme (ulusal) Seçimli derleme Olay eksenli derleme	1996	URL arama	İzlandaca İngilizce (sınırlı)
	Finlandiya Milli Kütüphanesi	Alan adı derleme (ulusal ve bölgesel) Tematik derleme	2006	URL arama Tam metin arama	Fince İsveççe İngilizce
	İsveç Milli Kütüphanesi	Alan adı derleme (ulusal) Yığın derleme Seçimli derleme	1997	URL arama	İsveççe

	Danimarka Milli Kütüphanesi	Alan adı derleme (ulusal) Yığın derleme Seçimli derleme Olay eksenli derleme	2005	URL arama	Danca İngilizce
	Norveç Milli Kütüphanesi	Alan adı derleme (ulusal) Olay eksenli derleme	2001	Anahtar kelime arama	Norveççe
	Katalonya Milli Kütüphanesi	Alan adı derleme (bölgesel) Yığın derleme Olay eksenli derleme Tematik derleme	2005	URL arama Anahtar kelime arama Alfabetik arama Konu arama	Katalanca İspanyolca İngilizce
	Slovenya Milli Kütüphanesi	Alan adı derleme (ulusal) Seçimli derleme Tematik derleme	2007	URL arama Alfabetik arama Konu arama	Slovençe
	British Library (İngiltere)	Seçimli derleme Olay eksenli derleme Tematik derleme	2005	URL arama Alfabetik arama Tam metin arama Konu arama	İngilizce
	Hollanda Milli Kütüphanesi	Seçimli derleme	2005	URL arama Alfabetik arama Tam metin arama	Flemenkçe
	Çek Cumhuriyeti Milli Kütüphanesi	Alan adı derleme (ulusal) Olay eksenli derleme Seçimli derleme	2000	URL arama Konu arama	Çekçe İngilizce
	Avusturya Milli Kütüphanesi	Alan adı derleme (ulusal) Yığın derleme Seçimli derleme Olay eksenli derleme Tematik derleme	2000	URL arama Konu arama	Almanca
	İsviçre Milli Kütüphanesi	Seçimli derleme Olay eksenli derleme Tematik derleme	2008	URL arama Anahtar kelime arama Tam metin arama Konu arama	Almanca Fransızca İtalyanca İngilizce
	Letonya Milli Kütüphanesi	Seçimli derleme Olay eksenli derleme Tematik derleme	2005	URL arama Anahtar kelime arama Konu arama	Letonca Rusça İngilizce
	Almanya Milli Kütüphanesi	Seçimli derleme Yığın derleme Olay eksenli derleme Tematik derleme	2006	URL arama Tam metin arama	Almanca
	Kore Cumhuriyeti	Seçimli derleme	2005	URL arama Anahtar kelime arama	Korece

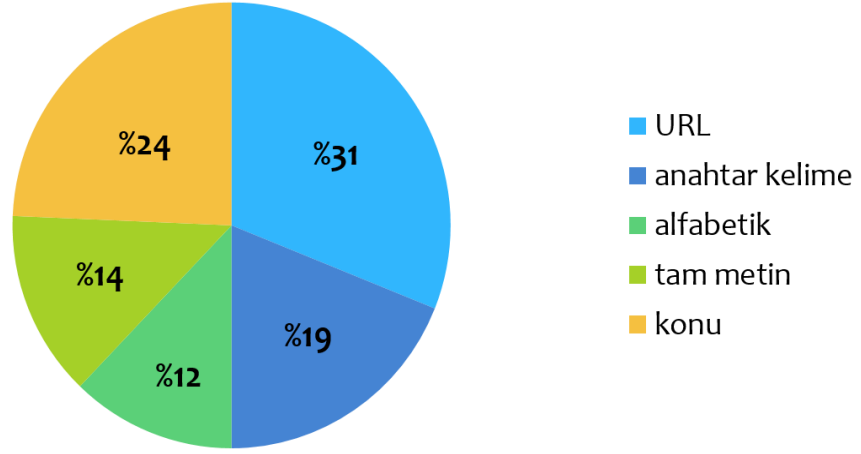
Asya	Milli Kütüphanesi			Konu arama	
	Japonya Milli Kütüphanesi	Seçimli derleme Yığın derleme	2002	Anahtar kelime arama Tam metin arama Konu arama	Japonca
	İsrail Milli Kütüphanesi	Alan adı derleme (ulusal) Olay eksenli derleme Seçimli derleme	2006	URL arama Konu arama	İbranice İngilizce
	Çin Halk Cumhuriyeti Milli Kütüphanesi	Olay eksenli derleme Seçimli derleme Tematik derleme	2005	Tam metin arama Anahtar kelime arama Konu arama	Çince İngilizce
	Singapur Milli Kütüphanesi	Alan adı derleme (ulusal) Seçimli derleme Tematik derleme	2006	Anahtar kelime arama Konu arama	İngilizce
Okyanusya	Yeni Zelanda Milli Kütüphanesi	Seçimli derleme	1999	URL arama Anahtar kelime arama Alfabetik arama Konu arama	İngilizce
	Avustralya Milli Kütüphanesi	Seçimli derleme Olay eksenli derleme	1996	URL arama Anahtar kelime arama Alfabetik arama Tam metin arama Konu arama	İngilizce

- Web arşivleme uygulamalarına faaliyet süreçlerine yer veren 27 ulusal kütüphaneden 15'i alan adı derleme, 22'si seçimli derleme, 16'sı olay eksenli derleme, 13'ü tematik derleme ve 7'si yığın derlemeyi tercih etmiştir. Söz konusu ulusal kütüphanelerin tercih ettiği Web derleme stratejileri arasında seçimli derleme, olay eksenli derleme ve alan adı derleme en çok tercih edilen Web derleme stratejileri olarak sıralanmaktadır. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Web Derleme Stratejileri Yüzdesel Dağılım Grafiği Şekil 11'de sunulmaktadır.



Şekil 11. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Web Derleme Stratejileri Yüzdesel Dağılım Grafiği

- Alan adı derlemenin ulusal Web sitelerini toplamayı hedeflemesi ve seçimli ve olay eksenli derlemenin istenilen kapsamda Web koleksyonları oluşturmaya olanak tanınması, söz konusu Web derleme stratejilerinin ulusal kütüphaneler tarafından sıklıkla tercih edilmelerinde etkili olmuştur. Yığın derleme ise ulusal mirasın da da ötesine geçerek tüm Web'i toplamayı hedeflemesi ve yüksek depolama maliyeti oluşturması nedeniyle ulusal kütüphaneler tarafından daha az tercih edilmiştir. Bu bağlamda Türkiye Milli kütüphanesinin Web arşivleme uygulamasına yer vermesi durumunda seçeceği Web derleme stratejilerinin seçimli derleme, olay eksenli derleme ve alan adı derleme olması ulusal Web mirasının korunması adına önemli görülmektedir.
- Web arşivleme uygulamalarına faaliyet süreçlerine yer veren 27 ulusal kütüphaneden 23'ü URL arama, 14'ü anahtar kelime arama, 9'u alfabetik arama, 10'u tam metin arama ve 18'i konu arama erişim yöntemini desteklemektedir. Ulusal kütüphanelerin desteklediği erişim yöntemleri arasında URL arama, konu arama ve anahtar kelime arama en çok tercih edilen erişim yöntemleri olarak sıralanmıştır. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Erişim Yöntemleri Yüzdesel Dağılım Grafiği Şekil 12'de sunulmaktadır.



Şekil 12. Web Arşivleme Uygulamalarında Bulunan Ulusal Kütüphanelerin Erişim Yöntemleri Yüzdesel Dağılım Grafiği

- Dizinleme işlemi gerektirmemesi söz konusu erişim yöntemleri arasında URL aramayı ön plana çıkarmıştır. Konu arama, küratöryal çaba gerektirmekle birlikte erişim konusunda istenilen sonuçlara daha yüksek oranda ulaşma imkânı sunar. Anahtar kelime arama ise otomatik ve/veya manuel yollarla üstveri oluşturmayı gerektirmekle birlikte olası veri madenciliği uygulamalarını desteklemez. Elde edilen bulguların bir sonucu olarak, Türkiye Milli Kütüphanesinin Web arşivleme uygulamasına yer vermesi durumunda seçeceği erişim yöntemlerinin URL arama, konu arama ve anahtar kelime arama olması bilgiye erişim bağlamında ihtiyaçları karşılama adına önemli görülmektedir.
- Çalışma kapsamında ele alınan ulusal kütüphanelerin uyguladığı erişim sınırlamaları değişkenlik göstermektedir. Ulusal kütüphanelerin bir bölümü Web arşivlerini Web arayüzü aracılığıyla, bir bölümü yalnızca kütüphanenin belirlediği lokasyonlarda kamuya sunmakta; diğer bir diğer bölümü ise koleksiyonun bir bölümünü veya tamamını kamusal erişime kapatmaktadır. Erişim konusunda yasal derleme mevzuatının toplayıcı kuruma verdiği yetki ve yayıncılardan alınan izinler belirleyici olmaktadır. Bununla birlikte uygulamada yayıncıların izinlerinin alınması yoluyla bazı Web sitelerinin kamusal erişime açık hale getirildiği veya yayıncıların erişim engellemesi talep etmesi durumunda ilgili içeriğin tamamen erişime kapatıldığı veya hak sahiplerinin kapatma talebi nedeniyle yalnızca yerinde erişime izin verildiği durumlar da mevcuttur. Ayrıca kimi durumlarda kullanıcılara yalnızca izleme yetkisi verilmekte ve Web arşivi içeriğinin kopyalarının alınması yasaklanmaktadır. Koleksiyonunu Web arayüzü aracılığıyla kamuya açık biçimde erişime sunan ulusal kütüphanelere Estonya, Hırvatistan, İzlanda, Çek Cumhuriyeti Milli Kütüphaneleri örnek verilebilir. Yalnızca yerinde erişime izin veren kütüphanelere ise Fransa, Finlandiya, İsveç, Avusturya, İsviçre, Letonya ve İsrail Milli Kütüphaneleri örnek verilebilir. Web arşiv koleksiyonunun bir bölümünün kamusal erişime kapalı olduğu ulusal kütüphanelere ise Slovenya Milli Kütüphanesi örnektir.

- Ülkemizde çevrimiçi materyallerin toplanmasına yönelik yetki, derleme yasası aracılığıyla Türkiye Milli Kütüphanesine verilmiştir. İlgili yasada Web sitelerinin arşivlenmesine yönelik doğrudan bir hüküm bulunmasa da Web sitelerinin çevrimiçi materyal sınıfına girmesi nedeniyle Türkiye Milli Kütüphanesi'nin gelecekte Web arşivleme uygulamalarına faaliyetlerinde yer vereceği öngörülmektedir. Türkiye Milli Kütüphanesi'nin Web arşivleme uygulamalarına faaliyetlerinde yer vermesi durumunda, telif ve mülkiyet haklarından doğan sınırlamalar nedeniyle toplanacak materyallerin bir bölümünün yerinde erişim yoluyla, diğer bölümünün ise Web arayüzü üzerinden sunulması olanaklıdır.

Çalışma kapsamında belirlenen araştırma alt hipotezleri ve alt araştırma problemi aşağıda sunulmaktadır:

Araştırma Alt Hipotezi 1: “Küresel rekabet düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Araştırma Alt Hipotezi 2: “İnsani gelişme düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Araştırma Alt Hipotezi 3: “Bilgi ve iletişim teknolojileri gelişmişlik düzeyi yüksek olan ülkelerde ulusal kütüphanelerin Web arşivleme uygulamalarına yer verme sıklığı fazladır.”

Alt Araştırma Problemi: “Ülkelerin küresel rekabet, insani gelişme ve bilgi ve iletişim teknolojileri gelişmişlik düzeyi ile ulusal kütüphanelerinin Web arşivleme uygulamalarına yer vermesi arasında bir ilişki var mıdır?”

Çalışma kapsamında belirlenen araştırma alt hipotezlerinin test edilmesi ve alt araştırma probleminin yanıtlanması amacıyla 2016 Küresel Rekabet Endeksi, 2016 İnsani Gelişime Endeksi ve 2016-2017 Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi verileri bir araya getirilerek Tablo 4 aracılığıyla aktarılmaktadır.

Tablo 4. Ulusal Kütüphanelerde Web Arşivleme Uygulamalarına Yer Veren Ülkelerin Küresel Rekabet Endeksi, İnsani Gelişme Endeksi, Bilgi Ve İletişim Teknolojileri Gelişmişlik Endeksi Sıralamaları [150]

ÜLKE	Küresel Rekabet Endeksi Sırası (2016-2017)	İnsani Gelişme Endeksi Sırası (2016)	Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi Sırası (2016)
Almanya	5	4	12
Amerika Birleşik Devletleri	3	10	15
Avustralya	22	2	14
Avusturya	19	24	23
Birleşik Krallık	7	16	5
Çek Cumhuriyeti	31	28	32
Çin Halk Cumhuriyeti	28	90	-
Danimarka	12	5	3
Estonya	30	30	18
Finlandiya	10	23	17
Fransa	21	21	16
Kore Cumhuriyeti	26	18	1
Hırvatistan	74	45	41
Hollanda	4	7	8
İspanya	32	27	26
İsrail	24	19	30
İsveç	6	14	7
İsviçre	1	2	4
İzlanda	27	9	2
Japonya	8	17	10
Kanada	15	10	25
Letonya	49	44	40
Norveç	11	1	9
Yeni Zelanda	13	13	13
Singapur	2	5	20
Slovenya	56	25	33
Toplam Ülke Sayısı	138	188	167
Ulusal Kütüphanelerde Web Arşivleme Uygulamalarına Yer Veren Ülkelerin Ortalama Sırası	20,61	19,57	16,96
Türkiye	51	71	72
*Katalonya bağımsız bir ülke olmaması nedeniyle ilgili endekslerde yer almadığından hesaplamaların dışında tutulmuştur.			
**Türkiye Milli Kütüphanesi Web arşivleme uygulamalarına yer vermediği halde, ilgili endekslerdeki sıralamalarının aktarılmasının konunun önemine dikkat çekmek bakımından yararlı olacağı düşünülerek tabloya eklenmiştir.			

Ülkelerin verimlilik düzeyini belirleyen etmenleri ortaya koyup rekabet edebilirliğini ölçen Küresel Rekabet Endeksi, yıllık olarak Dünya Ekonomik Forumu (WEF) tarafından yayımlanmaktadır. Endeks'in amacı ülkelerin güçlü ve zayıf yönlerinin belirleyerek ve ulusal rekabet edebilirliğin mikroekonomik ve makroekonomik temellerini ölçmektir. Endeks'in son olarak yayımlanan 2016-2017 raporunda 138 ülke yer almıştır. Çalışma kapsamında belirlenen belirlenen birinci araştırma alt hipotezi test edildiğinde ilgili rapordan aşağıdaki sonuçlara ulaşılmıştır:

- 2016-2017 Küresel Rekabet Endeksi'nde yer alan 138 ülke arasında, ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren 26 ülkenin ortalama sırası 20,61'dir. Bu sonuç ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren ülkelerin Küresel Rekabet Endeksi'nde ön sıralarda bulunduğunu göstermektedir.
- 2016-2017 Küresel Rekabet Endeksi'ne göre ilk beş sırayı İsviçre (1), Singapur (2), Amerika Birleşik Devletleri (3), Hollanda (4), Almanya (5) almaktadır ve bu ülkelerin tamamı ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yer vermektedir.

Ülkelerin insani gelişme düzeyi uzun ve sağlıklı bir yaşam, bilgiye erişim ve insana yakışır yaşam koşulları olarak belirlenen insani gelişimin üç temel kriterlerini temel alarak ölçülmekte ve sonuçlar Birleşmiş Milletler Kalkınma Programı (UNDP) tarafından yayımlanan İnsani Gelişme Endeksi'nde sunulmaktadır. Yeni bir kalkınma ölçümü geliştirilmesi amacıyla hazırlanan İnsani Gelişme Endeksi'ne ilişkin son rapor toplamda 188 ülkenin kapsamına dâhil edilmesiyle 2016 yılında yayımlanmıştır. Çalışma kapsamında belirlenen belirlenen ikinci araştırma alt hipotezi test edildiğinde ilgili rapordan aşağıdaki sonuçlara ulaşılmıştır:

- 2016 yılı İnsani Gelişme Endeksi'nde yer alan 188 ülke arasında, ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren 26 ülkenin ortalama sırası 19,57 dir. Bu sonuç ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren ülkelerin İnsani Gelişme Endeksi'nde ön sıralarda yer aldığını göstermektedir.
- 2016 yılı İnsani Gelişme Endeksi'ne göre ilk beş sırayı Norveç (1), Avustralya (2), İsviçre (2), Almanya (4), Danimarka (5) ve Singapur (5) paylaşmaktadır ve bu ülkelerin tamamı ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yer vermektedir.

Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi, bilgi ve iletişim teknolojileri alanında yaşanan gelişmeleri gözlemlemeyi, ülkeler arasında karşılaştırma yapmayı ve ülkelerin dijital farklılık düzeylerini ortaya koymayı amaçlar. Birleşmiş Milletler Uluslararası Telekomünikasyon Birliği'nin (ITU) yayımladığı Endeks, ülkelerin bilgi ve iletişim teknolojilerine erişim, onları kullanım ve bu alandaki diğer yeteneklerinin ölçülmesi yoluyla hesaplanmaktadır. Endeks'in 2016 tarihli son raporunda 167 ülke kapsama dâhil edilmiştir. Çalışma kapsamında belirlenen belirlenen üçüncü araştırma alt hipotezi test edildiğinde ilgili rapordan aşağıdaki sonuçlara ulaşılmıştır:

- 2016 yılı Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi'nde yer alan 167 ülke arasında, ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren 25 ülkenin ortalama sırası 16,96'dır. Bu sonuç ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yeren ülkelerin Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi'nde ön sıralarda yer aldığını göstermektedir.
- 2016 yılı Bilgi ve İletişim Teknolojileri Gelişmişlik Endeksi'ne göre ilk beş sırayı Kore Cumhuriyeti (1), İzlanda (2), Danimarka (3), İsviçre (4) ve Birleşik Krallık (5) almıştır. Söz konusu ülkelerin tamamı ulusal kütüphaneleri aracılığıyla Web arşivleme uygulamalarına yer vermektedir.

3.13. Türkiye Milli Kütüphanesine Yönelik Bir Değerlendirme

Türkiye Milli Kütüphanesi'nin Web üzerinde yer alan ulusal nitelikteki dijital bilgi ve kültür nesnelerini arşivlemesi ve böylelikle ulusal Web mirasını gelecek nesillere aktarması oldukça önemlidir. Web arşivleme bağlamında Türkiye Milli Kütüphanesi'ndeki mevcut duruma bakıldığında, ulusal bir Web arşivinin oluşturulmasına yönelik öngörülerin var olduğu ve konunun değerlendirme aşamasında olduğu söylenebilir.

Türkiye Milli Kütüphanesi ulusal Web mirasını sahiplenme amacıyla Web arşivleme uygulamalarını gerçekleştirecekse ilk olarak oluşturacağı Web arşiv koleksiyonunun kapsamını ve seçeceği derleme stratejisini belirlemelidir. Koleksiyona Web materyallerini kazandırma amacıyla yürütülecek seçme ve derleme aşamalarında, ilk olarak seçimli derlemenin ardından alan adı ve olay eksenli derlemenin uygulanması tarafımızca önerilmektedir. Seçimli derleme ile doğrudan hedef yönelik Web sitelerinin toplanması ve sınırları net olarak çizilebilen koleksiyonlar oluşturmak mümkündür. Alan adı derleme ile “.tr” alan adına sahip Web sitelerinin toplanması mümkün hale gelir. Ardından, koleksiyon bütünlüğünü sağlama amacıyla “.tr” alan adına sahip olmayan fakat

ulusal kültür mirası ürünlerini barındıran Web sitelerinin de toplanması amaçlanabilir. Bu noktada Türkiye Milli Kütüphanesi URL puanlama prensibiyle çalışan bir tür Web analiz aracı olan ve Çek Cumhuriyeti Milli Kütüphanesi'nde de kullanılan “Web analyzer” adlı yazılım aracını veya onunla benzer işlevselliğe sahip bir yazılım aracını devreye alabilir. Seçimli ve alan adı derlemeye ek olarak toplumsal veya stratejik önemi bulunan olaylara yönelik oluşturulan Web sitelerinin toplanması amacıyla olay eksenli derleme stratejisinin devreye alınması da yararlı olabilir. Uygulama örneklerine bakıldığında seçimli ve olay eksenli derlemenin sınıflama ve kataloglama gibi küratöryal faaliyetlerle aktif bir biçimde desteklendiği görülmektedir.

Web materyallerin derleme işlemleriyle sağlanmasının ardından bilgi nesneleri organize edilerek önceden belirlenen depolama ortamlarına aktarılabilir. Türkiye Milli Kütüphanesi'nde Web arşiv materyallerinin depolanması sürecinde WARC konteyner dosya formatının kullanımı tarafımızca önerilmektedir. Bunun nedeni WARC konteyner dosya formatının güvenli depolama ve dosya yönetimi gibi işlevsellikleri sunması ve uzun süreli dijital koruma uygulamalarını desteklemesidir. Veri depolama işleminin ardından materyal tanımlamaya ve üstveri oluşturmaya yönelik adımların atılması önerilir. Üstveriler Web sitelerini yakalama esnasında otomatik olarak çekilebileceği gibi seçimli veya olay eksenli derleme ile oluşturulacak koleksiyonlar için manuel olarak da oluşturulabilir. Üstveri oluşturma sürecinde küratörlerden destek alınması tercih edilebilir.

Web materyallerinin sunumunda arayüzün hangi erişim yöntemlerini ve arama işlevselliklerini destekleneceği kullanıcılar açısından önemlidir. Türkiye Milli Kütüphanesi Web arşiv arayüzü üzerinde URL arama ve tam metin aramanın desteklenmesi aranan içeriğe hızlı ve isabetli bir şekilde erişim ve verimlilik açısından önem taşıyacaktır. Desteklenmesi uygun görülen işlevsellikler ise veri madenciliği, kişiselleştirilmiş Web hizmetleri, farklı arama parametreleri sunma veya mükerrerlik kontrolü olarak sıralanabilir.

Web arşiv materyallerinin sunumunda açık erişim materyallerin doğrudan kamuya paylaşımı tarafımızca önerilmektedir. Bununla birlikte yayıncılardan alınan izinler kapsamında erişim engeli bulunmayan Web materyallerinin Web arşivi arayüzü üzerinden sunumu kamuya açık olacak şekilde gerçekleştirilebilir. Erişim engeli bulunan Web materyallerinin ise yerinde erişime açık hale getirilerek yalnızca Türkiye Milli Kütüphanesi'nin seçtiği lokasyonlarda kullanıma sunulması tarafımızca önerilmektedir. Türkiye Milli Kütüphanesi'nin Web arşiv koleksiyonlarında bulunacak materyallerin her biri, telif ve mülkiyet hakları ve yayıncılardan alınan izinler kapsamında görüntülemeye ve kopyalamaya açık veya kapalı hale getirilmeli ve bu süreç kontrollü bir biçimde yürütülmelidir. Erişim aşamasının ardından dijital korumaya yönelik önceden belirlenen göç, öykünüm veya yedekleme gibi stratejiler beliren ihtiyaca göre devreye alınabilir ve

seçilen koruma stratejileri sistemli olarak sürdürülebilir. Arşivlenen Web materyallerinin aktif bir şekilde yönetimi ve yapılandırılması adına çalışma kapsamında sunulan Archive-It Web Arşivleme Yaşam Döngüsü Modeli, Türkiye Milli Kütüphanesi için bir referans kaynak olarak görülebilir ve iş süreçlerine adapte edilebilir.

Ülkemizde yürürlükte olan yasal derleme mevzuatında çevrimiçi materyallerin derlenmesi hususunda Türkiye Milli Kütüphanesi'nin yetkili ve sorumlu tutulması, Kurum'a ihtiyaç duyduğu yasal derleme iznini vermekte ve Kurum'u desteklemektedir. Çoğaltılmış Fikir ve Sanat Eserlerini Derleme Kanunu'nda geçen "Elektronik ortamda üretilerek kullanıma sunulmuş elektronik yayın" ve Yönetmelik'te geçen "çevrimiçi ve çevrimdışı yayın" kavramları derlemenin kapsamını bir ölçüde belirlese de, Web derleme sürecinde kaynakların hangilerinin, neden, ne kadar sıklıkta ve ne ölçüde derleneceği ve hangilerinin dışlanacağı üzerinde dikkatlice düşünülmesi gereken konulardır. Yönetmelik'te "çevrimiçi elektronik yayın: İnternet veya ağlar üzerinden yaymak ve/veya umuma iletmek amacıyla erişime sunulan yayınlar" olarak ifade edilmiştir. Fakat bu ifade Web derlemenin kapsamını belirlemekten uzaktır. Söz konusu Kanun ve Yönetmelik birlikte düşünüldüğünde, Web derlemenin elektronik yayınları kapsayabileceği ve "çevrimiçi" ifadesinin sadece "yayınları" kapsadığı düşünülmelidir. Bu bağlamda örneğin kullanıcı izinlerinin alınması gerekliliği nedeniyle sosyal medya sitelerinin, şifreli içeriğin veya resmî yazışma ve resmî belgelerin dışlanması gerekebilir. Bununla birlikte, ülkemizde kanun koyucuların Web materyallerinin derlenmesine yönelik hükümlere yasal derleme mevzuatında daha açık ve net bir biçimde yer vermesinin Web arşivleme uygulamaları için daha güçlü bir yasal dayanak oluşturacağı düşünülmektedir.

Yasal derleme desteği almanın yanında Türkiye Milli Kütüphanesi'nin Web arşivlemeye yönelik bir program ve strateji ortaya koyması önemlidir. Oluşturulacak program ve strateji temelinde yer verilecek Web arşivi mimarisinin veri madenciliği uygulamaları başta olmak üzere çeşitli işlevsellikleri destekleyecek şekilde yapılandırılması önerilir. Ayrıca, Türkiye Milli Kütüphanesi'nin uygulamalarında yararlanacağı standartları, zorluk yaşanabileceği alanları ve kullanacağı istatistik ve kalite göstergelerini önceden belirlemesi ve söz konusu program ve strateji temelinde bunlara iş süreçlerinde detaylı olarak yer vermesi de önerilir. Ayrıca, iş birliği yapılacak kurumların belirlenmesi, uluslararası girişimlere katılım gösterilmesi ve bilgi ve tecrübe paylaşımı amacıyla çeşitli projelerde yer alınması Türkiye Milli Kütüphanesi'nin Web arşivleme faaliyetlerini destekleyecektir.

DÖRDÜNCÜ BÖLÜM

SONUÇ VE ÖNERİLER

4. SONUÇ VE ÖNERİLER

Bilgi ve kültür varlıkları bin yıllardır birer miras unsuru olarak gelecekte yeniden kullanılmak üzere çeşitli ortamlara kaydedilmekte ve insanlığın sürdürdüğü yaşamı şekillendirmektedir. Yaklaşık yarım yüzyıl öncesine kadar fiziksel ortamda üretilen ve kullanılan söz konusu varlıklar, bilgi ve iletişim teknolojilerinde yaşanan ilerlemeyle birlikte biçim değiştirerek dijital ortamda üretilmeye ve kullanılmaya başlanmıştır. Bilgi çağı olarak adlandırılan bu süreç dijital kültürün ortaya çıkışını ve yaygınlaşmasını desteklemiş ve bilgiyi yoğun bir biçimde kullanan toplumsal yapıları ortaya çıkarmıştır.

Bilgi çağında dijital bilgi ve kültür varlıklarının üretiminde ve kullanımında görülen hızlı artış, kültürel miras kurumlarını bu varlıkları koruma ve gelecek nesillere aktarma adına daha fazla sorumluluk almaya zorlamıştır. Bilginin dijital ortamda üretimi ve sürdürülen dijitalleştirme faaliyetleri neticesinde bilgi ve kültür nesneleri koleksiyonlarda daha fazla yer almaya başlamış ve başta ulusal kütüphaneler olmak üzere birçok kültürel miras kurumu hizmet çeşitliliklerini dijital hizmetleri de kapsayacak şekilde genişletmiştir. Ulusal bilgi ve kültür nesnelerine koleksiyonlarında yer veren ulusal kütüphaneler ise dijital bilgi ve kültür varlıklarının arşivlenmesi, korunması ve aktarımı noktasında daha fazla yetki ve sorumluluk alarak stratejik açıdan önemli ve öncelikli kurumlara dönüşmüştür.

Günümüzde Internet'in yaygınlaşmasıyla değeri giderek artan Web mirası, dijital kültür mirasının önemli bir parçası ve bilgi çağının hafızası konumundadır. Söz konusu hafızanın güçlendirilmesi ve kalıcı bir biçimde korunması ise Web materyallerinin doğru bir şekilde arşivlenmesi ve erişime sunulması ile mümkün olur. Fakat Web arşivleme faaliyeti kültürel miras kurumları için oldukça zor ve karmaşık süreçleri içinde barındırmaktadır. Web'in milyarlarca Web sitesini içinde barındıran devasa hacmi, kısa ömürlü doğası ve dinamik yapısı, kültürel miras kurumları arasında koruma ve aktarım sorumluluğunu üzerlerinde en çok hisseden kurumlar olan ulusal kütüphaneleri Web

arşivleme faaliyetlerinin hayata geçirilmesi ve sürdürülmesi noktasında zorlamaktadır. Çalışmamızda yerel, ulusal veya evrensel nitelikteki kültürel miras varlıklarını koruma ve onları gelecek nesillere aktarma sorumluluğu ile ulusal kütüphanelerin bir bölümünün Web arşivleme faaliyetlerinde bulunduğu, diğer bir bölümünün ise ya Web arşivleme faaliyetlerini planladığı ya da faaliyetlerin oldukça uzağında olduğu gözlemlenmiştir.

Kültürel miras varlıklarının kaybolması ve kuşaklar arasında dijital bilgi boşluğunun oluşma riski, ülkelerde dijital mirası koruma amacıyla bir derleme yasası oluşturulmasını gerekli kılmıştır. Bunun bir sonucu olarak yasal mevzuat desteği altında Web arşivleme uygulamalarına yer veren kurumlarda Web arşivleme faaliyetlerinin hız kazandığı söylenebilir. Web arşivleme kapsamında ulusal derleme mevzuatının desteğini alan ulusal kütüphanelere Kanada, Estonya, Danimarka, Slovenya, Avusturya, Letonya, Almanya, Japonya Ulusal Kütüphaneleri örnek verilebilir. Ülkelerdeki derleme mevzuatının toplayıcı kuruma açıkça yetki vermediği durumlarda ise Web arşivleme faaliyetleri genellikle yayıncıların veya hak sahiplerinin izinlerinin alınması yoluyla gerçekleştirilmektedir. Web sitelerinin derlenmesinde yayıncılardan veya hak sahiplerinden izin alma yoluna giden ulusal kütüphanelere ise Çek Cumhuriyeti, Avustralya ve Hollanda Ulusal Kütüphaneleri örnek verilebilir.

Web arşivleme faaliyetlerine başlayan ulusal kütüphaneler ilk olarak Web'i hangi ölçekte arşivleyeceğine karar vermelidir. Bu süreçte ulusal kütüphaneler belirledikleri hedefe göre yığın, alan adı, seçimli, olay eksenli veya tematik derleme stratejilerinden en az birini tercih etmektedir. Çalışmamızda ulusal kütüphanelerin en sık tercih ettiği derleme stratejilerinin sırasıyla seçimli derleme, olay eksenli derleme ve alan adı derleme olduğu sonucuna varılmıştır. Yığın derlemenin ulusal kütüphanelerde söz konusu derleme stratejilerine göre daha az tercih edilmesinin nedeni ise, ülkelerdeki diğer Web arşivleme girişimlerinin yığın derlemeye faaliyet süreçlerinde yer vermesi veya yığın derlemenin büyük ölçeği nedeniyle yüksek maliyetli olmasıdır.

Kültürel miras kurumları Web arşivleme sürecini doğru bir biçimde yönetme ve yapılandırma amacıyla Web arşivleme yaşam döngüsü aşamalarından yararlanabilirler. Web arşivleme yaşam döngüsünde ilk olarak değerlendirme ve seçim işlemleri kapsamında koleksiyona dâhil edilecek Web materyallerinin neler olduğu belirlenir. Sonrasında gerçekleştirilen sağlama işlemi ile materyallerin otomatik, manuel veya yasal derleme gibi süreçlerden hangisi ile derleneceği belirlenir ve materyaller sağlanır. Sağlama aşamasının ardından Web içeriğinin organizasyonuna ve depolanmasına yönelik adımlara ve tercih edilen kaynak türlerine göre üstveri oluşturma adımlarına yer verilir. Arşivlendiğinde yeni ve kalıcı bir URL'ye dönüşen Web sitesine kullanıcılar URL arama, anahtar sözcük arama, tam metin arama, alfabetik arama, gelişmiş arama ve konu araması yöntemlerinden biriyle erişim sağlar. Çalışmamızda yer verilen ulusal kütüphanelerin en sık tercih ettiği erişim yöntemleri sırasıyla URL arama, konu arama, anahtar kelime arama

ve tam metin aramadır. Eriřim ařamasında Web ařivleri kaynakları kamuyla paylařma d zeylerine g re karanlık ařiv, gri ařiv ve  evrimi i ařiv yaklařımlarından birini sergilerler. Ařivlenen kaynakların Web aray z nde, k t phane katalogunda veya keřif ara larında yer alması durumunda Web materyallerinin g r n rl klerini artmakta ve bu durum k lt rel mirasın korunmasını desteklemektedir. Web ařivleme yařam d ng s  ařamalarında son olarak veriye y nelik iřlemlerin ger ekleřtirildiđi ve g   ve  yk n m gibi stratejilerin devreye alındıđı korumaya y nelik iřlemlere yer verilir ve koleksiyona kazandırılan Web siteleri i in yařam d ng s  tamamlanır.  alıřmamızda sık kullanılan iř akıřlarını sunmak ve kurumlar i in uygulanabilir bir program oluřturarak s re  y netimini kolaylařtırmak amacıyla Archive-It tarafından oluřturulan Web Ařivleme Yařam D ng s  Modeli'ne yer verilmiřtir.

Web i eriđini yakalama, toplama, koruma ve ona eriřim sađlama s re lerinde ulusal k t phaneler ve diđer k lt rel miras kurumları  eřitli yazılım ara larından istifade ederler. A ık kaynak kodlu yazılımların Web ařivleme faaliyetlerinde yaygın olarak kullanımı yazılım maliyetlerini d ř rm řtir. Fakat se ilen yazılım aracının kullanılan ařiv sistemiyle uyumlu hale getirilmesinde teknik uzmanlık bilgisine sahip personele duyulan gereksinim olduk a fazladır.  alıřmamızda sađlama, k rasyon, yeniden oynatma, depolama ve muhafaza, arama ve keřif ve analiz alanlarında kullanılan yazılım ara ları ve bazı yardımcı uygulamalar, iřlevlerine iliřkin kısa a ıklamalarla birlikte sunulmuřtur.

Web ařivleme s recinde  eřitli standartların kullanımı uygulayıcıların aldıđı sonu ları etkilemektedir. Bu standartlara robot.txt protokol   rnek olarak verilebilir. Web sitelerinde robot.txt protokol n n kullanımı durumunda ilgili protokol geređi hedef Web sitesi yakalanamaz ve i eriđin tamamı veya bir b l m  ařivlenemez. Web sitesinin ařivlenememesi sonsuza dek kaybedilme riski altında olması sonucunu dođurmaktadır. Web ařivlerinin bir ođunda dosyaların g venli bir řekilde depolanması ve y netilmesi amacıyla 2009 yılında bir ISO standardı olarak kabul edilen WARC standardı kullanılmaktadır.  alıřmamızda ayrıca kaynaklar arasında bađlantılar oluřturan bir bađlantı tanımlayıcısı olan ve 2015 yılında bir ISO standardı olarak kabul edilen ISLI'ye yer verilmiřtir.  alıřma kapsamında ayrıca MARC21, ISAD(G) ve Dublin Core gibi  stveri standartlarına ve HTML, XML, CSS, JavaScript ve HTTP gibi temel Internet standartlarına ve son olarak OAIS Referans Modeli'ne deđinilmiřtir.

Web ařivleme s recinde uygulayıcı kurumların  eřitli alanlarda zorluklar yařadıđı bilinmektedir. Web sitelerinin kısa  m rleri ve karmařık yapıları nedeniyle ařivlenmesi, y netilmesi ve korunması zor nesneler oluřu yařanan zorlukların en bařında yer alır. Bununla birlikte Web sitelerinin i erisinde yer alan dijital nesnelerin t r ve format  eřitliliđinin fazla olması ařivleme uygulamalarını zorlařtırmaktadır. Kurumların telif ve m lkiyet hakları ve ařivlenecek Web sitelerinin se imi gibi konularında sorunlar

yaşadığı görülmüştür. Ayrıca Web mimarisinin sınırlılıkları, tarayıcı problemleri, veri tekilleştirme ve sosyal medya arşivleme gibi teknik konularda da zorluklar yaşandığı gözlemlenmiştir. Web arşivlerinde ihtiyaç duyulan teknik personelin niceliksel azlığı da kurumları bu alanda zorlayan bir diğer konudur.

Web arşivleme faaliyetlerinde istatistik ve kalite göstergelerinden yararlanılması kurumların aldığı sonuçları pozitif yönde etkilemektedir. Faaliyetlere ilişkin elde edilen istatistiksel veriler ve belirlenen kalite göstergeleri Web arşivleme uygulamalarının başarılı olup olmadığını değerlendirme ve kaynak planlaması yapma adına önemlidir. İstatistik ve kalite göstergelerinden yararlanılmasıyla genel bir performans ölçümü sağlanabilir ve Web arşivleme alanında sunulan hizmetlerin diğer kurumların sundukları hizmetlerle karşılaştırılması olanağı doğar. Çalışma kapsamında yer verilen istatistik ve kalite göstergeleri aracılığıyla koleksiyon geliştirme, koleksiyon nitelendirme, koleksiyon kullanımı ve maliyet gibi alanlarda istatistiksel ölçümlerin nasıl yapılacağına ilişkin bilgiler sunulmaktadır. Bununla birlikte çalışmamızda Web arşivleme uygulamalarına faaliyet süreçlerinde yer veren kurumların önceden belirlenen ihtiyaçları ne ölçüde karşıladıklarını ölçmelerinde kullanabilecekleri çeşitli kalite göstergelerine de yer verilmiştir.

Web arşivlemede faaliyet süreçlerinin değerlendirilmesi ve kurum hizmetlerinin karşılaştırılması adına işlevsellik konusu öne çıkmaktadır. Bu bağlamda arama parametreleri, erişim politikası, kişiselleştirilmiş hizmetler ve veri madenciliği gibi kurumların Web arşivleme uygulamalarında yer verdiği işlevsellik örnekleri çalışmamıza aktarılmıştır. Ayrıca IIPC tarafından hazırlanan Web Arşivleri için İşlevsellik Test Aracı da çalışmamızda sunulmuştur.

Dünya genelinde birçok ülkede Web arşivleme uygulamaları kurumlar tarafından veya ortaklıklar aracılığıyla sürdürülmektedir. Söz konusu kurum ve ortaklıklara ulusal kütüphaneler, ulusal arşivler, üniversiteler ve özel şirketler ile bu kurumlar tarafından oluşturulan konsorsiyumlar örnek verilebilir.

İlk Web arşivinin kuruluşundan bugüne yaşanan değişimleri anlamak dijital arşiv toplumunun geleceğe yön vermesi açısından oldukça önemlidir. Bu bağlamda, çalışmamızda Web arşivlemenin evrimine, yürütülen araştırma projelerine ve sürdürülen uluslararası Web arşivleme girişimlerine yer verilmiştir. Bununla birlikte çalışmamızın merkezinde yer alan ulusal kütüphane uygulamalarını değerlendirme amacıyla, 27 ulusal kütüphanenin Web arşivleme uygulamaları değerlendirilmiştir.

Çalışmamızın amacı kültürel mirasın korunması bağlamında dünya genelinde ulusal kütüphanelerin Web arşivleme uygulamalarının incelenmesi, karşılaştırılması ve Web arşivleme faaliyetlerinin Türkiye Milli Kütüphanesi'nde uygulanmasının gerekliliğini ortaya koymaktır. Bu doğrultuda çalışmamızda kültürel miras ve Web arşivleme konuları geniş bir perspektifle ele alınmış, ulusal kütüphanelerin Web arşivleme uygulamaları incelenerek karşılaştırılmış ve Türkiye Milli Kütüphanesi'ne yönelik bir değerlendirmede bulunulmuştur. Böylelikle çalışma kapsamında belirlenen araştırma amacına ulaşılmıştır. Bununla birlikte çalışma kapsamında belirlenen ve aşağıda sıralanan araştırmanın diğer amaçlarına da ulaşıldığı söylenebilir:

- Dünya genelinde ve ulusal kütüphaneler özelinde Web arşivleme uygulamalarının mevcut durumunu izlemek ve yönelimleri ölçmek,
- Gelecekte Türkiye Milli Kütüphanesi'nde Web arşivleme uygulamalarına yer verilmesi durumunda benimsenecek yaklaşımları ve atılacak adımları belirlemek,
- Web arşivleme sürecinde önemli görülen derleme stratejileri, yaşam döngüsü aşamaları, yazılımlar, standartlar, zorluklar, istatistik ve kalite göstergeleri gibi alanlarda açıklayıcı bilgiler sunmak ve değerlendirmeler yapmak,
- Web arşivlemeye ilişkin kapsamlı bir analiz gerçekleştirmek ve muhtemel uygulayıcılara bir öngörü kazandırmak,
- Web arşivleme konusunda Türkçe bilimsel literatürü desteklemek ve gelecekteki araştırmalara katkı sağlamak,
- Web arşivlemeye yönelik kavramları analiz etmek ve araştırmacıların kullanımına sunmak üzere önemli görülen kavramları içeren bir kavram listesi oluşturmak.

Çalışma kapsamında belirlenen araştırma problemlerinin yanıtlanması ve ortaya konulan hipotezlerin test edilmesinde betimleme yönteminden ve gözlem tekniğinden yararlanılmıştır. Kültürel mirasa ve Web arşivlemeye ilişkin ulusal ve uluslararası literatürün gözden geçirilmesi ve ulusal kütüphane uygulamalarının karşılaştırılması neticesinde elde edilen veriler ve ortaya konulan bulgular, çalışma kapsamında belirlenen hipotezlerin tamamının doğrulandığını göstermektedir. Araştırma kapsamında elde edilen sonuçlarla bağlantılı olarak sunulan çeşitli önerilere aşağıda yer verilmektedir:

- Web arşivleme uygulamalarına faaliyet süreçlerinde yer veren kurumlar iş süreçlerinin ve planlamaların yansıtıldığı bir program ve strateji dokümanı oluşturmalıdır. İlgili dokümanın yıllık olarak güncellenmesi ve içerisinde bütçe, amaçlar, hedefler, kapsam, sınırlılıklar, iş akışları, hak ve izinler, personel bilgileri ve iş birliği yapılacak kurumlara yer verilmesi yararlı olacaktır.

- Web arşivlemede seçilen derleme stratejisi koleksiyonun kapsamını belirlemede oldukça etkilidir. Eğer seçimli, tematik veya olay eksenli derleme tercih edilecekse, yakalanacak Web içeriğini seçme, kataloglama ve ona üstveri atama süreçlerinde küratörlerden destek alınmalıdır. Bununla birlikte, Web arşivleme faaliyetlerine yeni başlayan kurumlar eğer teknik açıdan eğitilmiş personele sahip değilse diğer kurumlarla iş birliği yapmalı veya dışarıdan hizmet alımı yoluna gitmelidir. Kurulacak iş birlikleri ve dışarıdan hizmet alımları kurumun Web arşivleme sürecinde deneyim kazanmasını destekleyecektir.
- Web arşivleme sürecinde kurumlar ihtiyaç duydukları alanlarda yaygın kullanıma sahip açık kaynak kodlu yazılımları tercih edebilir. Kurumların diğer kurumların farklı iş akışlarında hangi yazılımlardan yararlandığını öğrenmesi tercihlerini etkileyebilir ve bu tarafımızca faydalı görülmektedir. Diğer kuruluşların farklı iş akışları için ne kullandıklarını öğrenme sürecinde IIPC Web sitesinden yararlanılabilir [138].
- Web arşivleme süreci sağlanan Web sitesinin türüne göre oldukça karmaşık hale gelebilir. Bunun nedeni Web'in derin Web olarak da ifade edilen veri tabanlarını, şifreli içeriği ve sosyal medya gibi arşivlenmesi zor alanları içerisinde barındırmasıdır. Web arşivleri söz konusu alanlarda yaşanan problemleri giderebilmek adına Web arşivleme alanındaki gelişmeleri yakından takip etmelidir.
- Arşivlenen bir Web sitesinin içeriğinin kontrol edilmesi neyin derlendiğini gözden geçirebilmeyi sağlar. İçerik kontrolü ayrıca mükerrer kopyaların koleksiyonda ne ölçüde yer aldığını belirleme adına da önemlidir. Web arşivleri koleksiyon bütünlüğünü sağlama ve mükerrerliğin oluşmasını engelleme amacıyla derlenen materyallere yönelik kontrolleri periyodik olarak gerçekleştirmelidir.
- Erişim arayüzü Web arşivlerinin kullanımında oldukça belirleyicidir. Web arşiv arayüzü tasarlanmadan önce arayüzün hangi erişim yöntemlerini ve hangi dilleri destekleyeceği belirlenmelidir. Ayrıca Web arşivinin görünürlüğü artırarak amacıyla, arayüzünün kurumun dijital koleksiyonlarıyla veya belirlenen keşif araçlarıyla entegre edilmesi desteklenmelidir.
- Web arşivlerinde arşivlenen veri miktarının sürekli artışı erişim sürecinde arzu edilen sonuçlara ulaşmayı güçleştirebilir. Bu nedenle kurumların büyük boyuttaki verilere erişme amacıyla daha etkili arama teknolojilerini takip etmesi, desteklemesi ve söz konusu alanlara daha fazla kaynak ayırması tavsiye edilmektedir.

- Web arşivlerinin kullanılabilirliği eksiksizliği ve tanıtımına bağlıdır. Web arşivlemede uygulayıcılara değerlendirme yapma olanağı sunan kalite göstergeleri alanında teknik ilerlemenin az olduğu söylenebilir. Kalite göstergeleri alanında daha geçerli değerlendirme yapma olanağı sağlayan araçlar geliştirmek için daha fazla dikkat gösterilmeli ve kaynak ayrılmalıdır.
- Gelecekte analiz araçları yardımıyla Web arşivlerinde bulunacak muazzam boyuttaki verinin yalnızca insan-insan etkileşimine değil insan-makine ve makine-makine etkileşimine kaynaklık edebileceği düşünülmektedir. Bununla birlikte söz konusu arşiv verilerinin veri madenciliğini ve büyük veri uygulamalarını destekleyeceği ön görülmektedir. Bu durumda Web arşivleme faaliyetlerine yönelik araştırma geliştirme çalışmalarına kurumların daha fazla katılım göstermesi önerilmektedir.
- Web arşivleme gün geçtikçe dünya genelinde daha fazla ilgi görmekte ve tanınırlık kazanmaktadır. Fakat Web arşivleme topluluğunda en son gelişmeler hakkında bilgi eksikliği vardır. Kurumların güncel bilgiye erişim için gelişim trendlerini daha yakından izlemesi önerilmektedir.
- Başta Web olmak üzere büyük miktardaki dijital veriyi arşivleme yükümlülüğü birçok ülkede ulusal kütüphanenin bu görevi yerine getirmesine izin veren bir telif hakkı yasası ve bir derleme yasası oluşturulmasını gerekli kılmıştır. Fakat oluşturulan yasaların Web kaynaklarını tam olarak sahiplenmemesi, Web arşiv koleksiyonlarının kapsamının daralmasına neden olmaktadır. Yasal derleme ve telif hakları mevzuatlarının toplayıcı kuruma toplanabilecek bütün Web sitelerini toplama yetkisi verecek şekilde düzenlenmesi ve Web arşivlemeyi kolaylaştırıcı hükümleri barındırması gerekmektedir. Yasal derleme ve telif hakları mevzuatlarının korunan nesnelere ve bundan kimlerin sorumlu olduğuna karar vermede şekillendirici bir kriter haline geldiği günümüzde, ihtiyaçlar ve gerekliliklere yönelik olarak yöneticilerin ve uzmanların politikacıları bilgilendirmeleri ve onlardan destek almaları önerilmektedir.
- Web arşivleme uygulamalarına faaliyet süreçlerinde yer veren kurumların yayıncılar ve hak sahipleriyle anlaşma yoluna gitmeleri ve uluslararası projeler dâhilinde kurumlar arası iş birliği faaliyetlerinde yer almaları, Web arşivleme uygulamaları hızlı bir biçimde yaygınlaştıracak ve kültürel miras ürünlerinin korunması ve paylaşımı kolaylaşacaktır. Bu durumda kurumlar arasında en iyi uygulamalar ve ortak çalışmalar teşvik edilmelidir.

- Ulusal kütüphanelerin oluřturdukları Web arřivleri ülkelerin kültürel mirasının bir parçası olarak görölmektedir. Bu nedenle dijital kültürel miras varlıkları uzun vadeli bir koruma stratejisi ile korunmalıdır. Dijital koruma hedeflerine ulaşma adına Web arřivleri tarafından göç, öykünüm, yedekleme gibi stratejiler periyodik olarak devreye alınmalıdır.
- Derleme yasaları Web arřivleme alanında toplayıcı kurumların ne ölçüde yetkili olduđunu belirler. Ülkemizdeki derleme yasasının Türkiye Milli Kütüphanesi'nin Web kaynaklarını derleme hakkına sahip olduđunu açıkça belirtilecek şekilde yeniden düzenlenmesi önerilmektedir. Söz konusu mevzuatta “derlenecek Web sitelerinin yayıncıları aracılığıyla Türkiye Mili Kütüphanesine gönderilmesi sağlanmalıdır” şeklinde bir hükmün yer alması Web arřivleme faaliyetlerini hızlandırabilir.
- Türkiye Milli Kütüphanesinde Web arřivleme uygulamalarına yer verilecekse derleme yaklaşımları arasında öncelikli olarak ulusal kültürel miras nesnelerini barındıran Web sitelerini derlemek için adı derlemeye ve seçimli derlemeye yer verilmelidir. Alan adı derlemede ilk olarak “.tr”, “org.tr” ve “edu.tr” uzantılarının yılda 2 veya 3 kez taranması tarafımızca önerilmektedir. “.tr” alan adı tahsis hizmeti Orta Dođu Teknik Üniversitesi (ODTÜ) tarafından verildiğinden ilgili Kurum'dan arřivlenecek alan adlarının bir listesi elde edilebilir. Seçimli derlemede bilgi ve kültür mirası ürünlerini içinde barındıran Web sitelerinin kurum kütüphanecileri, küratörleri veya konu uzmanlarınca belirlenerek toplanması ve arřivlenmesi önerilir. Bir sonraki aşamada ise olay eksenli derlemeye yer verilebilir ve arřivlenecek içeriğı seçme aşamasında toplumun yoğun ilgi gösterdiği konular tercih edilebilir.

KAYNAKÇA

Aksoy, A. ve Ünsal, D. (Ed.) (2012). Kültürel miras yönetimi. T.C. Anadolu Üniversitesi Yayını No: 2573 ve Anadolu üniversitesi yayını No: 1543. Erişim adresi: <http://kitap.okur-yazar.net/e-kitap/aof/KMT201U-kulturel-miras-yonetimi.pdf>

Aldemir, A. ve Oğuz, E. S. (2006). Sayısal (dijital) kültürün korunması: Web arşivleme. *Türk Kütüphaneciliği* 20(3), 283-312. Erişim adresi: <http://www.tk.org.tr/index.php/TK/article/view/72/72>

Arioğlu İ. E. ve Aydoğdu Atasoy Ö. (2015). Somut olmayan kültürel miras kapsamında geleneksel el sanatları ve Kültür ve Turizm Bakanlığı. *International Periodical for the Languages, Literature and History of Turkish or Turkic* 10(16), 109-126. Erişim adresi: <http://www.turkishstudies.net/DergiTamDetay.aspx?ID=9058>

Bailey, J., Grotke, A., McCain, E., Moffatt, C. ve Taylor, N. (2017). Web archiving in the United States: A 2016 survey (An NDSA report). Erişim adresi: http://ndsa.org/documents/WebArchivingintheUnitedStates_A2016Survey.pdf

Başlar, K. (2012). Kırkıncı yılında Dünya Mirası Sözleşmesi (1972-2012). *Uluslararası Hukuk ve Politika* 7(27), 25-52. Erişim adresi: http://usak.org.tr/images_upload/files/uhp%2027_2.pdf

Bayter, M. (2008). Türkçe Web belgelerinin kataloglanması: Bir işbirliği modeli önerisi (Yayınlanmamış doktora tezi). Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara. Erişim adresi: <http://docplayer.biz.tr/6620880-Turkce-web-belgelerinin-kataloglanmasi-doktora-tezi-mustafa-bayter.html>

5632 sayılı Milli Kütüphane Kuruluşu Hakkında Kanun. (1946, 29 Mart). *T. C. Resmî Gazete* (7469). Erişim adresi: <http://www.mevzuat.gov.tr/MevzuatMetin/1.3.5632.pdf>

Beunen, A. ve Schiphof, T. (2006). Legai aspects of Web archiving from a Dutch perspective. Rapor. Hague, Hollanda. Erişim adresi: http://www.kb.nl/sites/default/files/docs/KB_Legal_Aspects_WebArchiving_EN.pdf

Bhat, M.H., (2009). Missing Web references: A case study of five scholarly journals. *LIBER Quarterly*. 19(2), 131–139. DOI: <http://doi.org/10.18352/lq.7957>

Bragg, M. ve Reed, S. (2014). ARIH Glossary of Web Archiving Terms. 1-3. Çevrimiçi ön yayın.

<https://webarchive.jira.com/wiki/display/ARIH/Glossary+of+Web+Archiving+Terms>

Brügger, N. (2005). Archiving Websites: general considerations and strategies. 1-64. Århus, Danmarka: The Centre for Internet Research. Erişim adresi: http://cfi.au.dk/fileadmin/www.cfi.au.dk/publikationer/archiving_underside/archiving.pdf

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2008). Bilimsel araştırma yöntemleri. Ankara: Pegem Akademi.

Cadavid, J., A., P. (2014). Copyright challenges of legal deposit and Web archiving in the National Library of Singapore. *Alexandria*. 25(1-2), 1-19. <http://journals.sagepub.com/doi/10.7227/ALX.0017>

Costa, M., Gomes, D. ve Silva, M. J. (2016). The evolution of Web archiving. Berlin, Almanya: *International Journal on Digital Libraries*. 18(3), 1-16. Erişim adresi: The_evolution_of_web_archiving/links/5767cd1b08ae421c448dc558/The-evolution-of-web-archiving.pdf

Council of Europe. (2005). The framework convention on the value of cultural heritage for society. *Council of Europe Treaty Series* (No: 199) 1-9. Erişim adresi: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680083746>

Daigle, L., van Gulik, D., Iannella, R. ve Faltstrom, P. (2002). Uniform Resource Names (URN) namespace definition mechanisms. Erişim adresi: <http://dl.acm.org/citation.cfm?id=RFC3406>

Dougherty, M., Meyer, E.T., Madsen, C., van den Heuvel, C., Thomas, A. ve Wyatt, S. (2010). Researcher engagement with Web archives: State of the art. London: JISC. Erişim

adresi:

http://ecommons.luc.edu/cgi/viewcontent.cgi?article=1014&context=communication_fa_cpubs

4848 Sayılı K lt r ve Turizm Bakanlıęı Teşkilat ve G revleri Hakkında Kanun. (2003, 16 Nisan). T. C. *Resm  Gazete* (25093). Eriřim adresi: <http://www.mevzuat.gov.tr/MevzuatMetin/1.5.4848.pdf>

Fabian, C. (2014). Cultural heritage in the digital world: From the treasure house to global traveller. Uluslararası K lt rel Mirasın ve K lt rel Bellek Kurumlarının Y netimi Kongresi'nde sunulan Bildiri (17-20 Eyl l 2014). 160-166. Ankara:  niversite ve Arařtırma K t phanecileri Derneęi. Eriřim adresi: <https://www.unak.org.tr/upload/userfiles/files/UNAK2014bildiri.pdf>

Fetterly, D., Manasse, M., Najork, M. ve Wiener, J. L. (2003). A large-scale study of the evolution of Web pages. *Software- Practice And Experience*. 1(1), 1-27. Eriřim adresi: https://www.researchgate.net/profile/Mark_Manasse/publication/200111282_A_large-scale_study_of_the_evolution_of_Web_pages/links/00b4952279706941a3000000.pdf

Galatasaray  niversitesi ve International Committee of the Red Cross. (t.y.). 12 Aęustos 1949 tarihli Cenevre S zleşmeleri ve ek protokolleri. (1949). Ed. M. B. Yamaner ve dięerleri. 1-306. İstanbul: Galatasaray  niversitesi Hukuk Fak ltesi yayınları No: 42, Eriřim adresi: <https://www.icrc.org/eng/home/languages/turkish/files/sozlesmeleri-protokolleri-conventions-protocols.pdf>

Glanville, L. (2010). Web archiving: ethical and legal issues affecting programmes in Australia and the Netherlands. *The Australian Library Journal*. 59(3), 128-134. <http://dx.doi.org/10.1080/00049670.2010.10735999>

Gomes D., Miranda J. ve Costa M. (2011). A survey on Web archiving initiatives. International Conference on Theory and Practice of Digital Libraries, Bildiri. 408-420. Eriřim adresi: https://doi.org/10.1007/978-3-642-24469-8_41

Gomes, D., Santos, A. L. ve Silva, M. J. (2006, Nisan). Managing duplicates in a Web archive. SAC 2006 Konferansı bildirisi, Dijon, Fransa.  evrimi i  n yayın. Eriřim adresi: <http://visibilidade.net/daniel/docs/papers/gomesManaging.pdf>

Gökkurt, Ö. ve Demirtel, H. (2014). Sosyal medyada kurumsal arşivleme sorunu. Uluslararası Kültürel Mirasın ve Kültürel Bellek Kurumlarının Yönetimi Kongresi, Bildiri. (226-234). Üniversite ve Araştırma Kütüphanecileri Derneği. Urla, Türkiye. Erişim adresi: <http://ozlemgokkurt.com/sosyal-medyada-kurumsal-arsivleme-sorunu/>

Grotke, A. (2011). Web archiving at the Library of Congress. *Computers in Libraries*. 31(10), 15-19. Erişim adresi: <http://www.infotoday.com/cilmag/dec11/Grotke.shtml>

Grotke, A. ve Jones, G. (2010). DigiBoard: A tool to streamline complex Web archiving activities at the Library of Congress. 10th International Web Archiving Workshop bildirisi. Vienna, Avusturya. Erişim adresi: <http://iawaw.europarchive.org/10/IWAW2010.pdf>

Guenther, R. ve Myrick, L. (2007). Archiving Web sites for preservation and access: MODS, METS and MINERVA. *Journal of Archival Organization*, 4:1-2, 141-166. http://dx.doi.org/10.1300/J201v04n01_08

Güzel, M. (2006). Küreselleşme, Internet ve gençlik kültürü. *Küresel İletişim dergisi* 1(1) 1-16. Erişim adresi:

Hakala, J. (2004). Archiving the Web: European experiences, *Program*, 38, 176-183. <http://dx.doi.org/10.1108/00330330410547223>

Hayırsever Baştürk F. ve Dimitriyadis I. (2012). Ekonomik tetikliyeciler olarak sigorta ve kültürel miras değerleri. *Organizasyon ve Yönetim Bilimleri Dergisi* 4(1), 107-116 Erişim adresi: http://www.sobiad.org/ejournals/dergi_ybd/arsiv/2012_1/feride_hayirsever_basturk.pdf

Heydon, A. ve Najork, M. (1999). Mercator: A scalable, extensible Web crawler. 2(4), 219-229. <https://doi.org/10.1023/A:1019213109274>

Hockx-Yu, H., Crawford, L., Coram, R ve Johnson, S. (2010). Capturing and replaying streaming media in a Web archive: A British Library case study. 1-5. Çevrimiçi ön yayın. Erişim adresi: <http://www.ifs.tuwien.ac.at/dp/ipres2010/papers/hockxyu-44.pdf>

Huurdeman, H. C., Ben David, A. ve Samar, T. (2013). Sprint methods for Web archive research. 182-190. WebSci '13: 5th Annual ACM Web Science Conference, Bildiri. ACM New York, NY, ABD. Erişim adresi: <https://doi.org/10.1145/2464464.2464513>

International Organization for Standardization. (2009). BS ISO 28500:2009 Information and documentation: WARC file format. Cenova, İsviçre: International Organization for Standardization (ISO). Erişim adresi: <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/000000000030167516>

International Organization for Standardization. (2012a). ISO/DTR 14873 Information and documentation: Statistics and quality issues for Web archiving (taslak teknik rapor). Cenova, İsviçre: International Organization for Standardization (ISO). Erişim adresi: http://netpreserve.org/resources/IIPC_project-SO_TR_14873_E_2012-10-02_DRAFT.pdf

International Organization for Standardization. (2013). PD ISO/TR 14873:2013(E) Information and documentation: Statistics and quality issues for Web archiving (resmî belge). Cenova, İsviçre: International Organization for Standardization (ISO). 1-54. Erişim adresi: <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/000000000030274605>

International Organization for Standardization. (2015). BS ISO 17316:2015 Information and documentation: International standard link identifier (ISLI). Cenova, İsviçre: International Organization for Standardization (ISO). Erişim adresi: <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/000000000030246898>

International Organization for Standardization. (2014). ISO/TR 17797:2014 Electronic archiving: selection of digital storage media for long term preservation. Cenova, İsviçre: International Organization for Standardization (ISO). 1-26. Erişim adresi: <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/000000000030281201>

International Organization for Standardization (2012b). BS ISO 14721:2012 Space data and information transfer systems: Open archival information system (OAIS) Reference model. 1-135. Cenova, İsviçre: International Organization for Standardization (ISO). <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/000000000030234187>

Internet Archive, The Archive-It Team. (2013). The Web archiving life cycle model. San Francisco, CA: Internet Archive. Erişim adresi: http://ait.blog.archive.org/files/2014/04/archiveit_life_cycle_model.pdf

2863 sayılı K lt r ve Tabiat Varlıkları Koruma Kanunu. (1983, 23 Temmuz). *T.C. Resm  Gazete* (18113). Eriřim adresi: <http://www.mevzuat.gov.tr/MevzuatMetin/1.5.2863.pdf>

İslamođlu, H. (2009). Sosyal bilimlerde arařtırma y ntemleri. İřletme- Ekonomi Dizisi: 332, Yayın No: 2153. İzmit: Beta. ISBN 978-605-377-071-8

Kanun H km nde Kararname. (1984, 18 Nisan). Sayı: 227, *T.C. Resm  Gazete* (18435). Eriřim adresi: <http://www.tskdv.org.tr/p01/227khk.pdf>

Karatař Ateř, E. (2015). Dijitalleřtirme s reci ve dijital k t phane uygulamalarına bir bakıř. *International Periodical For The Languages, Literature And History Of Turkish or Turkic* 10(10), 561-574. Eriřim adresi: http://turkishstudies.net/Makaleler/1136638971_28Karata%C5%9FAte%C5%9FEsmeray-sos-561-574.pdf

Kim, Y. ve Ross, S. (2012). Digital forensics formats: Seeking a digital preservation storage container format for Web archiving. Edinburg, İsko ya: *The International Journal of Digital Curation*. 7(2), 21–39. Eriřim adresi: <http://dx.doi.org/10.2218/ijdc.v7i2.227>

Kirschenbaum, M. G., Ovenden, R. ve Redwine, G. (2010). Digital forensics and born-digital content in cultural heritage collections. Council on Library and Information Resources. 1-93. Washington, Amerika Birleřik Devletleri. Eriřim adresi: <https://www.clir.org/pubs/reports/pub149/pub149.pdf>

Lasfargues, F., Oury, C., ve Wendland, B. (2008). Legal deposit of the French Web: Harvesting strategies for a national domain. IAWAW '08: 8th International Workshop for Web Archiving, Bildiri. 1-14. Aarhus, Danimarka. Eriřim adresi: <https://hal.archives-ouvertes.fr/hal-01098538/document>

Lasfargues F., Chlo  Martin C. ve Medjkoune, L. (2012). Archiving before losing valuable data?: Development of Web archiving in Europe. *Bibliothek Forschung und Praxis*. 36, 118-125. Eriřim adresi: <https://www.degruyter.com/view/j/bfup.2012.36.issue-1/bfp-2012-0014/bfp-2012-0014.pdf>

Lawrence, S., Coetzee, F., Glover, E., Pennock, D., Flake, G., Nielsen, F., Krovetz, R., Kruger, A., Giles. C. L. (2001). Persistence of Web references in scientific research. *IEEE*

Computer. 34(2), 26–31. Erişim adresi: <http://lexicalresearch.com/persistence-computer01.pdf>

Lecher, H. E. (2006). Small scale academic Web archiving: DACHS. In *Web Archiving* 213-226. Berlin: Springer Verlag. Erişim adresi: https://link.springer.com/chapter/10.1007%2F978-3-540-46332-0_10

Lyman, P. ve Varian, H. R. (2003). "How much information", Araştırma Raporu. Çevrimiçi ön yayın. Erişim adresi: http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf

Masanés, J. (Ed.) (2006). *Web archiving*. Berlin: Springer. 1-53. DOI: [10.1007/978-3-540-46332-0](https://doi.org/10.1007/978-3-540-46332-0)

McCown, F., Chan, S., Nelson, M. L. ve Bollen, J. (2005). The availability and persistence of Web references in D-Lib Magazine. *5th International Web Archiving Workshop (IWAW05)*, Viyana, Avusturya. Çevrimiçi ön yayın. Erişim adresi: <https://arxiv.org/abs/cs/0511077v1>

Ministry of Culture of the People's Republic of China. (2016). Interim regulations on the administration of Internet culture. *Chinese Law And Government*. 48(2), 134–141. Erişim adresi: <http://www.tandfonline.com/doi/full/10.1080/00094609.2015.1048160>

Murray, K. R. & Hsieh, I. K. (2008). Archiving Web-published materials: A needs assessment of librarians, researchers, and content providers. *Government Information Quarterly* 25(1). <https://doi.org/10.1016/j.giq.2007.04.005>

National archives. (2011). *Web Archiving Guidance*. Birleşik Krallık: National Archives. Erişim adresi: <https://nationalarchives.gov.uk/documents/information-management/web-archiving-guidance.pdf>

National Archives. (2017). *The UK Government Web Archive: Guidance for digital and records management teams*. Erişim adresi:

<http://www.nationalarchives.gov.uk/documents/web-archiving-technical-guidance.pdf>

Nelson, M. L., Sanderson, R., Balakireva, L., Shankar, H. (2009). Memento: time travel for the Web. Çevrimiçi ön yayın. Erişim adresi: <https://www.researchgate.net/publication/45882713>

Niu, J. (2012a). Functionalities of Web Archives. *D-Lib Magazine*, 18(3-4). Erişim adresi: <http://www.dlib.org/dlib/march12/niu/03niu2.print.html>

Niu, J. (2012b). An Overview of Web Archiving. *D-Lib Magazine*, 18(3-4). Erişim adresi: <http://www.dlib.org/dlib/march12/niu/03niu1.print.html>

Noronha, N., Campos, J. P., Gomes, D., Silva, M. J. ve Borbinha, J. (2001). A deposit for digital collections. 5td. *European Conf. Research and Advanced Technology for Digital Libraries*. Bildiri. 200–212. Erişim adresi: https://link.springer.com/chapter/10.1007/3-540-44796-2_18

Ntoulas A., Cho J., ve Olston C. (2004). What's new on the Web? The evolution of the Web from a search engine perspective. 13th international conference on World Wide Web, Bildiri. 1-12. ACM Press. Erişim adresi: <http://infolab.stanford.edu/~olston/publications/webstudy.pdf>

Oğuz, E. S. (2006a). Sayısal kültürün korunması: Web arşivleme. *Türk Kütüphaneciliği* 20(3), 283-312. Erişim adresi: <http://www.tk.org.tr/index.php/TK/article/view/72/72>

Oğuz, E. S. (2006b, Ekim). Web arşivleme yaklaşımları ve örneklerle Web arşivleri. *ÜNAK'06 Bilimsel İletişim ve Bilgi Yönetimi Sempozyumu* sunumu. 1-48. Ankara. Erişim adresi: <https://www.slideshare.net/esinsultan/web-arivleme-yaklamalar-ve-rneklerle-web-arivleri>

Oğuz, E. S. (2011a). Toplum bilimlerde kültür kavramı. *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, 28(2), 123-139. Erişim adresi: http://www.bby.hacettepe.edu.tr/e-bulten/dosyalar/file/mart2012/esinsultan_efd.pdf

Oğuz, E. S. (2011b). Kültür politikaları ve kütüphaneler: kuramsal bir yaklaşım. *Hacettepe Üniversitesi Sosyolojik Araştırmalar E-Dergisi*. 1-31. Erişim adresi: <http://www.sdergi.hacettepe.edu.tr/makaleler/EsinSultanoguzmak.pdf>

Ölçer Özünel, E. (2013). Yeni miraslar ve uluslararası sözleşmelerde sürdürülebilir kalkınma stratejileri. *Milli Folklor*, 25(100), 14-30. Erişim adresi: <http://www.millifolklor.com/tr/sayfalar/100/02-.pdf>

Özbağ D. (2010). Ulusal dijital kültür mirasının korunması ve arşivlenmesine yönelik kavramsal bir model önerisi (Yayınlanmamış Yüksek Lisans Tezi). Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü. Erişim adresi: <http://www.bby.hacettepe.edu.tr/yayinlar/dosyalar/ozbag.pdf>

Özel, N. (2014). Kültürel miras varlıklarının dijitalleştirilmesi sürecinde planlama çalışmalarının ve üstveri tanımlamalarının önemi. N. Özel ve T. Çakmak (Ed.). *Uluslararası Kültürel Mirasın ve Kültürel Bellek Kurumlarının Yönetimi Kongresi* (17-20 Eylül 2014) bildirisi. 160-166. Ankara: Üniversite ve Araştırma Kütüphanecileri Derneği. Erişim adresi: <https://www.unak.org.tr/upload/userfiles/files/UNAK2014bildiri.pdf>

Patel, K. (2007). Authors v. archivers: the copyright infringement battle over Web pages. *Journal of the Patent and Trademark Office Society*, 89(5), 410–428. Erişim adresi: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=965199

Pennock, M. (2013). Web-archiving (DPC Technology Watch Report 13-01 March 2013). 1-45. Birleşik Krallık: Digital Preservation Coalition. <http://dx.doi.org/10.7207/twr13-01>

Pietrobruno, S. (2013). YouTube and the social archiving of intangible heritage. *New Media & Society* 15(8), 1259–1276. Erişim adresi: <http://nms.sagepub.com/content/early/2013/01/08/1461444812469598>

Reitz, J. M. (2002). Online Dictionary of Library and Information Science (ODLIS). 1-733. Erişim adresi <http://vlado.fmf.uni-lj.si/pub/networks/data/dic/odlis/odlis.pdf>

Ruusalepp, R. (2003). AHDS Digital Preservation Glossary, 1-13. Erişim adresi: <http://docplayer.net/12899990-Ahds-digital-preservation-glossary.html>

SalahEldeen, H. ve Nelson, M. (2012). Losing my revolution: how many resources shared on social media have been lost? *16th International Conference on Theory and Practice*

of *Digital Libraries*, Bildiri. 125–137. Erişim adresi: <http://www.cs.odu.edu/~mln/pubs/tpdl-2012/tpdl-2012.pdf>

Stirling, P. ve Illien, G. (2011). The state of e-legal deposit in France: looking back at five years of putting new legislation into practice and envisioning the future. 77th IFLA general conference, bildiri. Erişim adresi: <https://www.ifla.org/past-wlic/2011/193-stirling-en.pdf>

Subaşıoğlu, F. (2001). Dijital kütüphaneler: tanımlama sorunu. *Türk Kütüphaneciliği* 15(1), 45-54. Erişim adresi: http://yunus.hacettepe.edu.tr/~soyda/bby153_2012/4&5/DijitalKutuphaneler-Subasioglu.pdf

Şencan, İ. ve Doğan, G. (Haz.). (2017). Bilimsel Yayınlarda Kaynak Gösterme, Tablo ve Şekil Oluşturma Rehberi: APA 6 Kuralları. Ankara: Türk Kütüphaneciler Derneği. Erişim adresi: http://www.tk.org.tr/APA/apa_2.pdf

The National Library Of China. (2009). 2009 CDNLAO Country Report. 2009 CDNLAO Konferansı'nda yayımlanan rapor. Hanoi, Vietnam: Conference of Directors of National Libraries in Asia and Oceania. Erişim adresi: http://www.ndl.go.jp/en/cdnla0/meetings/pdf/CR2009_Fiji.pdf

Tonta, Y. (2002). Dijital hizmetlere ve kaynaklara erişim. PULMAN-XT Türkiye Ulusal Toplantısı Bildirisi (16-19 Kasım 2002). Ankara: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü. Çevrimiçi ön yayın. Erişim adresi: <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/pulman-dijital-hizmetler.pdf>

Tonta, Y., Bitirim, Y. ve Sever, H. (2002). Türkçe arama motorlarında performans değerlendirme. Total Bilişim Ltd. Şti., Ankara. Erişim adresi: <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-bitirim-sever-arama-motorlari.pdf>

Tonta, Y. (2009). Dijital yerliler, sosyal ağlar ve kütüphanelerin geleceği. *Türk Kütüphaneciliği*, 23(4), 742–768. Erişim adresi: <http://www.tk.org.tr/index.php/TK/article/view/534/527>

Tonta, Y. (2014). Kültürel mirasın geleceği. *Uluslararası Kültürel Mirasın ve Kültürel Bellek Kurumlarının Yönetimi Kongresi* bildirisi (17-20 Eylül 2014). 1-13. Ankara:

Üniversite ve Araştırma Kütüphanecileri Derneği. Erişim adresi:
<https://www.unak.org.tr/upload/userfiles/files/UNAK2014bildiri.pdf>

United Nations Educational, Scientific and Cultural Organization. (1974). Universal bibliographic control, D. Anderson (Haz.). *Intergovernmental Conference On The Planning Of National Documaentation, Library and Archives Infrastructures* bildirisi. (23-27 Eylül 1974). 1-67. Paris. Erişim adresi:
<http://unesdoc.unesco.org/images/0000/000098/009864EB.pdf>

United Nations Educational, Scientific and Cultural Organization ve International Federation of Library Associations and Institutions. (1977). International congress on national bibliographies, rapor, 12–15 September 1977. Paris: UNESCO. Erişim adresi:
<http://unesdoc.unesco.org/images/0003/000300/030062eo.pdf>

United Nations Educational, Scientific and Cultural Organization. (2003). *Charter on the preservation of digital heritage: 32nd general conference*. 1-196. Paris, Fransa. Erişim adresi: <http://unesdoc.unesco.org/images/0013/001331/133171e.pdf#page=80>

United Nations Educational, Scientific and Cultural Organization. (2013). The Hangzhou Declaration: Placing culture at the heart of sustainable development policies. *Culture: Key to Sustainable Development* uluslararası kongresi bildirgesi. (1–6 May 2013). China. Erişim adresi:
<http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/images/FinalHangzhouDeclaration20130517.pdf>

United Nations Educational, Scientific and Cultural Organization. (2015). Draft recommendation concerning the preservation of, and access to, documentary heritage including in digital form. Taslak önerisi. Paris, Fransa. Erişim adresi:
<http://unesdoc.unesco.org/images/0023/002339/233916e.pdf>

United Nations Educational, Scientific and Cultural Organization. (2001). Kültürel çeşitlilik evrensel bildirgesi. Yeliz Özay (çev.). Erişim adresi:
http://www.unesco.org.tr/dokumanlar/kulturel_ifadelerin_cesitliliği/EVRENSEL_B%C4%B0LD%C4%B0RGE.pdf

United Nations Educational, Scientific and Cultural Organization. (1982). World Conference on Cultural Policies: Mexico City Declaration on Cultural Policies (26 Temmuz - 6 Ağustos 1982). 1-6. Çevrimiçi ön yayın. Erişim adresi:

http://portal.unesco.org/culture/en/files/12762/11295421661mexico_en.pdf/mexico_en.pdf

Zan, B. U., (2009). Elektronik yayınların derlenmesi: Legal deposit of electronic publications. *Türk Kütüphaneciliği* 23(2) (s. 299-320). Erişim adresi: https://www.researchgate.net/profile/Burcu_Zan/publication/47334316_Legal_Deposit_of_Electronic_Publications_Elektronik_Yayınların_Derlenmesi/links/56994f8e08aeeea985946e29/Legal-Deposit-of-Electronic-Publications-Elektronik-Yayınların-Derlenmesi.pdf

Zan, B. U., (2006). Derleme olgusu ve elektronik yayınlar. (Yayımlanmamış Yüksek Lisans Tezi). Ankara: Ankara Üniversitesi Sosyal Bilimler Enstitüsü. Erişim adresi: https://www.researchgate.net/profile/Burcu_Zan/publication/279444564_DERLEME_OLGUSU_VE_ELEKTRONIK_YAYINLAR/links/56994ef108aeeea985946deb/DERLEME-OLGUSU-VE-ELEKTRONIK-YAYINLAR.pdf

WEB KAYNAKLARI

- [1] https://www.wikizero.com/tr/World_Wide_Web
- [2] <http://www.worldwidewebsite.com/>
- [3] <https://www.deepweb-sites.com/how-big-is-the-deep-web/>
- [4] <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=21779>
- [5] http://turk_halkbilimi.gazi.edu.tr/posts/view/title/somut-olmayan-kulturel-miras-sozlesmesi-12924
- [6] <http://mevzuat.meb.gov.tr/html/tcanayasasi2.html>
- [7] <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/images/FinalHangzhouDeclaration20130517.pdf>
- [8] <http://archivemati.ca/2005/11/08/digital-archives/>
- [9] <https://support.office.com/tr-tr/article/Ba%C5%9Fka-konumlardan-federe-arama-sonu%C3%A7lar%C4%B1n%C4%B1-birle%C5%9Ftirme-486fc16d-f566-4e10-b39a-4563573fe94b>)
- [10] <https://en.wikipedia.org/wiki/ARPANET>
- [11] <http://www.internetlivestats.com/total-number-of-websites/>
- [12] <https://www.cnnturk.com/2009/bilim.teknoloji/teknoloji/04/02/web.1.0.web.2.0.web.3.0.ve.web.4.0/520633.0/index.html>
- [13] <https://blogs.loc.gov/digitalpreservation/2011/11/the-average-lifespan-of-a-webpage/>
- [14] <http://www.worldwidewebsite.com/>

- [15] <http://netpreserve.org/about-us>
- [16] https://www.loc.gov/preservation/digital/formats/content/webarch_quality.shtml
- [15] <http://netpreserve.org/about-us> var
- [16] <http://crawler.archive.org>
- [17] <http://www.httrack.com/>
- [18] <http://www.hanzoarchives.com/>
- [19] www.isi.edu/in-notes/rfc1321.txt
- [20] <https://addons.mozilla.org/en-US/firefox/addon/archivefacebook/>
- [21] <https://github.com/internetarchive/brozzler>
- [22] <https://github.com/justinlittman/fbarc>
- [23] <https://github.com/ludios/grab-site>
- [24] <https://webarchive.jira.com/wiki/display/Heritrix/Heritrix;jsessionId=E8F4A354C3D13B2C67BEBF8F401CF5E4>
- [25] <https://github.com/steffenfritz/html2warc>
- [26] <http://www.httrack.com/>
- [27] <https://github.com/NCSU-Libraries/lentil>
- [28] <http://mementoweb.github.io/SiteStory/>
- [29] <https://github.com/docnow/twarc>
- [30] <http://warcreate.com/>
- [31] <https://machawk1.github.io/wail/>
- [32] <https://github.com/helgeho/Web2Warc>
- [33] <https://webrecorder.io/>
- [34] <http://www.gnu.org/software/wget/>
- [35] <https://github.com/chfoo/wpull>
- [36] <https://github.com/oduwsdl/warrick>
- [37] <http://cinch.nclive.org/Cinch/>
- [38] <https://sbforge.org/display/NAS/Releases+and+downloads>
- [39] <http://webcurator.sourceforge.net/>
- [40] <http://sourceforge.net/projects/deeparc/>
- [41] <https://github.com/ikreymer/pywb>
- [42] <https://github.com/iipc/openwayback/>
- [43] <http://code.google.com/p/httrack2arc/>

- [44] <https://sbforge.org/display/JWAT/JWAT>
- [45] <https://bitbucket.org/jhove2/main/wiki/Home>
- [46] <https://wiki.umiacs.umd.edu/adapt/index.php/WarcManager>
- [47] <https://github.com/ukwa/webarchive-discovery>
- [48] <http://archive-access.sourceforge.net/projects/nutchwax/>
- [49] [http://coptr.digipres.org/WERA_\(Web_ARchive_Access\)](http://coptr.digipres.org/WERA_(Web_ARchive_Access))
- [50] <http://netpreserve.org/openwayback>
- [51] <https://sourceforge.net/projects/xinq/>
- [52] <https://github.com/helgeho/ArchiveSpark>
- [53] <https://github.com/lintool/warcbase>
- [54] <https://github.com/webrecorder/har2warc>
- [55] <https://sbforge.org/display/JWAT/JWAT>
- [56] <https://archivebrowser.c3.cx/>
- [57] <http://unarchiver.c3.cx/unarchiver>
- [58] <https://github.com/chfoo/warcats>
- [59] <https://github.com/helgeho/WarcPartitioner>
- [60] <http://www.robotstxt.org/orig.html>
- [61] <http://www.robotstxt.org/norobots-rfc.txt> .
- [62] https://en.wikipedia.org/wiki/International_Standard_Link_Identifier
- [62] https://en.wikipedia.org/wiki/International_Standard_Link_Identifier var
- [63] <https://www.loc.gov/marc/bibliographic/>
- [64] https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf(<https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>)
- [65] <http://dublincore.org/documents/dces/>
- [66] <http://id.loc.gov/authorities/subjects.html>
- [67] <http://www.oclc.org/content/dam/oclc/dewey/versions/print/intro.pdf>
- [68] <https://www.loc.gov/bibframe/>
- [69] <http://www.loc.gov/standards/mets/>
- [70] <https://www.loc.gov/standards/premis/>
- [71] <https://www.iso.org/standard/57284.html>
- [72] <https://tr.wikipedia.org/wiki/PageRank>

- [73] https://www.symantec.com/security_response/glossary/define.jsp?letter=f&word=fqdn-fully-qualified-domain-name
- [74] <https://bsol.bsigroup.com/Bibliographic/BibliographicInfoData/00000000003027460>
- [75] <https://github.com/iipc/awesome-web-archiving#tools--software>
- [76] <https://www.webarchive.org.uk/ukwa/visualisation>
- [77] <https://archive-it.org/>
- [78] <http://archivethe.net/en/index.php/about/>
- [79] http://www.cdlib.org/services/uc3/old_was.html
- [80] <http://aleph-archives.com/>
- [81] <https://www.hanzo.co/>
- [82] <http://www.businesswire.com/news/home/20120305005214/en/Reed-Tech-Launches-Web-Preserver-Mobile-Web>
- [83] http://cordis.europa.eu/project/rcn/95489_en.html
- [84] <http://timetravel.mementoweb.org/>
- [85] <http://netpreserve.org/about-us>
- [81] <https://www.hanzo.co/>
- [86] <http://netpreserve.org/about-us/members/>
- [87] <http://www.collectionscanada.gc.ca/webarchives/index-e.html>
- [88] <https://www.loc.gov/websites/>
- [89] <http://www.bne.es/en/LaBNE/ArchivoWeb/index.html>
- [90] <http://www.bne.es/es/Inicio/index.html>
- [91] http://www.bnf.fr/m/fr/m_anx_articles/a.archives_internet.html
- [92] <http://www.bnf.fr/fr/acc/x.accueil.html>
- [93] <http://veebiarhiiv.digar.ee/>
- [94] <http://www.nlib.ee/en>
- [95] <http://haw.nsk.hr/>
- [96] <http://www.nsk.hr/en/>
- [97] <http://vefsafn.is/>
- [98] <https://landsbokasafn.is/index.php/english>
- [99] <http://webarchive.nationallibrary.fi/va/>
- [100] <https://www.kansalliskirjasto.fi/en>
- [101] <http://www.kb.se/om/projekt/Svenska-webbsidor---Kulturarw3>
- [102] <http://www.kb.se/>

- [103] <http://netarkivet.dk/in-english/>
- [104] <http://www.kb.dk/en/>
- [105] <http://www.nb.no/English>
- [106] <http://www.padicat.cat/>
- [107] <http://www.bnc.cat/eng>
- [108] <http://arhiv.nuk.uni-lj.si/>
- [109] <http://www.nuk.uni-lj.si/eng/>
- [110] <http://www.webarchive.org.uk/ukwa/>
- [111] <https://www.bl.uk/>
- [112] <https://www.kb.nl/>
- [113] <https://webarchiv.cz/index.php?lang=en&action=default>
- [114] <http://www.en.nkp.cz/>
- [115] <https://webarchiv.onb.ac.at/>
- [116] <https://www.onb.ac.at/en/>
- [117] <https://www.e-helvetica.nb.admin.ch/pages/main.jsf>
- [118] <https://www.nb.admin.ch/?lang=en>
- [119] <http://webarchivs.lndb.lv/>
- [120] <https://www.lnb.lv/en>
- [121] http://www.dnb.de/DE/Netzpublikationen/Webarchiv/webarchiv_node.html
- [122] http://www.dnb.de/EN/Home/home_node.html
- [123] <http://www.oasis.go.kr/>
- [124] <https://www.nl.go.kr/english/>
- [125] <http://warp.da.ndl.go.jp/search/>
- [126] <http://www.ndl.go.jp/en/>
- [127] http://web.nli.org.il/sites/NLI/English/collections/israel-collection/internet_archive/Pages/default.aspx
- [128] <http://web.nli.org.il/sites/nlis/en>
- [129] <http://www.nlc.cn/newen/>
- [130] <http://eresources.nlb.gov.sg/webarchives/>
- [131] <https://www.nlb.gov.sg/>
- [132] <http://natlib.govt.nz/collections/a-z/new-zealand-web-archive>
- [133] <https://natlib.govt.nz/>

- [134] <http://pandora.nla.gov.au/>
- [135] <http://www.nla.gov.au/>
- [136] <http://www.resmigazete.gov.tr/eskiler/2012/02/20120229-1.htm>
- [137] <http://www.resmigazete.gov.tr/eskiler/2012/08/20120818-9.htm>
- [138] <http://netpreserve.org/web-archiving/tools-and-software/>
- [139] https://tr.wikipedia.org/wiki/Active_Directory
- [140] <http://www.dictionary.com/browse/data-warehouse>
- [141] <https://en.wikipedia.org/wiki/Logfile>
- [142] <https://tr.godaddy.com/help/web-sitem-kok-dizini-nedir-16187>
- [143] <https://www.webopedia.com/TERM/S/script.html>
- [144] <http://info.ssl.com/article.aspx?id=10241>
- [145] https://en.wikipedia.org/wiki/Monitoring#In_computing
- [146] <https://en.wikipedia.org/wiki/Z3.9.50>
- [147] <https://sourceforge.net/>
- [148] <http://www.tdk.gov.tr/>
- [149] <http://www.kulturvarliklari.gov.tr/TR,44423/dunya-miras-listesi.html>
- [150] <http://www.tuik.gov.tr/UstMenu.do?metod=istendeke>
- [151] <https://github.com/oduwsdl/warrick>
- [152] <https://sourceforge.net/projects/deepar/>
- [153] <http://tureng.com/tr/turkce-ingilizce/rootkit>
- [154] <http://archive.org/about/faqs.php#Archive-It>
- [155] <http://info.cern.ch/hypertext/WWW/TheProject.html>
- [156] <http://teftis.kulturturizm.gov.tr/TR,204537/milli-kutuphane-baskanligi-elektronik-yayinlarin-derlen-.html>
- [157] <http://bilgisayarkavramlari.sadievrenseker.com/2011/04/23/n-gram/>

EKLER

EK 1: WEB ARŞİVLEME KAVRAM LİSTESİ

Terim	Türkçe Karşılığı	Kavram Analizi	Kaynaklar
Access	Erişim	Bir kütüphane veya onun koleksiyonuna giriş hakkı. Daha genel bir anlamda ise, herkesin kullanımına açık olmayan ve ücretsiz yararlanılamayan bir kaynağı kullanma hakkı veya olanağı.	(Reitz, 2002:5)
		Bir kütüphane tarafından sağlanan çevrimiçi hizmetlere başarıyla sağlanan talep	(ISO, 2013:9)
Active Directory	Aktif Dizin	İnsanların, şirketlerin, kurumların, organizasyonların tüm bilgilerini saklayan ve ihtiyaçlar ve standartlar doğrultusunda şekillendirilebilen veri tabanı veya izin hizmeti.	[139]
Access Tool	Erişim Aracı	Arşivlenen Internet kaynaklarını bulmak, onlara erişmek ve onları yeniden göstermek için kullanılan yazılım	(ISO, 2013:10)
Administrative Metadata	Yönetmel Üstveri	Bir depo içerisindeki dijital nesnelerin doğru bir şekilde yönetimini sağlamak için gerekli olan bilgi	(ISO, 2013:10)
ARC File	ARC Dosyası	Web arşivlerinde kullanılan saklama dosya formatı	(NA, 2015:14)
Authentication	Doğrulama	Bilgisayar sistemlerine veya ağlarına giriş için güvenlik amaçlı kimlik doğrulama işlemi	(Reitz, 2002:53)
Backup	Yedekleme	Veri işleme sürecinde, önemli bir veri dosyasının kaybolma, hasar görme veya yok olma ihtimaline karşı ikinci kopyasını üretme	(Reitz, 2002:61)
Bit Stream	Veri Akışı	Dijital bir dosyayı oluşturan sıfır ve bir sayı dizileri	(ISO, 2013:2)
Bulk Crawl (Bulk Harvest)	Yığın Tarama (Yığın Derleme)	Bir ya da birkaç üst seviye domainin veya alt kümelerinin tamamını toplamayı amaçlayan tarama	(ISO, 2013:2)
Cache	Önbellek	Veri bloklarının geçici olarak depolanmasını sağlayarak sistem performansını iyileştirmeyi amaçlayan bir mikro bilgisayarın içine yerleştirilen yüksek hızlı belleğin atanmış küçük bir kesiti.	(Reitz, 2002:107)
Capture	Yakalama	Depolama veya arşivleme amacıyla Web'den bir depolama havuzuna dijital bilgiyi alma işlemi	(Bragg ve Reed, 2014)
		Bir tarayıcı kullanarak Web' den Web arşivine dijital bilgiyi kopyalama işlemi	(NA, 2015:14)

Cookie	Çerez	Bir Web sunucusu tarafından üretilmiş Internet'e bağlı bir bilgisayara iletilen ve kendi Web tarayıcısı üzerindeki çerez dosyasında depolanan küçük veri dizgisi	(Reitz, 2002:171)
Crawl	Tarama	Bir tarayıcı kullanarak kaynakları tarama ve kopyalama işlemi	(ISO, 2013:3)
Crawl Settings (Crawl Parameters)	Tarama Ayarları (Tarama Değişkenleri)	Toplanması gereken kaynakların ve kaynaklar için gereksinim duyulan sıklık ve derinliğin belirlenmesi	(ISO, 2013:3)
Crawler (Spider, Robot, Scraper)	Tarayıcı	Bağlantıları takip ederek ve verileri ayıklayarak keşifte bulunmak amacıyla Web'i dolaşan yazılım. Web arşivlemede, hedef alan adları üzerindeki kaynakları takip etmek ve Web tabanlı kaynakları yakalamak üzere yapılandırılmıştır.	(NA, 2015:14)
		World Wide Web üzerinde yer alan enformasyonu akıllı bir biçimde tarayabilen robot yazılım	(Reitz, 2002:181)
Crawler Trap	Tarayıcı Tuzağı	Az değeri olan veya hiçbir anlamlı içeriği olmayan kaynakların referanslarını sonsuza dek takip etmesine veya bir tarayıcının çökmesine neden olan bir veya bir dizi Web sayfası	(ISO, 2013:3)
Curator	Küratör	Web temelli kaynaklar için özellikle koleksiyon oluşturmaktan sorumlu kimse	(Bragg ve Reed, 2014)
Curator Tool	Kürasyon Yazılımı	Depolama sürecini destekleyen ve Web tarayıcısının üzerinde çalışan uygulama	(ISO, 2013:3)
Data Processing	Veri İşleme	Makinece okunabilir formata dönüştürülmüş veri üzerinde merkezi işlem birimi tarafından bir veya bir dizi işlemin sistematik uygulanması.	(Reitz, 2002:194)
Data Mining	Veri Madenciliği	Farklı görünüm ve boyuttaki nicel veriyi analiz ederek örüntüleri ayıklama, kategorize etme ve olası ilişki ve etkileri özetlemeye yarayan yazılım işlemi	(ISO, 2013:3)
Data Warehouse	Veri Ambarı	Bir organizasyonun içerisindeki farklı birimlerden toplanan sayısal verinin büyük ve merkezi bir koleksiyonda toplanması.	[140]
Database	Veri tabanı	Tek biçim formattaki belgeleri içeren, bilgiye erişimi ve taramayı hızlandıran ve kolaylaştıran, veri tabanı yönetim sistemi yazılımı ile idare edilen, spesifik bir konu veya alanla ilişkili sayısallaştırılmış enformasyonun düzenli olarak güncellendiği büyük dosya yapısıdır.	(Reitz, 2002:193)
Deep Web (Hidden Web, Invisible Web)	Derin Web	Özellikle dinamik olarak veya şifre ile korunan kaynakları içeren, arama motorları tarafından indekslenemeyen ve taranamayan Web' in bir bölümü.	(ISO, 2013:3)

Descriptive Metadata	Tanımlayıcı Metadata	Bir dijital nesnenin entelektüel içeriğini tanımlayan bilgi	(ISO, 2013:3)
Default	Varsayılan	Kullanıcının özel talimatlarının olmadığı bir donanım ya da yazılımın içerisinde otomatik olarak seçilmiş bir değer, seçenek veya ayar.	(Reitz, 2002:196)
Deposit	Derleme	Koruma amacıyla bir arşivden alınmış bir veya daha çok dijital kaynak.	(Ruusalepp, 2015:5)
Digital	Sayısal	0 ve 1 ikili sayılarıyla ifade edilen voltaj darbelerinin süreksiz ve aralıklı aktarılmasıyla oluşan veri	(Reitz, 2002:205)
Digital Archive	Dijital Arşiv	Kullanıcı erişimini arttırma veya koruma amacıyla makinece okunabilir formata dönüştürülmüş arşivsel malzeme.	(Reitz, 2002:206)
Digital Divide	Dijital Bölünme	Bilgi kaynak ve hizmetlerine ulaşmak için bilgisayar donanımı ve yazılımını satın alma konusunda yetersiz olan veya Internet erişimi olmayan düşük gelirli ailelerin ve toplulukların erişim imkânlarındaki yetersizliği ifade eder.	(Reitz, 2002:206)
Digital Imaging	Dijital İmgeleme	Dijital fotoğrafçılık, tarama ve nokta esaslı grafikleri içeren, bilgisayar biliminde yakalamanın, depolamanın, yönlendirmenin, aktarımın ve imgeleri göstermenin tüm yönlerini kapsayan alan.	(Reitz, 2002:206)
Digital Preservation	Dijital Koruma	Dijital kaynakların uzun dönemli depolanması, bakımı ve erişimidir.	(Ruusalepp, 2015:5)
Digital Signature	Dijital İmza	Şifreleme tekniklerini kullanarak bir veri kaynağının gerçekliğini, güvenilirliğini ve doğruluğunu gösteren enformasyon.	(Ruusalepp, 2015:5)
Domain Name	Alan Adı	Internet üzerindeki belirli bir siteyi tanımlayan adres.	(Reitz, 2002:218)
		Alan Adı Sistemi (DNS) tarafından kuralları ve yöntemi tayin edilen ya da özerklik ve yetki alanları belirlenmiş yer tanıma dizgesi	(ISO, 2013:3)
Domain Name System (DNS)	Alan Adı Sistemi	Internet'e bağlanmaları amacıyla varlıkları tanımlamak için kullanılan hiyerarşik ve dağınık küresel adlandırma sistemi.	(ISO, 2013:3)
Download	Karşıdan Yükleme, İndirme	Bir veya daha fazla dosyayı ana bilgisayardan uç birime, ağ sunucusundan istemciye veya sabit disk sürücüsünden başka bir depolama ortamına kopyalama işlemidir.	(Reitz, 2002:222)
Emulation	Öykünüm	Mevcut bilgisayar sistemi üzerinde yazılım kullanarak eski bir sistemin davranışını ve işlevselliğini yeniden üretme	(ISO, 2013:4)

Encryption	Şifreleme	Veriyi yetkili alıcı haricindeki herkese okunamaz kılmak için kamu iletişim kanallarından aktarılmadan önce gizli bir kodu içeren bir mesajın içerisinde dönüştürme işlemi	(Reitz, 2002:236)
End-user	Son Kullanıcı	Kütüphane kaynaklarının veya diğer bilgi hizmetlerinin sunulduğu kişiler.	(Reitz, 2002:238)
Exceptional Crawl	İstisnai Tarama	Web arşivleme yapan kuruluşun takdirine bağlı ve düzenli tarama programına ek olarak istisnai durumlarda yapılan tarama.	(NA, 2015:14)
Ethernet	Ethernet	Dünya genelinde en yaygın biçimde kullanılan yerel ağ teknolojisi.	(Reitz, 2002:246)
Extranet	Genişletilmiş Ağ	Kamuya açık olmayan, organizasyonun dışındaki seçilmiş insanlara ve kuruluş çalışanlarına çeşitli seviyelerde hizmet sağlaması için tasarlanmış özel bilgisayar ağı.	(Reitz, 2002:253)
Firewall	Güvenlik Duvarı	Sınır güvenliği fonksiyonu olan genellikle daha büyük bir ağdan yerel ağa verinin transfer edildiği ve ağı bir bölümünden diğerine geçişte oluşan trafiği bloklayan özel bir bilgisayar veya yazılım.	(Reitz, 2002:269)
File Transfer Protocol (FTP)	Dosya Transfer Protokolü	Veri dosyalarının e-mail de olduğu gibi Internet üzerindeki bir bilgisayardan diğerine kopyalanmasına izin veren platform bağımsız bir TCP/IP protokolüdür	(Reitz, 2002:286)
Full-text Search	Tam Metin Arama	Bibliyografik bir veri tabanında arama ifadelerinde yer alan metnin tüm kayıt veya metin içerisinde aranabildiği erişim sağlama yöntemi.	(Reitz, 2002:288)
Harvest	Derleme, Toplama	Taramanın bir parçası olarak Web içeriğini yakalama eyleminin diğer adı.	(Bragg ve Reed, 2014)
Heritrix	Heritrix	Internet Archive'ın açık kaynaklı, genişletilebilir, Web ölçekli arşivleme amaçlı Web tarayıcı projesidir.	(Bragg ve Reed, 2014)
Host	Ana Bilgisayar, Sunucu	Diğer terminaller veya bilgisayarlar için veri kaynağı olarak hizmet veren bir bilgisayardır.	(Reitz, 2002:320)
HTML (Hypertext Markup Language)	Hipermetin İşaretleme Dili	Ham metine yapısal ve anlamsal bilgiyi eklemeye kullanılan bileşenleri içeren, Web sayfaları için temel işaretleme dilidir.	(ISO, 2013:4)
HTTP (Hypertext Transfer Protocol)	Hipermetin Aktarım Protokolü	Web üzerinde bilgi transferini sağlayan istemci - sunucu iletişim protokolüdür.	(ISO, 2013:4)
Hyperlink (Link)	Bağlantı	Internet üzerindeki Web sayfalarını birbirine bağlamak için kullanılan bağlantı yapısı	(ISO, 2013:4)
Hypertext	Hipermetin	Dijital enformasyonu art arda sıralamak yerine veri bileşenlerinin ve ilişkili dosyaların birleştirilmesine izin veren sunum metodu.	(Reitz, 2002:324)

Information Retrieval	Bilgi Erişim	Bir veri dosyasından kayıtlı bir enformasyonu seçerek geri çağırma yöntem, metot ve süreçleri.	(Reitz, 2002:336)
Junk (spam)	Önemsiz İçerik, İstenmeyen İçerik	Uzun süre koruma değerine sahip olmayan ve ilgisiz görülen istenmeyen içerik	(ISO, 2013:4)
Knowledge Base	Bilgi Tabanı	Bir kişi veya sistem tarafından birleştirilmiş, kişiye veya sisteme alınan bilgi algılama becerisi kazandıran bir takım enformasyon.	International Organization for Standardization [ISO], 2012b:12)
Link	Bağlantı	Bir hipermetin dokümanı içerisinde yer alan, metin içerisindeki bir kelime, sözcük grubu ya da bir Mouse veya işaret aygıtının kullanılmasıyla harekete geçen, sembol, ikon veya diğer grafik bileşen olarak görünen başka bir dokümanın Internet adresine (URL) geçişi sağlayan bağlantı.	(Reitz, 2002:393)
Link Mining	Bağlantı Madenciliği	Örüntü ayıklama ve bağlantılardan elde edilen bulgulara odaklanan işlem ve analizler	(ISO, 2013:4)
Live Web Leakage	Canlı Web Kaçağı	Web arşivindeki arşivlenen sürüm yerine canlı Web sitesindeki güncel kaynağın çözülmesinden kaynaklanan arşiv kaynaklarını dönüştürme problemi.	(ISO, 2013:4)
Local Area Network (LAN)	Yerel Ağ	Yüksek hızlı bir sunucu ve tek kullanıcı geliştirmiş mikro bilgisayarların, bir ağ işletim sisteminin ve iletişim bağlantısının bulunduğu kolej, üniversite, kuruluş kampüsü gibi tek bir binanın veya bir grup bitişik binanın bulunduğu nispeten küçük bir coğrafi alandaki iletişim ağıdır.	(Reitz, 2002:398)
Log File	Günlük Dosyası	Bilgisayarlarda hem işletim sisteminde veya çalıştırılan yazılımlarda gerçekleşen olayları hem de iletişim yazılımlarında farklı kullanıcılar arasındaki mesajları kaydeden dosya türü.	[141]
		Sunucu tarafından otomatik olarak üretilen ve sunucu aktivitelerinin kaydını tutan dosya	(ISO, 2013:5)
Mainframe	Ana Bilgisayar	Bağımsız işlem yeteneği olmayan ve hatırı sayılır ölçüde bilgisayar gücüne ihtiyaç duyan karmaşık uygulamaları çalıştıran, birçok terminali (uç birimi) destekleme kapasitesine sahip büyük bilgisayar sistemi.	(Reitz, 2002:407)
Merge	Birleştirmek	Mantıksal sıralama içerisindeki iki veya daha fazla ayrı enformasyon dosyasını otomatik ya da manuel işlemlerle tek bir dosyada birleştirmek.	(Reitz, 2002:421)
Metadata	Üstveri	Kelime karşılığı veri hakkında veridir. Fonksiyonel anlamıyla bilgi kaynakları ve varlıklarını çeşitli amaçlar için tanımlayan yapılandırılmış veridir.	(Reitz, 2002:421)

		Dijital nesnenin bağlam, içerik ve yapısı ile yaşam süresi boyunca yönetimini gösteren veri.	(ISO, 2013:5)
Microsite	Mikro Site	Ana Web adresi altında yer alan kök dizine bağlı durumdaki alt dizin.	(NA, 2015:14)
Migration	Göç	Bir donanımın/ yazılımın veri kaynağının bir diğer nesile dönüştürülmesi veya biçimlendirmesi.	(Ruusalepp, 2015:8)
		Var olan özelliklerini barındıracak şekilde bir donanımı veya yazılımı biçimini değiştirmeden bir diğerine taşıma.	(International Organization for Standardization [ISO], 2014:1)
Mirror Website	İkiz Web Sitesi	Genellikle dünyanın başka bir bölgesinden veya ülkesinden gelen talebi karşılamak için, içeriği bir ana makinede muhafaza etmek yerine bir sunucu üzerinde tutulan bir Web sitesinin özdeş kopyası	(Reitz, 2002:428)
Multiple Access	Çoklu Erişim	Erişimin yalnızca bir noktadan sağlanmasının aksine dosya verisinin birden fazla erişim noktasında sunumu. Örnek vermek gerekirse bir bibliyografik veri tabanı ya da kütüphane kataloğu yazar, başlık, konu, anahtar kelime vb. alanları kullanılarak erişim sunabilir.	(Reitz, 2002:435)
Nomination	Aday	Web arşivine eklenmesi düşünülen kaynak	(ISO, 2013:6)
Offline	Çevrimdışı	Ağ ile bağlantısı olmayan otomatikleştirilmiş servis	(Reitz, 2002:464)
Online	Çevrimiçi	Internet, Intranet veya başka bir ağa telekomünikasyon bağlantılarıyla bağlı olma, bağımsız olmama.	(Reitz, 2002:467)
Open Source	Açık Kaynak	Genellikle Internet üzerinde yer alan, bir sahibin veya lisans sahibinin zorlaması olmaksızın kaynak kodları kullanıma uygun halde bulunan bilgisayar programları.	(Reitz, 2002:470)
Partial Crawl	Kısmi Tarama	Web arşivleme sürecinde tarayıcının hedef alan adından ayrılarak sınırlı derinlikteki başka bir Web sitesini arşivlemek üzere taraması	(NA, 2015:14)
Permission	İzin	Canlı bir Web sitesini tarama ve/veya içeriğini bir Web arşivi üzerinde gösterebilme yetkisi	(ISO, 2013:5)
Plug-in	Eklenti	Temel yazılım paketinin kapasitesini arttırıcı, kurulumu kolay bir destek program	(Reitz, 2002:510)
Portal	Portal	Aslen gazeteler, hava durumu, rehber bilgisi, Web tarama, ücretsiz e-mail hesabı, chat grupları, çevrimiçi alışveriş ve diğer siteler ile bağlantı gibi geniş çapta hizmetler ve kaynaklar sunan genel amaçlı bir Web sitesidir.	(Reitz, 2002:516)
Provenance (Context) Metadata	Kaynak Üstveri	Bir kaynağın yaşam döngüsüyle ilişkili oluşum ve süreçleri tanımlama bilgisi.	(ISO, 2013:10)

		(Örneğin: Yapılandırma veya günlük dosyaları)	
Proxy Server	Vekil Sunucu	Bir bilgisayar ağında istemci ve sunucu arasında çalışan, genellikle koruma sağlamak amacıyla güvenlik duvarı olarak kurulan veya kullanıcı isteklerini doğrulama ya da yetkilendirme kontrolü gibi sunucunun kendisinin yürüttüğü ön işlem görevlerini uygulamada erişim hızını arttıran uygulama programıdır.	(Reitz, 2002:538)
Registered User	Kayıtlı Kullanıcı	Kütüphane içinden veya dışından koleksiyon ve/veya hizmetlere belirlenen yetkilendirme dâhilinde erişim sağlayan kayıtlı kişi veya kurum.	(ISO, 2013:5)
Quality Assurance (QA)	Kalite Güvencesi	Otomatik ve manuel araç ve tekniklerin bir kombinasyonunu kullanarak yakalama işleminin tamlığını kontrol süreci	(NA, 2015:15)
Refreshment	Yenileme	Depolama donanımı ve yazılımının fonksiyonelliğini değiştirmeden, depolama ortamının eşiti ile değiştirilmesinin sağlandığı veri aktarım uygulaması	(ISO, 2014:1)
Remote Harvesting	Uzaktan Derleme	Web genelinden Web içeriğini derleme eylemi.	(NA, 2015:15)
Request	İstek, İstem	URL tarafından kimliklendirilmiş belirli bir kaynak için İstemci sistem (browser veya crawler) tarafından uzak sunucuya gönderilen HTTP formundaki mesaj.	(ISO, 2013:6)
Response	Yanıt	Bir kaynak için bir HTTP isteğine, istenen kaynağı veya başka bir URL' ye yönlendirmeyi veya nedenleriyle birlikte olumsuz (hata) yanıtı içeren uzak bir sunucu tarafından verilen cevap.	(ISO, 2013:6)
Response Code (Status Code)	Yanıt Kodu, Durum Kodu	İstemci sunucuya istenen kaynağın durumunu belirten üç basamaklı numara	(ISO, 2013:6)
Rights Metadata	Haklar Üstverisi	Bir nesnenin mülkiyetini ve yasal olarak kullanım iznini tanımlayan bilgi.	(ISO, 2013:10)
Robot.txt (Robots Exclusion Standart)	Robot Engelleme Standardı	Bir Web sitesinin tamamına veya bir bölümüne Web tarayıcılarının erişimini engellemek için kullanılan protokol	(ISO, 2013:6)
Root	Kök Dizin	Web arşivlemede tarayıcı alan adına erişim sağladığında yüklediği içeriktir.	[142]
Root URL	Kök URL	Bir Web sitesi üzerindeki genellikle URL'nin en kısa şeklidir (Örneğin: http://www.gov.uk/).	(NA, 2015:15)
Router	Yönlendirici	Elektronik mesajlardan oluşan küçük dijital veri paketlerini en verimli yolu tercih ederek, ağ düğümü olmuş bir bilgisayar	(Reitz, 2002:587)

		ağından diğerine yönlendirmek için tasarlanan donanım aygıtı.	
Scheduled Crawl	Zamanlanmış Tarama	Web sitelerinin önceden belirlenen tarih aralıklarında taranması işlemi. Bu aralıklar 1-12 ay arasında düşünülebilir.	(NA, 2015:15)
Script	Betik	Herhangi bir programlama dilinde yazılmış, kullanıcı etkileşimi olmadan çalışabilen kodlardan oluşan komut listesidir.	[143]
Second Level Domain	İkinci Düzey Alan Adı	İlgi alanları veya kurumların özel kategorileri için üst seviye alan adlarının içindeki alt bölümler	(ISO, 2013:6)
Secure Sockets Layer (SSL)	Güvenli Yuva Katmanı	Bir Web sitesi ile tarayıcı arasında şifreli bir bağlantı kurmayı sağlayan standart güvenlik teknolojisi.	[144]
Seed (Targeted URL)	Kaynak, Kaynak Liste	Tarayıcının yakalama talimatı verdiği URL.	(NA, 2015:15)
		Bir Web Tarayıcı tarafından başlangıç noktası olarak kabul edilen, taranacak belirli bir kaynağın konumuna karşılık gelen URL	(ISO, 2013:6)
Selective Crawl (Selective Harvest)	Seçimli Tarama	Belirli ölçütlere göre seçilen kaynakları derleme amacıyla yapılan tarama	(ISO, 2013:6)
Server	Sunucu	Aynı ağa bağlı istemci bilgisayardan alınan, veri veya program dosyaları indirme isteklerini yanıtlamak için programlanmış ağ üzerindeki ana bilgisayar.	(Reitz, 2002:612)
Storage Medium	Depolama Ortamı	Dijital bilginin üzerine kaydedilebileceği aygıt.	(ISO, 2014:1)
Structural Metadata	Yapısal Üstveri	Mantıksal birimleri birleştirmek için nesnelerin nasıl bir araya getirileceğini tanımlayan bilgi	(ISO, 2013:6)
Sub-domain	Alt Domain	Kök Web adresinden önce isimlendirilen domainidir. Örnek verecek olursak: http://mysite.domain.gov.uk/	(NA, 2015:15)
Supplementary URL List	Tamamlayıcı URL Listesi	Bir Web sayfası üzerindeki özellikle tarayıcı teknolojisinin erişimde zorlandığı ya da harici ana bilgisayarla yer alan URL'leri listeleyen metin dosyası. XML site haritasına bir alternatiftir.	(NA, 2015:15)
Syntax	Sözdizim	Aramayı yürüten bilgisayar tabanlı bilgi erişim sisteminde dizilimi belirleyen, kategorilere ayrılmış anahtar sözcük sistemindeki komut arama terimleri ve Boolean operatörleri sıralı listeleridir.	(Reitz, 2002:660)
Target	Hedef	Bir veya daha fazla çekirdek kaynak ve tarama ayarları belirlendikten sonra derlenecek anlamlı kaynak dizisi.	(ISO, 2013:7)

Technical Metadata	Teknik Üstveri	Dijital bir nesnenin teknik özelliklerini tanımlayan bilgi.	(ISO, 2013:2)
Uniform Resource Identifier (URI)	Tek Biçimli Kaynak Tanımlayıcı	İnternet üzerindeki bir kaynağı adlandıran veya tanımlayan genişletilebilir karakter dizisidir.	(ISO, 2013:7)
Uniform Resource Locator (URL)	Tek Biçimli Kaynak Konumlayıcı	Yer belirleme amacıyla İnternet üzerindeki bir kaynağı tanımlayan benzersiz adres.	(Reitz, 2002:695)
		Bir kaynağın yerini ve ona erişim için uygun protokolü belirten URI' ın alt kümesidir.	(ISO, 2013:7)
		Kaynağın Web üzerindeki tekil lokasyonudur.	(Bragg ve Reed, 2014)
Update	Güncelleme	Sıklıkla var olan bir içeriği revize ederek veya yerine yeni bir malzeme koyarak bir veri dosyasını veya bilgi kaynağını güncel duruma getirme.	(Reitz, 2002:699)
Upload	Yüklemek	Genellikle yerel bir bilgisayardan uzak bir bilgisayarın diskine bir veya daha çok dosyanın kopyasını iletmek.	(Reitz, 2002:699)
Validation	Geçerlilik	Bir dosyanın veya dijital kaynağın içerik kontrolünü gerçekleştiren ve aynı zamanda depolanmasını ve korunmasını sağlayan kalite teminatı.	(Ruusalepp, 2015:11)
Virtual	Sanal	Yalnızca sayısal formda (siber uzayda) ortaya çıkan fiili fiziksel gerçekliği olmayan bir e-mail kutusu veya elektronik alışveriş kartı gibi aktiviteler, nesneler, varlıklar ve yerlere karşılık gelen bir sıfat.	(Reitz, 2002:711)
Virtual Reality	Sanal Gerçeklik	Özellikle bilgisayar kullanıcıları için üretilmiş, fiziksel maddelerden yoksun fakat üç boyutlu gerçekliğin görsel görünümüne uyarlanmış yazılımın kullanıldığı elektronik ortam.	(Reitz, 2002:711)
Virus	Virüs	Kullanıcının bilgisi ve talimatı olmadan çalışan yüklenmiş kod veya bilgisayar programı.	(Ruusalepp, 2015:12)
Vocabulary Mapping	Kelime Haritalama	Bazı bibliyografik veri tabanlarının tarama yazılımı içerisinde inşa edilen, uygun konu başlığını veya anahtar sözcüğü kullanıcıya özel tarama terimiyle ilişkilendiren, kaç tane başlığın listelendiğini ve o başlığa ait hiyerarşik ağaç yapısını bir kapsam notuyla birlikte veren, taramadaki daha geniş terim, dar terim ve alt başlıkların seçimine yardım eden fonksiyon. Veri tabanı için kontrollü terimler sözlüğü diğer bir deyişle thesaurus oluşturma fonksiyonudur.	(Reitz, 2002:713)
WARC Format	Web Arşiv Dosyası Formatı	Çoklu dijital dosyaları özel bir metot yardımıyla birleştiren, arşiv dosyasını ilişkili olduğu bilgilerle bir araya getiren dosya formatı.	(ISO, 2013:7)

Wayback Machine	Wayback Makinesi	Internet Archive'ın genel arşivi olarak bilinir. Wayback Machine, arşivlenen Web sitelerini canlı Web' de yer alır gibi taranabilir hale getiren bir uygulama yazılımıdır.	(Bragg ve Reed, 2014)
Web	Internet Ağı	Üç ana standart (URI, HTTP ve HTML) tarafından geçerli kılınan Internet'in temel yayın uygulaması	(ISO, 2013:7)
Web Archive	Web Arşivi	Arşiv içerisinde saklanan Web kaynaklarının koleksiyonudur.	(NA, 2015:15)
		Zamanla Web' den taranarak depolanan kaynakların tamamıdır.	(ISO, 2013:2)
Web Browser	Web Tarayıcısı	Metin formatının yanı sıra Internet üzerindeki mevcut dokümanları ve diğer veri dosyalarını grafik olarak da gösterebilen, Web sayfalarındaki kodları yorumlayabilen istemci yazılımı.	(Reitz, 2002:718)
Web Page	Web Sayfası	Bir URL tarafından kimliklendirilen, diğer kaynaklarla ilişkisi bulunan veya bulunmayan, insanlar tarafından okunabilmesi amacıyla yapılandırılmış kaynak	(ISO, 2013:5)
Web-published Materials	Web Materyali	Web aracılığıyla sunulan ve erişilen, yok olma riski altındaki hem dinamik ve hem geçici materyallerdir. Web materyalleri kültürel miras spektrumunu ve bir dizi materyal türünü kapsar.	(Bragg ve Reed, 2014)
Web Search Engine	Web Arama Motoru	Hiyerarşik konu dizini veya anahtar sözcük kullanarak, Dünya Çağında Ağ (World Wide Web) üzerinde kullanımda olan Web sitelerinde yer alan enformasyonun yerini bulma konusunda kullanıcıya yardım etmesi için tasarlanmış bilgisayar yazılımı.	(Reitz, 2002:600)
Website	Internet Sitesi	Yasal ve/veya editoryal olarak birbirine bağlanmış Web sayfaları kümesi	(ISO, 2013:7)
Weblog	Web Günlüğü	Ters kronolojiyle düzenlenmiş genellikle güncel konularla ilişkili yorum ve bağlantıların listesini sunan Web sayfası.	(Reitz, 2002:719)
Webmaster	Web Sorumlusu	Bir Web sitesini yöneten ve sürekliliğini sağlayan genellikle onu tasarlayan, adı sıklıkla anasayfa altında veya karşılama ekranında bir iletişim linkiyle verilmiş kişi.	(Reitz, 2002:719)
Website Monitoring	Web Sayfası İzleme	Bir Web sitesiyle ve uygulamasıyla etkileşime giren son kullanıcıları doğrulama veya süreci test etme işlemi.	[145]
Wireless	Kablosuz	Tel veya kablo yerine elektromanyetik yayın araçları yardımı ile Internet bağlantısı sağlama metodu.	(Reitz, 2002:723)
World Wide Web (WWW)	Dünya Çapında Ağ	İçeriği birleştirmeye izin veren, yerel veya uzaktan kumanda edilebilen, HTML olarak adlandırılan betik dilinde yazılmış dokümanlara erişim sağlayan Internet sunucularının küresel ağıdır. Web, Tim	(Reitz, 2002:727)

		Berners-Lee tarafından 1989 yılında Geneva’da CERN yüksek enerji fiziği laboratuvarında tasarlanmıştır.	
Worm	Solucan	Kendisiyle bağlantısız kaynak ve alanları tüketen, sabit disk ve hafıza arasında kendi kopyasını türeterek bir sistemi iş yapamaz duruma getiren bir tür bilgisayar virüsü.	(Reitz, 2002:728)
Extensible Markup Language (XML)	Genişletilebilir İşaretleme Dili	HTML (HyperText Markup Language) ile benzer şekilde kullanılan fakat onun yerine verinin nasıl görüntülediğiyle ilgilenen etiketlerin yer aldığı SGML (Standard Generalized Markup Language) dilinin bir alt kümesidir.	(Reitz, 2002:251)
XML Sitemap	Genişletilebilir İşaretleme Dili Site Haritası	Dizinlenen içerikte tarayıcıları ve özellikle arama motorlarını destekleyen XML’de kodlanmış URL listesi	(NA, 2015:15)
Z39.50	Z39.50	TCP/IP bilgisayar ağı üzerindeki bir veri tabanından tarama yapmak ve bilgi erişim sağlamak için oluşturulmuş bir uluslararası istemci-sunucu uygulama katmanı iletişim protokolüdür.	[146]

EK 2: WEB ARŞİVLEME İÇİN KALİTE GÖSTERGELERİ (ISO, 2013:39-47)

Kalite Göstergesi Türü	Yönetim
Gösterge Numarası	1
Adı	Toplanan URL başına maliyet
Amacı	Web arşivleme sürecinin verimliliğini değerlendirmek.
Ön Koşullar	• Web arşivlemenin toplam maliyeti • Taranan URL'lerin toplam sayısı
Yöntem	Toplanan URL başına maliyet: A/B A , belirli bir zaman periyodunda Web arşivlemenin toplam maliyetidir. B , aynı zaman periyodunda toplanan URL'lerin toplam sayısıdır. Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Toplanan URL başına maliyet düşükse bu Web arşivleme sürecinin verimliliğinin yüksek olduğunu gösterir. Toplanan URL başına maliyet yüksek ise bu ileri seviye küratörlük yapıldığını gösterebilir. Bu gösterge benzer boyut ve amaçtaki koleksiyonları karşılaştırmak için kullanılır.
Kalite Göstergesi Türü	Yönetim
Gösterge Numarası	2
Adı	Web arşivinde çalıştırılan personel yüzdesi
Amacı	Web arşivlemeye kurumsal bağlılığı göstermek
Ön Koşullar	• Web arşivlemede çalıştırılan tam zamanlı personel sayısı • Kütüphanede çalışan toplam tam zamanlı personel sayısı
Yöntem	Web arşivinde çalıştırılan personel yüzdesi A/B * 100 A , Web arşivlemenin seçim, depolama, erişim sağlama ve koruma aşamalarında çalışan tam zamanlı personel sayısını ifade etmektedir. B , sürekli, geçici ve proje bağlantılı çalışan toplam tam zamanlı kütüphane personeli sayısını ifade eder. En yakın tam sayıya yuvarlayın. Yarı zamanlı çalışanların çalışma süresini hesaplarken kendi beyanlarını veya zaman loglarını kullanın.
Yorumlar	-
Kalite Göstergesi Türü	Toplama İşleminin Kalitesi
Gösterge Numarası	3
Adı	Belirli bir zaman diliminde canlı Web'de kaybolan kaynakların yüzdesi
Amacı	Web arşivinin değerini tespit etmek
Ön Koşullar	• Arşivdeki hedeflerin sayısı • Arşivdeki canlı Web'de kaybolan hedeflerin sayısı İkinci sayı bulunamayan DNS ve 404 yanıtlarıyla otomatik olarak toplanır veya canlı Web'den manuel olarak kontrol edilerek bulunabilir.
Yöntem	Belirli bir zaman diliminde canlı Web'de kaybolan kaynakların yüzdesi A/B * 100 A , yok olan hedeflerin sayısı B , arşivdeki hedeflerin sayısı Sonucu tek ondalıklı sayıya yuvarlayın.

Yorumlar	Kaybolan bir hedefi belirlemek zordur. Bazen bir hedef alan adı yok olmaz, yeni ve farklı bir alan adına dönüşür. Bazı durumlarda ise hedefin yalnızca bir bölümü yok olur. Bu senaryo uygulamanın dışında tutulur. Bu göstergede bir DNS yanıtı yoksa veya hedefin Internet adresi 404 bulunamadı yanıtı veriyorsa hedef yok olmuştur. Eğer mümkünse canlı Web'i manuel olarak kontrol en uygun yöntemdir.
Kalite Göstergesi Türü	Toplama İşleminin Kalitesi
Gösterge Numarası	4
Adı	Zorunlu tutulan kapsamın arşivlenme yüzdesi
Amacı	Web arşivleme sonuçlarının direktiflere uygun olup olmadığını değerlendirmek
Ön Koşullar	- Her bir yıl için kuruluş tarafından depolanan hedeflerin sayısı - Zorunlu kapsamda tanımlanan veya elde edilen hedeflerin sayısı
Yöntem	Zorunlu tutulan kapsamın arşivlenme yüzdesi $A/B * 100$ A , Kütüphanenin yıllık depoladığı hedef sayısı B , Web arşivleme için zorunlu tutulan (örneğin: derleme yasası ile), tanımlanan veya elde edilen kapsam dâhilindeki hedeflerin sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Web arşivleme için direktifler sıklıkla ulusal veya kurumsaldır. Ulusal derleme yasasının bulunduğu yerlerde genellikle bu bir ulusal alan adının bölgesel sınırlarını tanımlar. Ulusal üst seviye alan adları kullanılarak sunulan Web kaynakları için bu saptama basittir. Ayrıca belirli bir faaliyet alanı kapsamındaki kaynaklar her zaman açık bir coğrafi referans altında sunulmayabilir. Örneğin .fr tescil dairesi ANFIC'e göre Fransız Web sitelerinin yalnızca üçte biri .fr üst seviye alan adı altındadır. Bölgesel sınırları ulusal üst seviye alan adları haricinde başka yollarla belirleme gereksinimi vardır. Bu gereksinim alan adı sağlayıcılardan bilgi edinmeyi gerektirebilir.
Kalite Göstergesi Türü	Toplama İşleminin Kalitesi
Gösterge Numarası	5
Adı	İstekler neticesinde hak sahiplerince sağlanan izinler ve yapılan anlaşmaların yüzdesi
Amacı	İzin taleplerinin etkinliğini değerlendirmek
Ön Koşullar	- Hak sahiplerine gönderilen anlaşma ve izin isteklerinin sayısı - Hak sahiplerince sağlanan izinler ve yapılan anlaşmaların sayısı
Yöntem	İstekler neticesinde hak sahiplerince sağlanan izinler ve yapılan anlaşmaların yüzdesi $A/B * 100$ A , Hak sahiplerince sağlanan izin ve yapılan anlaşmaların sayısı B , Hak sahiplerine gönderilen izin ve anlaşma taleplerinin sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Yüksek bir oran izin talebi aktivitesinin başarısını gösterir. Aynı zamanda açık retlerin ve olmayan yanıtların sayısını kaydetmek önerilir. Web sitesi üreticilerine yönelik iyi bir iletişim ve savunma başarıyı arttırmak için bir çözüm olabilir. Üreticileri ikna etmek için kilit

	parametreleri belirleyen ayrıntılı bir plan hazırlanabilir ve bilgi dağıtımı için en iyi kanalların belirlenmesi üzerine karar verilebilir.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	6
Adı	Son kullanıcıların erişebildiği kaynak yüzdesi
Amacı	Web arşivinin kullanılabilirliğini değerlendirmek
Ön Koşullar	• Web arşivindeki toplam kaynak sayısı • Çevrimiçi kullanılabilir kaynak sayısı (çevrimiçi arşiv) • Yerinde kullanılabilir kaynak sayısı (gri arşiv) Yukarıdakiler URL veya byte cinsinden ölçülebilir.
Yöntem	Son kullanıcıların erişebildiği kaynak yüzdesi $(A + A')/B * 100$ A , Web arşivindeki çevrimiçi kullanılabilir kaynak sayısı A' , Web arşivindeki yerinde kullanılabilir kaynak sayısı B , Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Yüksek bir oran Web arşivinin yüksek görünürlüğünü ve erişilebilirliğini gösterir. Son kullanıcıların doğrudan kullanılabilirliğini değerlendirmek için, gösterge sadece çevrimiçi ortamda mevcut olan kaynakları hesaplamada da kullanılabilir. Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	7
Adı	Tam metin indekslenen kaynak yüzdesi
Amacı	Web arşivinin aranabilirliğini değerlendirmek
Ön Koşullar	• Web arşivindeki toplam kaynak sayısı • Tam metin indekslenen kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	Tam metin indekslenen kaynak yüzdesi $A/B * 100$ A , Tam metin indekslenen kaynak sayısı B , Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Tam metin arama bir Web arşivinin aranabilirliğini ve kullanılabilirliğini büyük ölçüde artırır. Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	8
Adı	Kataloglanan kaynak yüzdesi
Amacı	Web arşivinin kürasyon seviyesini ve aranabilirliğini değerlendirmek
Ön Koşullar	• Web arşivindeki hedef sayısı • Katalog kaydı olan hedef sayısı
Yöntem	Kataloglanan kaynak yüzdesi $A/B * 100$ A , Katalog kaydı olan hedef sayısı B , Web arşivindeki toplam hedef sayısı

	Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Web arşivindeki kaynakları kataloglamak kaynakların erişilebilirliğini ve kullanılabilirliğini artırır. Bu ayrıca kütüphanenin sahip olduğu diğer kaynaklarla Web arşivi kaynaklarının entegre edilmesine yardımcı olur. Kataloglama masraflıdır ve yığın derleme aracılığıyla toplama yapılan büyük ölçekli Web arşivleri için mümkün olmayabilir. Bu göstergeyi hesaplarken kaynakları toplamada kullanılan derleme stratejisinin raporlanması tavsiye edilir.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	9
Adı	Yıllık erişilen kaynak yüzdesi
Amacı	Web arşivinin gerçek kullanım genişliğini değerlendirmek
Ön Koşullar	• Web arşivindeki toplam alan adı sayısı • Bir yıl içerisinde en az bir kez görüntülenen alan adlarının sayısı
Yöntem	Yıllık erişilen kaynak yüzdesi $A/B * 100$ A , Bir yıl içerisinde en az bir kez görüntülenen alan adlarının sayısı B , Web arşivindeki toplam alan adı sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Yüksek bir oran Web arşivinin geniş kapsamlı kullanıma sahip olduğunu gösterir. URL'ler yerine alan adlarının kullanılma nedeni sadece sınırlı sayıdaki alan adının altındaki kaynakların aktif olarak kullanımın mümkün olmasıdır. Alan adı başına hesaplama kullanım genişliği ve arşivin kapsamlılığını gösterir. Kamunun yanı sıra araştırmacılara yönelik iyi iletişim, yüzdeyi arttırmanın bir diğer yoludur. Gösterge 5'te belirtildiği gibi, değerli kaynaklar için daha ikna edici örnekleri ve bilgi dağıtımı için en iyi yolları belirlemek bir iletişim planıyla ayrıntılı bir biçimde sağlanabilir.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	10
Adı	Web arşivine yapılan bir ziyareti de kapsayan kütüphaneye ziyaretlerinin yüzdesi
Amacı	Kütüphane ziyaretçileri (yerinde veya çevrimiçi) tarafından Web arşivinin kullanımını değerlendirmek
Ön Koşullar	• Kütüphane ziyaretlerinin toplam sayısı (fiziksel ve sanal ziyaretler) • Web arşivine yapılan ziyaret sayısı
Yöntem	Web arşivini kullanan kütüphane ziyaretçileri yüzdesi $A/B * 100$ A , Web arşivine yapılan ziyaret sayısı (örneğin, aylık- yıllık) B , Kütüphane ziyaretlerinin toplam sayısı (örneğin, aylık- yıllık) Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Yüksek bir oran Web arşivinin ve diğer kütüphane hizmetlerinin sık kullanıldığını gösterir. Fiziksel ziyaretler kütüphane binalarına yapılan girişlerdir. Sanal ziyaretler kütüphane Web sitesine yapılan girişlerdir.
Kalite Göstergesi Türü	Erişilebilirlik ve Kullanım
Gösterge Numarası	11

Adı	Ziyaret başına sayfa görüntüleme yüzdesi
Amacı	Kullanıcıların arşive olan ilgisini değerlendirmek
Ön Koşullar	- Bir raporlama sürecindeki toplam sayfa görüntüleme sayısı - Aynı raporlama sürecindeki Web arşivi ziyaretlerinin toplam sayısı
Yöntem	Ziyaret başına sayfa görüntüleme yüzdesi $A/B * 100$ A, Raporlama sürecindeki toplam sayfa görüntüleme sayısı B, Raporlama sürecindeki ziyaretlerinin toplam sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Yüksek bir oran Web arşivinin yoğun bir biçimde kullanıldığını gösterir. Bu aynı zamanda arşiv kullanıcılarının ilgilendiği içeriğin koleksiyonda bulunduğunu gösterir. Kullanıcıların ve araştırmacıların ihtiyaç ve beklentilerini dikkate almak, örneğin anket uygulayarak, bu oranı yükseltmenin bir yoludur.
Kalite Göstergesi Türü	Koruma
Gösterge Numarası	12
Adı	En az bir adet kopyası olan kaynak sayısı
Amacı	Veri akışını koruma yeteneğini değerlendirmek
Ön Koşullar	- Web arşivindeki toplam kaynak sayısı - En az bir kopyası olan kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	En az bir adet kopyası olan kaynak sayısı $A/B * 100$ A, En az bir kopyası olan kaynak sayısı B, Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.
Kalite Göstergesi Türü	Koruma
Gösterge Numarası	13
Adı	Kayıbolan veya bozulan kaynak yüzdesi
Amacı	Web arşivinin depolama kapasitesini değerlendirmek
Ön Koşullar	- Web arşivindeki toplam kaynak sayısı - Kayıp veya bozuk kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	Kayıp veya bozulan kaynak yüzdesi $A/B * 100$ A, Kayıp veya bozuk kaynak sayısı B, Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Düşük bir yüzde yüksek güvenliği gösterir. Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.
Kalite Göstergesi Türü	Koruma
Gösterge Numarası	14
Adı	Dosya formatı tanımlanmış olan kaynak yüzdesi
Amacı	Web arşivinin organizasyonel bilgisini ve koruma kapasitesini değerlendirmek

Ön Koşullar	• Web arşivindeki toplam kaynak sayısı • Dosya formatı tanımlanmış olan kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	Dosya formatı tanımlanmış olan kaynak yüzdesi $A/B * 100$ A , Dosya formatı tanımlanmış olan kaynak sayısı B , Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır. Bu yüzde tanımlama araçlarını veya tanımlama işlemini geliştirerek artırılabilir (performansı arttırmak için).
Kalite Göstergesi Türü	Koruma
Gösterge Numarası	15
Adı	Formatı tanımlı bir koruma stratejisine sahip kaynak yüzdesi
Amacı	Web arşivinin mantıksal korumasına kurumsal bağlılığı değerlendirmek
Ön Koşullar	• Web arşivindeki toplam kaynak sayısı • Formatı tanımlı bir koruma stratejisine sahip kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	Formatı tanımlı bir koruma stratejisine sahip kaynak yüzdesi $A/B * 100$ A , Formatı tanımlı bir koruma stratejisine sahip kaynak sayısı B , Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.
Kalite Göstergesi Türü	Koruma
Gösterge Numarası	16
Adı	Virüs taraması yapılmış kaynak yüzdesi
Amacı	Diğer koleksiyonlar ve kullanıcı aygıtları karşısında Web arşivinin kullanım güvenliğini değerlendirmek
Ön Koşullar	• Web arşivindeki toplam kaynak sayısı • Virüs taraması yapılmış kaynak sayısı Hesaplama URL veya byte cinsinden yapılabilir.
Yöntem	Virüs taraması yapılmış kaynak yüzdesi $A/B * 100$ A , Virüs taraması yapılmış kaynak sayısı B , Web arşivindeki toplam kaynak sayısı Sonucu tek ondalıklı sayıya yuvarlayın.
Yorumlar	Virüslerin tespiti konusunda, kurumların virüsleri Web arşivlerinden kaldırmayan fakat etkilenen kaynaklara erişimi engelleyen bir politikası olabilir. Hesaplama için kullanılan ölçüm birimi URL veya byte cinsinden raporlanmalıdır.

EK 3: WEB ARŞİVLERİ İÇİN İŞLEVSELLİK TEST ARACI

Web arşivi ile ilgili temel bilgiler:

- İsim:
- URL:
- Üretici:

Web arşivi tarafından hangi fonksiyonlar destekleniyor? Uygulananların tümünü seçin:

Arama parametreleri:

- URL'ye göre ara
- Anahtar Kelime İle Ara
- Boolean araması
- Arama sonuçlarını tarihe göre daralt
- Belirli bir arşivlenmiş Web sitesinde arama yapın
- MD5 parmak izi kullanarak aynı kopyaları arama
- Anahtar kelimeler yerine Web içeriği özelliklerine göre arama yapın (ör. güncelleme sıklığı, türü, medya türü)
- Web arşivine federe bir arama portalı üzerinden erişilebilir
- Kullanıcılar, http yanıt başlıklarını, indirme zamanlarını, aktarım hızlarını ve diğer ilgili bilgileri depolayan arşivlenmiş Web sitelerinin günlük dosyalarına erişebilir.
- Yetişkinlere uygun içeriği filtrelemek için güvenli arama
- Basit arama ve ileri arama arasındaki ayrım

Arama sonuçları:

- Arama sonuçları sorguyla alaka düzeyine göre sıralanır
- Arama sonuçları alan adına göre gruplanmıştır
- Kullanıcılar bir URL'nin arşivlenen tüm sürümlerini görebilirler
- Kullanıcılar, her URL'nin arşivlendiği tarihleri görebilirler
- Her döndürülen sonuçta bir içerik özeti vardır
- Arşivlenen bir Web sayfası belli bir süre boyunca değişip değişmediğini kullanıcılara bildirir
- Kullanıcılar, bir Web arşiv sayfasında hangi tarihler arasında hangi bölümlerin değiştirildiğinden haberdar edilir
- Kullanıcılar, arşivlenen bir Web sayfasının iki sürümü arasındaki farklılıkların bir özetini elde edebilir (ör. yüzde, kelime sayısı, bağlantı sayısı, resim sayısı)
- Kullanıcılar Internet'te olduğu gibi arşivlenmiş Web sitesinde gezinebilirler.
- Arşivlenen her Web sayfasında, referans olarak kullanılacak kalıcı bir tanımlayıcı vardır
- Web arşivi kullanıcıların alakalı sayfaları yazdırmasına olanak tanır
- Web arşivi, çıktıların güvenilirliğinin imzalı beyanını sunar
- Web sitesinin arşivlendiği gösterilir
- Arşivlenen Web siteleri, sitelerin canlı versiyonları ile aynıdır
- Bir Web sitesine nasıl erişileceği ifade edilir (çevrimiçi veya yerinde)

Göz atma:

- Web arşivindeki koleksiyonlar bir hiyerarşi halinde düzenlenmiştir; Kullanıcılar koleksiyonun hiyerarşisine göz atabilir, zaman periyodu ve alan adına göre daraltma yapabilir ve istenilen sonuca göz atabilir

Politikayla ilgili işlevler:

- Kullanıcılar belirli URL'lerin arşivlenmesini isteyebilir
- Belirli Web içeriğine genel erişim, sahibinin isteği üzerine engellenebilir

Kişiselleştirilmiş hizmetler:

- Web arşivi kişiselleştirilmiş kullanıcı arabirimi sunar (kişisel hesap oluşturmaya izin verir). Örneğin, kullanıcıların referans ve tekrar arama için gerçekleştirilen aramalarını kaydetmelerine, bazı Web sayfalarını kişisel bir yer imi klasörüne koymalarına ve bazı belgeleri Web arşivlerinde başka amaçlarla kaydetmelerine olanak tanır.
- Web arşivi, Web arşivindeki güncellemeleri kullanıcılara e-postayla aktif olarak bildirir. Örneğin, belli bir sayfanın güncellendiği veya değiştirildiği derlemeleri ve belirli bir zaman zarfında bir URL'de yapılan belirli değişiklikleri bildirir.
- URL derlendiğinde ve Web arşivi aracılığıyla Web sayfasını kullanmak mümkün hale geldiğinde kullanıcı e-postayla bilgilendirilecektir.

Veri madenciliği:

- Web arşivi, arşivlenen belirli Web sitelerinin belirli bir süre içindeki belirli etkinliklerle nasıl ilişkilendirildiğini gösteren grafikleri döndürebilir.
- Web arşivi arşivlenmiş bir Web sayfası için bağlantı bilgileri (gelen bağlantılar, giden bağlantılar, dâhili bağlantılar) sağlayabilir.
- Web arşivi, kullanıcıların belirli kriterlere (zaman aralığı, alan adı ismi, dil, dosya formatı, tekli veya çoklu üstveriler) dayanarak arşivin bir alt kümesini çıkarmalarını ve çıkarılan alt kümeyi başka bir yerde işleme için aktarmalarını sağlar.

Kaybolan Web sitelerinin yeniden yapılandırılması:

- Kullanıcılar, orijinal Web sitesinin yapısını koruyarak kayıp bir Web sitesinin belirli bir sürümünü yeniden oluşturmak için Web arşivlerini kullanabilir
- Bir Web sitesinin kanıtlanmış sahipleri, kayıp Web sitelerinin belirli bir sürümünü yeniden oluşturmak için Web arşivlerini kullanabilir.

ÖZET

Bilgi çağında yaşanan dijital dönüşümle birlikte geleneksel kültürel miras varlıklarının yerini dijital kültürel miras varlıkları almaktadır. Dijital kültürel mirasın büyük bir bölümünü Internet'in hayatın her alanına girmesi ve aktif bir biçimde kullanımı sonucunda ortaya çıkan Web mirası oluşturmaktadır. İçinde bulunduğumuz çağın hafızası ve dijital kültürel mirasımızın bir kaydı olan Web mirasını koruma ve gelecek nesillere aktarma sorumluluğu ise kültürel miras kurumlarına aittir. Web mirasını koruma amacıyla kültürel miras kurumları Web arşivleme uygulamalarını hayata geçirmekte ve bu süreçte koruma sorumluluğunu üzerlerinde en çok hisseden kurumlar olan ulusal kütüphaneler ön plana çıkmaktadır.

Araştırmamızın amacı kültürel mirasın korunması bağlamında dünya genelinde ulusal kütüphanelerin Web arşivleme uygulamalarının incelenmesi, karşılaştırılması ve Web arşivleme faaliyetlerinin Türkiye Milli Kütüphanesi'nde uygulanmasının gerekliliğini ortaya koymaktır. Bu bağlamda ilk olarak kültürel miras ve Web arşivlemeye ilişkin konular detaylı olarak ele alınmıştır. Ardından dünya genelinde ulusal kütüphanelerin Web arşivleme uygulamaları incelenmiş ve Türkiye Milli Kütüphanesi'ne yönelik bir değerlendirmede bulunulmuştur. Araştırma kapsamında sunulan araştırma problemlerini yanıtlama amacıyla ortaya konulan hipotezlerin tamamı doğrulanmış ve söz konusu süreçte betimle yönteminden ve gözlem tekniğinden yararlanılmıştır.

Araştırmada Türkiye Milli Kütüphanesi'nin ulusal kültürel mirasının önemli bir bölümünü oluşturan Web mirasını koruma gayesiyle Web arşivleme uygulamalarına faaliyet süreçlerinde yer vermesi gerektiği sonucuna ulaşılmıştır. Bununla birlikte, Web arşivleme sürecinde diğer ulusal kütüphanelerle karşılaştırıldığında Türkiye Milli Kütüphanesi'nin seçeceği derleme stratejisinin ulusal alan adı, seçimli ve olay eksenli derleme; erişim yönteminin URL ve tam metin arama olması gerektiği belirlenmiştir. Bununla birlikte geçen yıllar içerisinde Web arşivleme uygulamalarına faaliyet süreçlerinde yer veren ulusal kütüphanelerin sayısında bir artış yaşandığı sonucuna varılmıştır. Araştırmada ayrıca ülkelerin küresel rekabet, insani gelişme ve bilgi ve iletişim teknolojileri gelişmişlik düzeyi ile ulusal kütüphanelerinin Web arşivleme uygulamalarına yer vermesi arasında bir ilişki bulunduğu sonucuna ulaşılmıştır.

Anahtar Sözcükler

Kültürel miras, Web arşivleme, ulusal kütüphanelerde Web arşivleme uygulamaları

ABSTRACT

Along with digital transformation in the Information Age, traditional cultural heritage assets are replaced by digital cultural heritage assets. Web heritage forms a large part of the digital cultural heritage, as a result of the Internet's penetration into every aspect of life and its active use. The responsibility of preserving the Web heritage which is the memory of the age we are in and a record of our digital cultural heritage and transferring it to future generations belongs to cultural heritage institutions. In order to protect the Web heritage, cultural heritage institutions realize their Web archiving practices, and the national libraries, which are the most sensible institutions on the responsibility of protection, are at the forefront.

The aim of our research is to examine and compare the Web archiving practices of national libraries around the world in the context of cultural heritage preservation and to test whether the Web archiving activities can be implemented in the National Library of Turkey. In this context, firstly, issues related to cultural heritage and Web archiving are discussed in detail. Then, the Web archiving applications of the national libraries were examined and an evaluation was made about the National Library of Turkey. All of the hypotheses for responding to the research problems presented in the research are verified and the descriptive method and observation technique have been used in the process.

The research has reached the conclusion that the National Library of Turkey should include Web archiving practices in its activities in order to protect the Web heritage which constitutes a significant part of the national cultural heritage. However, when compared to other national libraries in the Web archiving process, it has been determined that the harvesting strategy chosen by the Turkish National Library should be national domain, selective and event-based harvesting and the access method should be URL and full-text search. However, in the last years it has come to the conclusion that there has been an increase in the number of national libraries involved in the activities of Web archiving. The research also reached the conclusion that there is a relationship between the level of development of countries' global competition, human development and information and communication technologies and the involvement of national libraries in Web archiving practices.

Keywords

Cultural heritage, Web archiving, Web archiving practices in national libraries

ÖZGEÇMİŞ

Metehan Samancı 18.01.1987 tarihinde Ankara’da doğdu. 2006 yılında Hacettepe Üniversitesi Edebiyat Fakültesi Bilgi ve Belge Yönetimi ve 2008 yılında Anadolu Üniversitesi İktisat Fakültesi Kamu Yönetimi Lisans Programına başladı ve 2012 yılında her iki fakülteden de mezun oldu. 2013 yılında atandığı Ankara Üniversitesi Ziraat Fakültesi Muhlis Erkmen Kütüphanesi’nde kütüphaneci olarak çalıştı. 2014 yılında Kültür ve Turizm Bakanlığı’na geçiş yapan Metehan Samancı Millî Kütüphane Başkanlığı’nda Kültür ve Turizm Uzmanı olarak görev yapmaktadır. Temel ilgi alanları yüzme ve futbol olup iyi derecede İngilizce bilmektedir.