



Bilgi Yönetimi Dergisi

Cilt: 2 Sayı: 1 Yıl: 2019

<https://dergipark.org.tr/by>



Hakemli Makaleler

Araştırma Makalesi

Makale Bilgisi

Gönderildiği tarih: 27.05. 2019

Kabul tarihi: 17.06. 2019

Yayınlanma tarihi: 28.06. 2019

Article Info

Date submitted: 27.05.2019

Date accepted: 17.06.2019

Date published: 28.06.2019

Anahtar sözcükler

*büyük veri, yapay zeka,
elektronik belge ve arşiv
yönetimi, arşivsel değer*

Keywords

*big data, artificial
intelligence, electronic
records and archive
administration, archival
value*

DOI numarası

10.33721/by.570634

ORCID

0000-0002-5763-0770 (1)

0000-0003-2393-5207 (2)

Belge ve Arşiv Yönetimi Süreçlerinde Büyük Veri Analitiği ve Yapay Zeka Uygulamaları

*Big Data Analytics and Artificial Intelligence Applications in
Records and Archive Administration Processes*

Mehmet Oytun CİBAROĞLU

*Bursa Teknik Üniversitesi Bilgi İşlem Daire Başkanlığı Öğretim Görevlisi,
oytun.cibaroglu@btu.edu.tr*

Bahattin YALÇINKAYA

*Marmara Üniversitesi Bilgi ve Belge Yönetimi Bölümü Öğretim Üyesi,
yalcinkaya@marmara.edu.tr*

Öz

Son yıllarda sıkça duyulan kavramlardan olan yapay zeka ve büyük veri, insan hayatının neredeyse her alanına etki edecek düzeye gelmiştir. Tüketici davranışlarından kişilik analizlerine, trafik problemlerinin çözümünden uzay çalışmalarına kadar birçok alanda kullanılan yapay zeka teknolojilerinin uygulama alanlarından biri de son dönemde belge ve arşiv yönetimi olmuştur. Yapılan çalışmalar ve yürütülen projeler, çok büyük miktarda içerik barındıran arşivler, kütüphaneler ve bilgi merkezlerinin, insan eliyle gerçekleştirilen süreçler yerine artık daha fazla doğruluk ve neredeyse sıfır hata ile işleyen otomatikleştirilmiş yazılımlar kullanmak istediklerini göstermektedir. Uygulama süresinin görece uzun olduğu işler artık terkedilmeye başlanmış ve yapay zeka temelli otomatikleştirilmiş süreçler, işlem sürelerini dakikalara hatta saniyelere indirmiştir. Bunun doğal bir sonucu olarak da bilgiye hızlı erişim sağlanırken, stratejik kararların alınmasında bilginin etkinliği de artmıştır. Bu çalışmada, yapay zeka ve büyük veri kavramları incelenmiş; bu teknolojilerin belge ve arşiv yönetimi süreçlerine uygulanabilirliği konusunda yapılan uluslararası çalışmalara değinilmiş ve alan uzmanlarının bu teknolojiler karşısında nasıl bir pozisyon alması gerektiği konuları ele alınmıştır. Bilgi ve Belge Yönetimi disiplini ile bu disiplinin ilişkide olduğu diğer bilimsel alanlardaki (Veri Bilimi, Yönetim Bilişim Sistemleri, Bilgisayar Bilimi ve Mühendisliği) literatür taramasından elde edilen veriler neticesinde örnek olaylar ve düşünceler incelenmiştir. Son zamanlarda otomatik dosya kodu sınıflandırma ve sevk işlemleri, otomatik arşivleme ve tasfiye süreçlerinin tasarımı ile ilgili çalışmalar yapılmaktadır. Belge yönetiminde yapay zeka ve büyük veri analitiğinin uygulanabilirliğine dair Türkiye’de yapılan çalışmaların sayısı oldukça sınırlıdır. Yapılan bu çalışma ile alandaki eksikliğin giderilmesine katkı sağlanması amaçlanırken, kurumsal belgelerin saklama süreçlerinde ve arşiv değeri olan malzeme seçiminde yapay zekanın Türkiye’de kullanım koşulları açıklanmaya çalışılacaktır. Aynı zamanda açık devlet verilerinin büyük veri analitiği aracılığıyla halkın ve ilgi gruplarının hizmetine sunulmasıyla ilgili uygulamalar ortaya konulacaktır. Çağın gereklerini yakalamak adına, Türkiye’de de bu alanda çalışan konu uzmanlarının ve koordinatör kurumların mutlaka bu teknolojilerden faydalanması gerektiği anlaşılmaktadır.

Abstract

Artificial intelligence and big data, which are frequently heard in recent years, have reached the level that affect almost every aspects of human being. Recently, records and archive management has been one of the application fields of Artificial Intelligence (AI) technologies used in many fields from consumer behavior to personality analysis, from traffic problems to space studies. The studies and projects show that archives, libraries and information centers from a

huge amount of content bars now want to use automated software that processes more accuracy and almost zero errors instead of man-made processes. Processes which are relatively long is now being abandoned and automated processes based on artificial intelligence reduced processing times to minutes or even seconds. As a natural consequence, while providing rapid access to information, the effectiveness of information in making strategic decisions has increased. In this study, the concepts of AI and big data are examined; international studies on the applicability of these technologies to records and archive management processes and the issues that the experts should take against these technologies are discussed. As a result of the information obtained from the literature review in the other scientific fields, (Data Science, Management Information Systems, Computer Science and Engineering) related with the discipline of Information and Records Management, case studies and thoughts were examined. Recently, automatic file code classification and referral processes, automatic archiving and design of disposal processes are being carried out. Regarding the applicability of AI and big data analytics number of studies records management in Turkey is quite limited. While this study aims to contribute to remedy the deficiencies in the field, it is also tried to explain that the process of preserving corporate records and the selection of archival material which has an archival value using in Turkey in the basis of AI. At the same time, the implementation of open government data to the public and interest groups through big data analytics will be introduced. In order to capture the requirements of the age, it is understood that the experts working in this field and coordinator institutions surely should benefit from this technology in Turkey.

1. Giriş

Yakın tarih boyunca arşivcilik ve belge yönetimi, teknolojik değişimlerden en çok etkilenen mesleklerden olmuştur. 19. yüzyılın ikinci yarısından itibaren müstakil bir yapıda faaliyet göstermeye başlayan modern anlamdaki arşivcilik, 1950'li yıllarda işletme ve yönetim alanında belge yönetimi adında yeni bir meslek grubunun kabul edilmesini sağlamıştır. Arşivcilik uygulamaları, teknolojik dönüşümün etkisi ile daha açık bir yapıya bürünmüş ve belge yönetimi ile aralarında bir iş bölümü oluşmuştur. 1990'ların başında yaşanan bilgisayar devrimi, bu iki meslek grubunun metodoloji ve bilimsel temeller ekseninde bazı yenilikleri tasarlama zorunluluğunu ortaya çıkartmıştır. 1990'ların sonunda ise tamamen elektronik ortamda oluşturulan belgeler için elektronik belge yönetimi kavramı işletme biliminin uygulamalarından ziyade bilişim sistemleri uygulamalarına yaklaşmıştır. Takip eden yıllarda elektronik arşiv kavramı ise önceleri bilgisayar mühendisliği uygulamaları ile anılırken daha sonra endüstri 4.0'ın etkisi ile veri bilimi ile anılmaya başlamıştır (Marciano vd., 2018, s. 197).

Günümüzde özellikle mobil cihazlar, internet sayfaları, kişisel bloglar, sosyal medya uygulamaları ve sabit veya hareketli cihazlarda bulunan sensörler vasıtası ile üretilen veriler inanılmaz boyutlara ulaşmıştır. Dünya çapında binlerce exabayt büyüklüğünde olan ve her geçen zaman katlanarak artan bu verileri yönetmek ve depolamak, işlemek, bilimsel veya ticari kullanım amacıyla analiz etmek ve sonuçlarını açıklayabilmek oldukça zordur. Bu tür zorlukları aşmak adına son yıllarda çeşitli yöntemler geliştirilmiştir. Büyük veri analitiği adını alan bu yöntemler ile yüksek hacimli verinin seçilmesi, depolanması ve işlenmesi sağlanmaktadır. 2020 yılı itibariyle internete bağlı yaklaşık 20-100 milyar arası cihaz olacağı tahmin edilmesinin yanında (Government Office for Science, 2014, s. 14), bu büyüklükteki bir ağın ürettiği verilerin analiz edilmesi, büyük verinin daha kolay anlaşılmasını sağlayacaktır. 2. Dünya Savaşı'ndan sonra araştırmalarına başlanan yapay zeka disiplini ise insan zekasının çok iyi tanımlanabileceği ve bir makinenin bu zekayı taklit edebileceği fikrinden yola çıkılarak oluşturulmuştur. Günümüzde spor müsabakaları, sağlık, otomotiv, video oyunları, finans ve ekonomi tahminleri, sürücüsüz (otonom) araçlar, insansı robotlar ve öğrenebilen bilgisayarlar vb. birçok alanda yapay zeka kullanılmaktadır. Yapay zeka ayrıca büyük veri analitiğinde de uygulanmaktadır.

Belge yönetimi disiplini ise en son teknolojik gelişmeleri takip etmek bir gereklilik olarak karşımıza çıkmaktadır. Bu durumda; belgelerin yaşam döngüsünde, arşivleme süreçlerinde ve uzun süreli saklama dönemlerinde büyük veri analitiği ve yapay zeka teknolojilerinden faydalanmak kaçınılmazdır. Bir yandan yapay zeka uygulamaları ile belge/arşiv süreçlerin otomatikleştirilmesi, diğer yandan büyük veri analitiği ile bu süreçlerden elde edilen verilerin anlamlandırılması neticesinde bilgiye erişim, katma

değer olarak sağlanmış olur. Bu çalışmada; büyük veri, büyük veri analitiği ve yapay zeka kavramları incelenmiş, bu teknolojilerin belge yönetimi ve arşiv süreçleriyle ilişkisi üzerinde durulmuş ve uygulama örnekleri açıklanmaya çalışılmıştır. Bunun yanında belge yöneticisi ve arşiv uzmanlarının bu teknolojileri kullanma konusunda alması gereken aksiyonlar belirtilmiştir.

2. Kil Tabletten Elektronik Belgeye Büyük Veri

Tarihi kökleri itibariyle arşivcilik, antik çağlardan beri var olan bir meslektir. Arşiv biliminin teorisyenlerinden Ernst Posner'ın "Antik Dünyadaki Arşivler" kitabında belge yönetimi ifadesi hiç geçmemesine rağmen MÖ. 4000 yıllarına kadar uzanan ve günümüzde de varlığını koruyan arşivcilik prensiplerinden bahsedilmiştir. Belge yönetimi kavramı arşivcilik kavramına göre henüz çok yeni bir kavram olarak kabul edilmektedir. 2. Dünya Savaşı'nın ardından enformasyon teknolojilerinin yaygın gelişimi ve buna bağlı olarak artan bilgi miktarı, belge yönetimi kavramının tartışılmasına ve müstakil bir yapı olarak ortaya çıkmasına zemin hazırlamıştır (Dollar, 1993). Arşivcilik ve Belge Yönetimi çalışmalarının temelinde yatan teori, metodoloji ve uygulamaların büyük ölçüde farklılık gösterdiği konusu, bu alanlarda çalışmalar yapan bilim insanları tarafından kabul görmüştür (Dollar, 1993, s. 38). Gerald Brown bu farklılığı açıklarken teknolojinin etkisi ile birlikte belge yöneticisinin bir işletme bilimine yakın, arşivcinin ise temelde bir tarihle birlikte olduğunu ifade etmiştir (Dollar, 1993, s. 38).

Enformasyon teknolojilerindeki gelişimin etkisi ile bilgi miktarındaki artış, belge üretiminin ve dolaylı olarak arşiv malzemesinin şekillenmesine etki edecek bazı durumların oluşmasına yol açmıştır. 1940'lı yıllarda, üretilen her materyalin arşiv materyali olacak kadar değerli olamayacağı konusu tartışılmaya başlanmıştır. 1940'lı yıllarda Avrupa kıtasında başlayan bu tartışmalara Amerikan Ulusal Arşivlerinin kurulması ile yeni bir boyut daha eklenmiştir. Amerika Birleşik Devletleri Ulusal Arşivi (NARA), arşivcilerin daha büyük yığınlarla çalışmasının önüne geçmek için bir yönetmelik yayınlayarak kalıcı ve tarihi değeri olmayan belgelerin envanterlerini Kongre'ye sunma yetkisini arşivcilere vermiştir. 1940'lı yılların başında müfettiş yardımcısı olarak çalışan bir arşivci olan Philip C. Brooks, "Hangi Belgeleri Saklayacağız?" başlıklı bir makalesinde belgelerin üretimi, yönetimi ve saklanması gibi süreçlerde hangi aşamalara ve nasıl bir işlem prosedürüne tabi olacağı ifade edilmiştir. Daha sonra, bu fikir, belgelerin yaşam döngüsü modeli olarak literatürde yer almıştır (Dollar, 1993, s. 39). Başka bir deyişle belgelerin aktif kullanım sürecindeki uygulamalar belge yönetimi aşamasına dâhil olurken, tarihe mal olmuş olanlar veya saklama değeri objektif bir anlam ifade edenlerden arşiv değeri olanların arşivciler tarafından yönetilmesi uygun görülmüştür.

Sonraki on yıllarda bilgi teknolojilerinin gelişimi, bilginin oluşturulmasını, işlenmesini, depolanmasını ve paylaşılmasını büyük ölçüde kolaylaştırmıştır. Bu doğrultuda, insanların çalışma, oyun oynama, sosyalleşme, örgütlenme, öğretme ve hatta savaş yapma biçimlerinde muazzam değişiklikler olmuştur. Bu noktada bilişim teknolojileri, arşivcilerin çalışmaları üzerinde nihai etkisini göstermeye başlamıştır. Bu etki aslında bir paradoksun oluşumunu da sağlamıştır. Bu paradoks; arşivcilerin retrospektif ve geçmiş yıllarda oluşmuş arşiv malzemesini nasıl düzenleyeceği, tanımlayacağı ve araştırmaya açacağı konusuna odaklanmış ama aynı zamanda bilgi toplumunun gelişmeye başladığı 1960'lı yıllarda bilgi ve iletişim teknolojilerine bağlı olarak gelecekte üretilecek potansiyel arşiv malzemesinin seçilmesinde izlenecek yöntemlerin de belirlenmesini zorunlu kılmıştır. Bilgi ve belge yöneticiliği adıyla anılan yeni bir disiplinler arası mesleğin ortaya çıkması da bu paradoksun bir sonucudur. Arşivciler sadece retrospektif malzemelere odaklanarak teknolojiye bağlı olarak üretilen ya da üretilmesi muhtemel belgelerin yönetimi konusunu bilgi ve belge yöneticilerine bırakmışlardır.

Sonraki yıllarda arşivciler ve belge yöneticileri süreç ve yönetim konusunda birbirini tamamlayan ancak birbirinden bağımsız hareket kabiliyeti kazanmışlardır. Belge yönetimi ve daha sonraki yıllarda da arşiv yönetimi konseptini değiştiren gelişmelerin başında; donanım ve depolama olanaklarının artması ile bilgisayarların daha da küçülüp daha geniş depolama alanı sunması, ortaya önceleri elektronik belge yönetimi ve daha sonra tıpkı fiziksel ortamda olduğu gibi elektronik arşiv yönetimi kavramlarının çıkmasını sağlamıştır. Belge yönetimi, 2000'li yıllarda elektronik ortama taşınarak bir anlamda boyut değiştirmiştir. Belge, yapılan her bir faaliyet neticesinde oluştuğu için kurumların elektronik belge yönetimini, dijital kurumsal dönüşüm aracı olarak yapılandırılmaları kaçınılmaz bir hal almıştır. Yaşanan

4. Sanayi Devrimi ile de, 1940'lı ve takip eden 50 yıl boyunca bilgi üretim miktarındaki artışla kıyaslanmayacak şekilde, bilgi ve belge miktarındaki devasa artışa neden olmuştur.

Endüstri 4.0'ın sunmuş olduğu büyük veri, yapay zeka vb. uygulamalar ve fırsatlar, arşivcilik ve belge yönetimini hiç şüphesiz derinden etkilemiştir. Bu etkinin önümüzdeki on yılda çok daha belirgin bir şekilde hissedileceği tahmin edilmektedir. Gerald Brown'un arşivci ve belge yöneticisini tanımlarken kullandığı becerilere; elektronik belge yöneticisi, yönetim bilişim uzmanı ve elektronik arşivcinin de temelde veri bilimcisi olması gerektiğini eklemek mümkündür. Bununla ilgili olarak 2010'lu yıllarda teori ve uygulamaların geliştirildiği ve temelde bilgisayar bilimi ve arşiv biliminin ortak çarpanlarının yer aldığı Hesaplamalı Arşiv Bilimi (Computational Archival Science-CAS) literatürdeki yerini almıştır (Lee, Zhang, Chen, Spencer, Cruz, Hong ve Marciano vd., 2016).

2.1. Veri, Doküman ve Belge İfadelerini Yapay Zeka ve Büyük Veri Kavramları Işığında Yeniden Düşünmek

Yapay zeka ve büyük veri kavramlarına geçmeden önce temel tanımların açıklanması faydalı olacaktır. Arşiv ve Belge Terminolojisi sözlüğüne göre veri; analiz edilmemiş ve orijinal haline dokunulmamış gerçekler, fikirler veya ayırık bilgi parçaları olarak tanımlanmaktadır (Moses, 2005, s. 102). Veri kavramı; çoğu zaman, bilgilerin en küçük formu olarak bulunan haline atıfta bulunmak için kullanılır. Örneğin; sentezlenmemiş veya herhangi bir yorum eklenmemiş sayılar veya gerçekler olarak bir belirteçten elde edilen ilk okumalar veya bir anketten elde edilenler göstergeler vb. Bu anlamda veriler, bilginin temeli olarak kullanılır; "Ham veri" ifadesi, orijinal verileri "rafine veri" den ayırt etmek için kullanılmaktadır. Veriler, elde edildiği ortamlardan bağımsızdır ve bir ortama kaydedilene kadar somut değildir. Belgelerden, dokümanlardan veya başka bir biçimden elde edilse bile, verinin sahip olduğu içerik, bağlı bulunduğu ortamdan farklı özellikler göstermektedir.

Belge (records); herhangi bir bireysel veya kurumsal fonksiyonun yerine getirilmesi için alınmış ya da fonksiyonun sonucunda üretilmiş, içerik, ilişki ve formatı ile ait olduğu fonksiyon için delil teşkil eden kayıtlı bilgidir (Çiçek, 2009, s. 106). Bir içerik, bağlam ve yapıya sahip; insan hafızasının bir uzantısı olarak veya hesap verebilirliği göstermek için kullanılan ve bazı ortamlarda sabitlenmiş veri veya bilgiler olarak da tanımlanabilir. Bireysel veya kurumsal faaliyetler sırasında oluşturulan veya sabit bir formda olan veri veya bilgi; ileride başvurulmak üzere bu faaliyetin kanıtı olarak saklanmış değerdir. Kaçınılmaz olarak belirli bir belge türüne atıfta bulunurken genel 'belge' teriminin kullanılması karışıklığa yol açar, çünkü belirli bir türün özellikleri genel olarak kullanılan terim tarafından iletilemez. Belgenin yasal tanımı; ilgili yasalara, kurallara ve düzenlemelere bağlı olarak değişebilir. Ancak belge ifadesi daha çok hukuki anlamda kullanılan ve içeriğinin değiştirilmesi mümkün olmayan, kayıtlı ve ispatlanabilir materyaller için kullanılır. Bu durumda belgelerde olmazsa olmaz temel unsurdan en önemlisi üreticisinin belli olmasıdır. Başka bir deyişle belgeyi üretenin ve içerikten sorumlu olanın belgeyi imzalamış olması şartı aranır (Kandur, 2006).

Doküman (document) ise; herhangi bir yazılı veya basılı çalışma; bir yazı olarak tanımlanmaktadır (Moses, 2005, s. 126-127). Ayrıca; bazı ortamlarda sabitlenmiş bilgi veya veriler; bazı ortamlarda sabit olan, ancak resmi kaydın bir parçası olmayan kayıt dışı bilgi veya veriler olarak da tanımlanabilir. Doküman; geleneksel olarak kâğıda sabitlenmiş metin anlamına gelir ve ayrıca tüm medyayı ve formatları (fotoğraflar, çizimler, ses kayıtları ve videolar ile kelime işlem dosyaları, elektronik tablolar, web sayfaları ve veritabanı raporları) içerebilir. Tıpkı belgeler gibi, dokümanlar da geleneksel olarak içerik, bağlam ve yapıya sahiptir. Ancak, bu nitelikler elektronik dokümanlarda değişebilir. Elektronik formatlar, bilgileri üç boyutlu olan veya doğrusal olmayan bir yapıya sahip karmaşık katmanlarda sunabilir. Özellikle elektronik ortamda dokümanların içeriği sabit değildir, zamanla değişebilir. Daha çok bilgi amacıyla üretilmiş ve belge kadar kesin delil hükmü taşımayan materyaller için doküman ifadesini kullanmak mümkündür. Bu dokümanlar, geleneksel ve sabit belgelerden ayırt edilmesi için "dinamik dokümanlar" olarak da tanımlanmaktadır. Bazı bağlamlarda, dokümanlar belgelerin ikincil kopyaları ve ticari faaliyetlerle doğrudan ilgili olmayan materyaller gibi kanıt niteliği olmayan bir öge olarak tanımlanır. Bu anlamda, dokümanlar, genellikle saklama planlarına dahil edilmez ve izinsiz olarak imha edilebilir. Bazı durumlarda ise, bir belge ile doküman arasında çok açık bir ayrım vardır. Örneğin, Amerika Birleşik Devletleri'ndeki hukuk davalarında, dokümanlar, federal kanıt kurallarına göre ticari veya resmi belge tanımına girerse delil olarak kabul edilebilmektedir (Moses, 2005). Belge

ve doküman arasındaki farklar ise Emery (2003, s. 3) tarafından şu şekilde açıklanmıştır: Bir doküman, çoğu kurumda kullanılan, yapılandırılmamış form olarak kabul edilen (yapılandırılmış belgelerin aksine bazı durumlarda farklı elektronik belgelerin içine yerleştirilebilen) temel iletişim aracıdır. Belgeler ise; bir kurum veya kuruluşun işlevlerine, politikalarına, kararlarına, prosedürlerine, operasyonlarına veya diğer faaliyetlerine veya bunların içindeki bilgilerin değeri nedeniyle kanıtlar sunar. Dokümanlar belge olarak değerlendirilmez, fakat yapılandırılabilir. Belgeler ise sabit bir formda olduklarından yapılandırılmaz. Dokümanlar, belgeler kadar sıkı bir sürece sahip değildir (Emery, 2003, s. 3).

Arşivci ve belge yöneticisi açısından bakılacak olursa belge, ispat gücü açısından hukuki olarak geçerli ve taşıdığı bilginin korunduğunu garanti altına tutan materyaldir. Doküman ise bilgi amaçlı üretilmiş ispat gücü açısından dolaylı delil vasfı taşıyan, üzerinde değişiklik yapmanın mümkün olduğu, daha çok bilgi taşıyıcılarının genel adıdır. Veri ise henüz işlenmemiş ham bilgi olarak kabul görmektedir. Arşivci veya belge yöneticisi artan veri, bilgi, doküman ve belge miktarına bakmaksızın ve bunların yönetilmesi ile ilgili herhangi bir seçim yapmaksızın büyük bir evrenle karşı karşıyadır. Bu evrene özellikle teknolojinin etkisi ile yönetilmeyi bekleyen çok miktarda materyal dâhil olmuştur.

2.2. Büyük Veri ve Yapay Zeka Kavramları Bağlamında Elektronik Arşivlerin Sahip Olması Gereken Genel Özellikler

Geleneksel belge ve dokümanın anlam, algı ve tanımları, elektronik arşivlerin varlığıyla değişmiş ve bu da, ilgili kavramların yeniden yorumlanıp anlaşılması gerekli kılmıştır. Elektronik arşivler gerek varlığı hukuki çerçevede ispatlanabilir, güvenilir ve özgün belgeleri korumak ve gerekse bilgi değeri taşıyan doküman, veri ve görsel-ışitsel materyalleri kapsayıcı bir misyona sahiptir. Üstelik sadece doğuştan elektronik ortamdakileri değil fiziksel ya da analog ortamda bulunan materyalleri de kapsayıcı bir özelliktedir. Elektronik arşivlerin büyük veri ve yapay teknolojileri bağlamında sahip olması gereken özelliklerden bazılarını şöyle özetlemek mümkündür:

- Arşiv malzemesinin sisteme yüklenmesi ve aktarılması ile ilgili belirgin metotlar oluşturulması ve sürdürülebilir bir yapının kurgulanması,
- Arşiv sisteminin etkin bir biçimde kullanılmasını sağlayacak ve erişim haklarını gözetecek bir erişim politikasının oluşturulması,
- Elektronik arşivdeki hassas bilgi, belge ve verilerin korunmasını sağlayacak şifreleme ve maskeleyen mekanizmasının geliştirilmesi,
- Elektronik arşive dahil edilen belgelerin değişmezliklerini garanti altına alabilecek, onları otantiklik ve güvenilirlik açısından uzun süreli koruyabilecek güvenilir bir yapının inşa edilmesi,
- Elektronik arşivden sorgulama yapılabilecek gelişmiş ve temel arama yöntem ve senaryolarının hazırlanması,
- Elektronik arşiv malzemesinin mevcut sistemden başka sistemlere aktarılmasını sağlayacak formatların ve donanımların desteklenmesi,
- Kalıcı değere sahip olmayan, emaneten ve belirli bir süre için elektronik arşivde bulunan arşiv malzemesinin imha ve tasfiye işlemlerine dair iş ve işlem süreçleri ile ilgili kuralların tanımlanması,
- Elektronik arşiv malzemesinin verimli bir şekilde saklanmasını sağlayacak teknik ve teknoloji politikalarının oluşturulması,
- Elektronik arşiv malzemesinin hacminde artış olması ihtimaline karşın ölçeklenebilir bir yapıda mimarinin kurgulanması,
- Teknolojik eskimeye karşı tedbir alınması vacak tedbirler ve arşiv malzemesine zaman içerisinde sürdürülebilir erişimi sağlanması (Mahon, 2015, s. 41-42).

Elektronik arşivlerin sahip olması gereken bu özelliklerin, organizasyonel bağlamda kalıcı olarak benimsenmesi, yapay zeka ve büyük veri teknolojileri ile desteklenecek bir yeni oluşumda süreklilik, inovasyon ve organizasyonel hafızanın etkin bir biçimde işlerlik kazanmasını sağlayacaktır.

3. Büyük Veri Analitiği ve Yapay Zeka Kavramlarının Değerlendirilmesi

3.1. Büyük Veri ve Yapay Zeka Kavramları

Büyük veri kavramı, literatürde ilk kez Michael Cox ve David Ellsworth tarafından 1997 yılında düzenlenen 8. IEEE Görselleştirme Konferansı'nda "Application Controlled Demand Paging for Out-of-Core Visualization" adlı makalede kullanılmıştır (Aktan, 2018, s. 3). Makalede veri setlerinin büyüklüğünden, bilgisayar sistemlerinin bellekleri ve diskleri aşırı kapasite ile doldurduğundan bahsedilmiş ve karşılaşılan bu soruna "büyük veri problemi" adı verilmiştir (Cox ve Ellsworth, 1997, s. 235). Literatürde büyük veri ve yapay zeka kavramlarına dair çeşitli tanımlamalar mevcuttur. Sakyi'ye göre (2016, s. 1) büyük veri, geleneksel hesaplama tekniklerinin veya algoritmaların üzerinde çalışmadığı hem yapılandırılmış hem de yapılandırılmamış büyük karmaşık veri kümelerini ifade etmektedir. Boyd ve Crawford'a göre (2012, s. 663) ise büyük veri, aşağıdaki etkileşimlere dayanan kültürel, teknolojik ve bilimsel bir olgudur. Bu yaklaşımla büyük verinin, teknolojik, analiz ve mitolojik boyutu ele alınabilir:

- Teknoloji: Büyük veri kümelerini toplamak, analiz etmek, birbiriyle ilişkilendirmek ve diğer veri setleri ile karşılaştırmak için hesaplama gücünün ve algoritmik doğruluğun en üst düzeye çıkarılmasını,
- Analiz: Ekonomik, sosyal, teknik ve yasal iddialarda bulunmak için kalıpları tanımlamada büyük veri setlerinden faydalanılmasını,
- Mitoloji: Büyük veri setlerinin; gerçeğin, tarafsızlığın ve doğruluğun havasını taşıyan, daha önce imkansız olan sezgileri ortaya çıkaran yüksek bir zeka ve bilgi formu sunduğuna dair yaygın bir inancı temsil etmesi.

Büyük veri, üç temel altyapı bileşeni gerektirir: Verinin düzenlenmesi, depolanması ve erişilebilir hale getirilmesi için bir platform; büyük ölçekli veri setlerini işleyebilecek bilgi işlem teknolojisi ve gücü; yapılandırılmış ve kullanılabilir veri formatları.

Levi (2013, s. 34), sosyal bilimlerde "veri" terimini kullanmanın "oksimoron" olduğunu belirtmiş ve bu alanda büyük verinin aslında her zaman var olduğu iddiasında bulunmuştur. Ona göre belgeler, nesnelere, resimler, fotoğraflar, kitaplar ve bunların içerdikleri bilgilerle ilgili yığınlar, kaçınılmaz bir büyük veri kaynağı olarak görülmelidir.

Literatürde, büyük veri kavramına tarihsel anlamda 2 farklı yaklaşım getirilmiştir. Bunlar:

- Retrospektif büyük veri: Kütüphanelerde, arşivlerde, müzelerde ya da dijitalleştirme yoluyla geriye dönük büyük veri yaratmalarına katkıda bulunan özel koleksiyonlarda bulunan analog kaynaklar,
- Gerçek büyük veri: Ağa bağlı (interaktif) büyük veriler (Twitter ve Facebook gibi) (Levi, 2013, s. 33).

Yapay zeka ise ilk kez 1956'da Stanford Üniversitesi Bilgisayar Bilimleri Bölümünden John McCarthy tarafından "akıllı makineler yapmanın bilimi ve mühendisliği" şeklinde tanımlanmıştır (Grewal, 2014, s. 10). Bu ilk tanım, yapay zekanın mekanik düzeydeki uygulanabilirliğine atıf yapmıştır. Bu alan, insanın merkezi bir özelliği olan zekanın, bir makine tarafından benzeşim gösterebilecek kadar kesin bir şekilde tanımlanabileceği üzerine kurulmuştur (Grewal, 2014, s. 10). Yapay zeka, ilk tanımlarından farklı olarak, bilgi ve iletişim teknolojilerinin (BİT) gelişmesiyle soyut bir sistem tasarımına doğru değişime uğramıştır. Bu tanımlar ışığında yapay zekanın genel amaçlarını; insan zekasını makinelere aktarmak ve gelişmiş bir sistem kurmak olarak açıklamak mümkündür.

Arşivler genellikle büyük miktarda veri ve çeşitli koleksiyonlar barındırmasına rağmen sınırlı insan kaynağına sahiptir. Sınırlı insan kaynağı arşivlerin istenilen düzeyde düzenlenmesine ve tanımlanmasına yönelik bazı engellerin oluşmasına ve bazı hatalı arşiv iş ve işlemlerinin yapılmasına neden olabilmektedir. Bu tür olumsuzluklar, yapay zeka gibi bir takım yeni teknikleri ve teknolojik imkanları kullanmayı gerekli kılmaktadır. Yapay zeka uygulamaları, belge yönetim süreçlerinden başlayarak arşiv yönetimi süreçlerine kadar etkin ve uygun çözümler üretebilecek bir yapıyı içerisinde barındırır. Aynı

¹ İki zıt anlamlı kelimenin bir arada kullanılması

zamanda verinin ilk elden ve tek olduğu arşivlerde büyük veri uygulamaları ile daha sağlıklı veri analizleri yapılabilir. Yapay zeka uygulamasının yaygınlaşmaya başlaması ile birçok alanda çalışmaların yürütüldüğü bilinmektedir. Özellikle arşivler açısından, gerek doğrudan elektronik doğan, gerekse fiziksel ortamlardan dijitalleştirilen büyük miktarda elektronik materyal dikkate alındığında, bu veri yığınlarının anlamlandırılması ve analizi için yapay zeka teknolojisinin kullanılması büyük bir kolaylık sağlamaktadır. Elektronik arşivlerde tarihe tanıklık etmek için üretilmiş belgelerin söylediklerini yorumlamak yapay zeka teknolojisi ile daha kullanışlı bir hale gelecektir.

Aynı zamanda arşivleme sürecinden önceki aşama olan belge yönetimi aşaması için yapay zeka teknolojisi, belgelerin saklama planlarının otomatikleştirilmesi, sınıflandırmanın yapay zeka ile yapılması, açık devlet verilerinin halka sunulması ve belge yaşam döngüsü yapay zeka tarafından kontrol edilmesinden bahsedilmektedir. (Jie, Linqian, Jing ve Huan, 2016; McDonald ve Leveille, 2014; Janssen ve Hoven, 2015; Yong, Liming ve Yongsheng, 2015)

3.2. Yapay Zeka, Belge Yönetimi ve Arşiv İlişkileri

Literatürde yapay zeka uygulamalarının belge yönetimi ve arşivleme süreçlerine uygulandığı çeşitli çalışmalar mevcuttur. Bu çalışmalardan biri Rolan, Humphries, Jeffrey, Samaras, Antsoupova ve Stuart tarafından (2018) Kasım-Aralık 2017'de gerçekleştirilen New South Wales Devlet Arşivleri (NSWSAR) pilot çalışmasıdır. Bu pilot çalışmanın temel amacı, makine öğrenme algoritmalarını daha önce imha süreçlerine analog olarak uygulanan arşiv belgeleri üzerinde test etmektir. Projede 8784 arşiv belgesi analiz edilmiştir. Birtakım sayısal işlemler uygulandıktan sonra metin görüntülerine iki farklı sınıflandırma algoritması uygulanmıştır: Çokterimli Naif Bayes (Multinomial Naif Bayes-MNB) ve Çok Katmanlı Algılayıcı (Multi Layer Perception-MLP). Elde edilen istatistikler, MLP algoritmasının en başarılı olduğu ve belge yığınının test kısmında maksimum %84 başarı oranı ile süreci tamamladığı görülmüştür. Bu sonuçlar, MLP'nin sınıflandırılmamış ve yapılandırılmamış verilerin sınıflandırılması ve imhası konusunda yardımcı olduğunu göstermektedir.

Türkiye'de ise Ankara Üniversitesinde başlatılan bir proje, yapay zeka-belge yönetimi ilişkisine örnek olarak gösterilebilir. Projede, yapay zekanın bilgi ve belge yönetimi süreçlerine uygulanabilirliği, kurum hafızasının yapay zeka destekli oluşturulması ve geçmişte alınan kararların yeni karar süreçlerine entegre edilebilmesi araştırılmaktadır. Ayrıca yapay zeka algoritmaları tarafından yönetilen iki sanal birim arasındaki yazışmaların gerçekleşmesi ve bu yazışmaların otomatik imzalanması da araştırılan diğer alanlardır (Ünal ve Özdemirci, 2017).

Konuyla ilişkili birkaç uygulama örneğinin Avusturalya'da yapıldığı görülmektedir. Bu çalışmalardan ilki Avustralya Ulusal Arşivlerinde uygulanmakta olan imha ve saklama prosedürlerinin Avustralya Hükümetindeki dijital işletmeyi destekleyecek bir biçimde nasıl oluşturulacağını ve verileceğini araştırmak üzerinedir. Mevcut durumda saklama ve imha kurallarını belgeleyen genel ve özel kayıt ilkeleri yayınlanmaktadır. Süreçte, arşiv ekibi tarafından otomatik üst veri çıkarma, semantik analiz, taksonomi ve ontoloji oluşturma ve bağlantılı veri yaklaşımları incelemiştir. Projenin ilk aşamasında, kavramsal bir model oluşturulmuş ve imha prosedürlerinin semantik analizi yapılmıştır. Ayrıca otomatik sınıflandırma, kümeleme ve indeksleme araçları da dahil olmak üzere çeşitli makine öğrenmesi türleri oluşturulmuştur. İlgili kurumun tüm genel evrak birimlerinin gelecekteki dijital yaklaşımlar için bir pilot çalışma alanı olması beklenmektedir (Rolan vd., 2018). Yine Avusturalya'da; imha ve saklama süreçlerinin makine öğrenme ve yapay zeka yöntemleri ile nasıl oluşturulacağını belirlemek amacıyla bir çalışma başlatılmıştır. Bu çalışmada mevcut belge yönetimi yapısındaki üst verilerin çıkartılması, semantik analiz, taksonomi, ontoloji oluşturma ve bağlı veri (linked data) yaklaşımları incelenmiştir (Rolan vd., 2018). Üçüncü örnek olay Avusturalya Maliye Bakanlığı'nın e-posta sunucularında biriken e-postaların yapay zeka aracılığı ile ayıklanması ve bakanlık ile ilişkili olan tüm kuruluşların bir anlamda hikayelerini ortaya çıkarma ve elde edilen veriyi gelecekte yatırım politikası üzerinde uygulamayı amaçladığı çalışmadır (Rolan vd., 2018).

Yapılan bu çalışmaların yakın bir gelecekte insan müdahalesini kısıtlayan veya tamamen ortadan kaldıran bir forma dönüşmesi muhtemeldir. Belge ve dokümanların yapay zeka tarafından işlenip Bu tahmini destekleyici bir çalışma olarak Oxford Üniversitesinden Frey ve Osborne'un (2017) çalışması

örnek olarak gösterilebilir. Araştırmacılar, Amerika Birleşik Devletlerindeki işlerin %47'sinin yirmi yıl içinde robotlar tarafından değiştirilme riski altında olacağını belirlemiştir (Frey ve Osborne, 2017). Algoritmaların, birçok işte insanlardan daha iyi performans gösterdiği ve bu durumun da “insan niteliklerinin ve yeteneklerinin yüzde 99'unun çoğu modern işin performansı için gereksiz olduğu” şartını yarattığı söylenmektedir (Harari, 2017). En yenilikçi teknolojik girişimcilerinden biri olarak kabul gören Elon Musk bile, defalarca yapay zekanın yaygın ve düzensiz uygulamasından kaynaklanacak olan “sosyal istikrarsızlaşma ve iş kaybı” konusunda uyarmıştır. Bunun yanında, bilgi merkezlerinde çalışan uzmanların yirmi yıl içerisinde yaklaşık %80 oranında robotlarla yer değiştirilebilme ihtimalini hesaplamıştır (Arlitsch ve Newell, 2017, s. 792). Kataloglamanın bile artık zorunlu olarak insanların alanı olmadığı düşüncesi ile hareket edilmeye başlanmıştır. Bu durumu açıklayan en güzel örnek ise Yewno adlı bir şirketin, milyonlarca yayınlanmış makaleyi ayrıştırmak, diğer çalışmalarla kavramsal ilişkileri kurmak ve bu ağı keşif katmanlarına ek olarak sunmak için yapay zekayı kullanmasıdır (Gramatica ve Pickering, 2017). Bugüne kadar Yewno, hepsi tek bir insan yapımı katalog olmayan, 100 milyondan fazla bilimsel makaleyi (Gutierrez, 2017) içselleştirmiş ve organize etmiştir. Yapay zeka, kullanıcılara hizmet etmek için gereken beceri ve araçların karışımını sonsuza dek değiştirerek bilgi merkezlerinin verdiği hizmetleri, kullanıcıların yaşamlarını ve toplulukların dinamiklerini değiştirecektir. Bu değişikliklerin tam olarak ne kadar yayılacağını tahmin etmek zor olmakla birlikte bir değişimin olacağı kesindir.

Büyük verilerin değerini, nasıl manipüle edilebileceğini, görselleştirildiğini ve analiz edildiğini öğrenmek için nicel ve analitik becerilerin edinilmesi gerekmektedir. Bu da bilgi merkezleri ve çalışanları için büyük bir fırsat olarak karşımıza çıkmaktadır. Kısacası, makinelerin insanlık için çalıştırılmasının yolları bulunmaya devam edilmelidir.

3.3. Büyük Veri ve Belge Yönetimi Arşiv İlişkileri

Literatürde büyük verinin belge yönetimi ve arşivleme süreçlerine uygulandığı çeşitli çalışmalar mevcuttur. Bunlardan bir tanesi Cox, Shah, Frederick, Nelson, Thomas, Jansen, Dibert, Kurtz ve Marciano (2018) tarafından gerçekleştirilen Maryland eyaletindeki “Kölelik Mirası Projesi” adlı çalışmadır. Projenin misyonu, Maryland'deki köleleştirme ve Afrika kökenli Amerikalıların özgürlük öyküsünü anlatmaktır. Bunun için Maryland Eyalet Arşivi'nde bulunan on binlerce tarihi arşiv belgelerinden faydalanılmıştır. Eyalet arşivinde araştırmalar yapılırken nüfus sayımından oturma bilgilerine reklamlardan hapisane verilerine kadar 16 farklı kaynaktan bilgiler elde edilmiş; toplanan bu veriler kopyalandıktan sonra, analize hazır hale getirilmesi için birleştirilmesi, temizlenmesi ve bir veritabanında depolanması gerekmiştir. Bunun için de Apache Spark veritabanı kullanılmıştır. Belirli bir otomasyon sürecine bağlanan iş akışları ile belgeler ve ilişkili malzemeler (gazete, dergi vb.) OCR destekli tarayıcılarda taranmış ve ardından transkripsiyon ve veri temizliği yapıldıktan sonra, veritabanına aktarılmıştır. Veri tablolarının tümü başarıyla entegre edildikten sonra bir arayüz tasarlanarak birleştirilmiştir. Bu arayüz aracılığı ile 420.000'den fazla bilgi kümesine (ilçe ve kentlere göre özgürlük belgeleri-sertifika, cinsiyete göre özgürlük kazanma durumu, köle ve özgür olarak doğanların oranı vb.) erişim sağlanabilmiştir.

Büyük verinin arşiv yönetim süreçlerine uygulandığı örnekler arasında ise İngiliz Milli Kütüphanesi bibliyografik kayıtlarına uygulanan 2 farklı çalışma bu konuda göze çarpmaktadır. İlk çalışmada; üç milyondan fazla kayıt R istatistik programı ile analiz edilmiş ve veriler görselleştirilmiştir. Çalışma sonucunda koleksiyonda bulunan materyaller; tarih, bilim dalı, konu vb. alanlara ayrılmış ve hangi kaynakların hangi yılda, hangi konu başlığı altında yoğunlaştığı gösterilmiştir (Uribe, 2018). İkinci çalışma ise 150 yıldan fazla süredir büyüyen kütüphane kataloğundaki verilerin, büyük veri analitiği yöntemleri ile görselleştirilmesi çalışmasıdır. Bu sayede çeşitli süreçler daha kolay tanımlanmış, verilerdeki yeni ve mevcut sorunlar tanımlanarak gerekli iyileştirmeler yapılmıştır. Bunun yanında farklı veri katmanları tanımlanarak dijital erişim de kolaylaştırılmıştır (Prescott, 2013).

4. Büyük Veri ve Yapay Zeka Uygulamalarına Göre Yaşam Döngüsü

4.1. Temel İş Süreçlerinin Analiz Edilmesi

Kurumsal işletmelerde belgelerin sağlıklı bir şekilde ilgili yerlere ulaşması için belirli birtakım iş süreçlerine sahip olunması gerekmektedir. Bu konu belgelerin yaşam döngüsü modeli bağlamında ele alınabilir. Belgelerin yaşam döngüsü; belgelerin üretiminden iş süreçlerine entegre edilmesine, dosyalanmasına, dağıtımından saklanmasına ve imha edilmesine kadar olan süreçlerin toptan ele alınması ile ilgilidir. Belgeler her canlı gibi doğar büyür ve ölür. Belgelerin bir iş için kullanıldığı aşamalar; belgelerin oluşturulduğu ve iş süreçlerine dahil edilip aktif olarak kullanıldığı güncel aşama (current stage); belgelerin organizasyonlarda kullanım değerinin azaldığı yarı-güncel aşama (semi-current stage) ve belgelerin kullanım değerinin olmadığı veya çok az olduğu ancak başka amaçlarla kullanılabilirdiği güncel olmayan aşama (noncurrent stage) olarak sınıflandırılabilir (Shepherd ve Yeo, 2003, s. 5). Belgelerin güncel olmayan dönemdeki kullanımında 2 boyut bulunmaktadır: Nesnel ve öznel boyut. Nesnel boyut; bir belgenin idari mali ve hukuki açıdan değerini belirtir. Öznel boyut ise belgenin tarihi bir değere sahip olup olmadığına karar verilmesidir. Fakat bu duruma dair karar vermek genellikle çok zordur.

Bahsedilen bu süreçlerde yapay zeka uygulamaları, süreçlerin entegrasyonu ve otomatikleştirilmesi bağlamında ele alınabilir. Örnek olarak; kurum dışından gelen veya organizasyon içinde oluşturulan bir belgenin, yapay zeka uygulaması kullanan bir otomatik dosyalama algoritması ile üstveri ve içeriği taranarak analiz edilip, bu analiz sonucuna göre bir dosya numarası verilmesi ve içeriğe göre belgenin otomatik olarak ilgili birime sevk edilmesi düşünülebilir.

Fiziksel belgeler için yukarıda özetlenen süreçler, bilgisayar ortamında depolanan materyaller için her zaman geçerli olmayabilir. Bu tip belgelere erişim için belirli bir yetkilendirme prosedürü gerekmektedir. Bu nedenden ötürü her bir sürecin sistem analizinin yapılması, bu analizler sonucunda gerekli görülen süreçlerin revize edildikten sonra programlanması daha yerinde olacaktır.

4.2. Sınıflandırma, Saklama Planı ve Tasfiye Süreçleri

Arşiv materyallerinin otomatik olarak sınıflandırılması sorunsal literatürde son zamanlarda oldukça sık tartışılmaktadır. Arşive gelen materyallerin konu, tarih, oluşturulma yeri vb. sınıflandırma alanlarının tamamıyla yapay zeka kontrolüne bırakılması yakın bir gelecekte olası görünmektedir. Bu tür bir işlem, arşiv uzmanlarının mevcut işleyişteki hatalarını giderecek, zaman problemlerine çözüm olacak ve doğru materyali sınıflandırarak etkili bir bilgi erişimi sağlayacaktır.

Materyaller, arşive gelmeden önce birtakım değerlendirme ölçütlerine göre “arşive gidecekler” ve “saklanmasına gerek olmayanlar” şeklinde bir ayrıma tabi tutulmaktadır. Bu işlem de ilgili kanun, yönetmelik ve yönerge baz alınarak yapılmakta; arşive kabul edilen materyaller de belirli ölçütlere göre kendi içlerinde sınıflandırılmakta ve bu şekilde yapılan bir işlem sonucunda saklanmaktadır. Bu sürecin otomatik bir hale dönüştüğü düşünülecek olursa; dijitalleştirme işlemi yapılan materyallerin, karakter tanıma (OCR) ve metin sayısallaştırma işleminden geçirilerek bir yazılım aracılığı ile üstveri veritabanındaki verilerle otomatik olarak eşleşip dijital klasörlere nakli sağlanabilir.

Büyük veri ve yapay zeka bağlamında düşünüldüğünde; günümüzde saklama planlarının oluşturulması ve bu planların uygulanması anlamında birtakım kaygılar mevcuttur. Bunlar arasında; saklama ve tasfiye planlarının oluşturulmasında hangi alt süreçler ve kriterler olacağı; bu süreçler ile ilgili resmi programları geliştirmek, onaylamak ve uygulamak için mevcut yaklaşımların ne olduğu ve nelerin değişmesi gerektiği bulunmaktadır (McDonald ve Leveille, 2014, s. 100). Yeni dönemde bu süreçlerin nasıl olacağına karar vermek için, geçmişte kağıt bazlı oluşturulan materyallere uygulanan klasik yöntemden bu güne dek geliştirilen uygulamaları anlamak önemlidir. Bu, büyük veri ve yapay zeka çağında hangi hataların yapılmaması gerektiğini bize gösterecektir.

Belgeler ayıklandıktan sonra, tarihi, ulusal, ekonomik, hukuki vb. değerlere sahip olanlar, uzun süre saklamaya yönelik olarak ayrılıp arşive gönderilmektedir. Saklama süresi belgenin içeriğine ve değerine göre değişmekte olup, bu süreler birtakım kurallar neticesinde belirlenmiştir. Belgelerin oluşturulma

sürecinden itibaren en son adım olan saklanmasına kadar olan işlemler mutlaka belirli bir hiyerarşik düzende yapılmalıdır.

Bir diğer önemli nokta ise saklanmasına karar verilen materyallerin nasıl bir işleme tabi tutulacağıdır. Belgelerin, saklama süreleri boyunca arşivlerde düzensiz olarak yerleştirilmesi veya rastgele sıralanması en son istenecek şeydir. Bu durumda saklama süresi boyunca belirli kurallara sahip iş akışlarının izlenmesi veya oluşturulması gerekmektedir. Son 20-30 yıldır ise sadece fiziksel belgelerin değil, elektronik ortama transfer edilen dijital belgelerin de korunması konsepti düşünülmüş fakat mevcut süreçlerin sadece kâğıt veya fiziksel materyallere yıllardır belirli kalıplarda uygulanmasından ötürü, bunun, dijital belgelere uygulanma konusunda fazlasıyla yetersiz olacağı anlaşılmıştır.

Yaşam döngüsü, belgelerin durağan bir yapıda olmadıklarını; tıpkı canlı organizmalar gibi belgelerin de doğacağını, büyüyeceğini ve öleceğini öngören bir modeldir. Bu fikir ilk defa 1956 yılında Schellenberg tarafından kaleme alınan, belgelerin güncel kullanımı ve sonraki aşamalarını içeren “belgelerin yaşam süresi” konulu yazısında açıklanmıştır. Model; bir belgenin yaşamında farklı zamanlarda almış olduğu aksiyonların nasıl ilerlediğini beş aşamada göstermeyi amaçlamıştır: Oluşturma (creation), sağlama (capture), depolama (storage), kullanım(use) ve arşive transfer (disposal) (Shepherd ve Yeo, 2003, p. 5).

Her ne kadar bu döngüde bir değişikliğe gerek duyulmasa da; her bir aşamanın otomatikleştirilmesi yeniden ele alınabilir. Arşivlerde kesin kural ve yönergelerle bağlı olarak uygulanan ayıklama ve imha süreçlerini, gelişen BİT kapsamında yeniden ele almak gerekmektedir. Bunun nedeni, dijitalleşen dünyada belgelerin de dijitalleşmesi ve fiziksel ortamdaki kayıtlardan farklı olarak alım, sağlama, değerlendirme, saklama, ayıklama ve imha süreçlerinin doğal olarak farklılık gösterecek olmasıdır.

Özellikle çok sayıda arşiv belgesine sahip olan arşivlerde, bu tür süreçler çok zaman almaktadır. Her sene sürekli tekrarlanan bu işlemler, fazlasıyla masrafa yol açmakta; üstelik yanlış kayıt oluşturma veya önemli bir belgeyi yanlışlıkla imha etme gibi durumlarla karşılaşmaktadır. Süreçlerin elektronik ortama taşınması ile var olan sistem kökten değişikliğe uğramış, süreçlerin tamamlanması daha da kısalmış, doğruluk ve güvenilirlik oranı yükselmiştir.

Bu konuda Bentley Tarih Kütüphanesindeki entegrasyon projesi, süreçlerin yeniden tasarımı konusunda önemli bir örnek olarak karşımıza çıkmaktadır. Projenin temel hedefleri arasında; sistemler arası üst verilerin oluşturulması, yeniden kullanılması, içeriğin alınması ve depolanmasının kolaylaştırılması, değerlendirme süreçlerinin elektronik olarak tamamlanması, kod ve dokümantasyon verilerinin diğer ilişkili arşivlerle ve dijital koruma topluluklarıyla paylaşılması bulunmaktadır. Tüm değerlendirme süreçlerinin elektronik ortama taşındığı bu projede ArchivesSpace (ASpace), Archivematica ve DSpace gibi açık kaynak programları kullanılmıştır. ArchivesSpace programı ile dijital arşivlerle ilgili açıklayıcı ve idari meta veriler depolanmış; Archivematica ile içerikler sisteme alınmış ve ASpace'deki tanımlayıcı meta verilerle ilişkilendirilip depolama bilgi paketleri hazırlanmış; DSpace ise koleksiyonlar için bir koruma deposu ve erişim portalı görevi görmüştür. Temelde Gönderi Bilgi Paketi (Submission Information Package-SIP), Arşiv Bilgi Paketi (Archival Information Package-AIP) ve Dağıtım Bilgi Paketi (Dissemination Information Package-DIP) oluşturularak hem değerlendirme süreçleri tamamlanmış hem de kullanıcıların içeriği keşfetmesine, hassas verileri tanımlamasına ve içerdiği bilgiyi anlamak için dosyaları ön izlemesine izin verecek yeni bir işlevsel alan yaratılmıştır. Archivematica; dizin yapısı, dosya biçimi karakterizasyonu, hassas veri tanımlama, içerik ön izlemesi ve etiketleme gibi fonksiyonlar ile içeriğin incelenmesini ve değerlendirilmesini kolaylaştırmaktadır. Bazı önemli adımlar tamamen otomatikken (hassas verilerin tanımlanması ve dosya formatlarının karakterizasyonu gibi), değerlendirme sekmesi farklı araçlar kullanmak ve bilgi kaydetmek için insan müdahalesi gerektirmektedir (Shallcross, 2016).

4.3. Gelecek Öngörüsü: Akıllı Arşivler

Bilgi ve belgelerin saklanmasında geleneksel yaklaşımlar artık pratik değildir. Bu tür yaklaşımlar; yinelenen, eskimiş ve ilgisiz bilgilerle dolmuş çözümler ile sonuçlanır. Bu durum aşırı enerji tüketimine yol açar ve depolama maliyetlerini artırır. Daha küçük ve daha karmaşık bir dünyanın gerek ekonomik gerekse politik zorlukları göz önüne alındığında, hangi bilgilerin saklanması gerektiği ve kullanılan bilgi

altyapısının nasıl çalıştığını yeniden düşünmek şarttır. Bu bağlamda düşünüldüğünde yeni akıllı teknolojiler kullanılarak akıllı arşivlere geçiş olası görünmektedir.

Bilginin arşivlenmesi ihtiyacı, verilerin çok büyük boyutlara erişmesi, yasal gereklilikler, uyumluluk ve gittikçe artan karmaşık altyapılar nedeniyle hızla artmaktadır. Bilginin yönetimi ve saklanması etkileyen bu gereksinimler neticesinde kurum ve kuruluşlar, hangi bilgilere sahip olduklarını, neyin saklanacağını ve neyin imha edileceğini belirlemeye çalışmaktadır. Dolayısıyla organizasyonların aşağıdaki özellikleri gerçekleştiren akıllı bir arşiv stratejisine ihtiyacı bulunmaktadır:

- Organizasyon, kendi içinde hangi bilgilerin bulunduğunu daha iyi anlamaya çalışmalıdır. Bu; organizasyonun hayati özelliklerini bilmesini, yönetmesini ve güvenmesini kolaylaştırır,
- Organizasyon, uyumluluk çabalarını desteklemek ve riski yönetmeye yardımcı olmak adına önemli bilgilerin uygun şekilde saklanmasını ve korunmasını sağlamaya yardımcı olur,
- Organizasyon, çeşitli kaynaklardan gelen çoğu veri ve içerik türünü işlemek için tasarlanmış daha basit, birbirine bağlı, entegre çözümler kümesi ile genel bir arşivleme stratejisi sağlar,
- Organizasyon, bilgi altyapısının esnekliğini artırırken maliyetleri ve karmaşıklığı azaltır (IBM, 2010, s. 2).

Akıllı bir arşiv; optimize edilmiş ve birleştirilmiş belge sağlama süreçlerine ve keşif ile analitik temelli değerlendirme yoluyla hangi bilgi ve belgelerin arşivleneceği konusunda daha derin bir anlayışa sahiptir. Bunun yanı sıra, ortak sağlama ve sınıflandırma teknolojileri yoluyla veri ve içerik arşivlemeyi birleştirerek karmaşıklığı ve maliyeti azaltır.

Akıllı bir arşiv oluşturma stratejisinin faydaları ise şu şekilde özetlenebilir:

- BİT maliyetlerinin düşmesi: Entegre cihazlar ve çoklu dağıtım seçenekleri sayesinde arşivleme altyapısının basitleştirilmesi ve genel depolama ile güç gereksinimlerinin yanı sıra idari gereksinimlerin azaltılması,
- Performansın artması: Sistemin işlem ve kullanıcı verimliliğinin artırılmasına yardımcı olarak etkin olmayan veya az erişilen içeriğin akıllıca taşınması,
- Tutarlı bir yönetişimin desteklenmesi: Optimize edilmiş bilgi yönetimi yetenekleri ile uyumla ilgili bilgilerin organizasyon ömrü boyunca proaktif olarak yönetilmesi,
- Yasal durumların uyarlanabilmesi: Yetkili personelin olay ile ilgili bilgilere hızlı bir şekilde erişmesini ve verilerin analiz edilmesinin sağlanarak keşif maliyetlerinin düşürülmesi ve müdahale yeteneklerinin artırılması (IBM, 2010, s. 3).

Akıllı arşiv oluşturulmasına yönelik olarak Yong, Liming ve Yongsheng (2015), bir çalışmada Nanjing Arazi ve Kaynak Bürosu'nda akıllı dosya yönetim sisteminin kurulması ve dokümanlara daha hızlı erişim sağlamak üzerine büyük veri ortamında araştırmalar yapmıştır. Araştırmanın amacı, Spring-Struts-Hibernate (SSH)² çerçeve sistemine dayanan, erişim performansını optimize etmeyi ve verimliliğini arttırmayı amaçlayan Java Enterprise Edition (JEE) çerçevesinin kusurlarını açıklamak ve yeni bir model önermektir. Bunun yanında, artan arşiv belgelerinin, web tabanlı arşiv yönetim sistemine erişim verimliliğini azaltmasının nedenleri de incelenmiştir. Araştırmada veri madenciliğinde verileri sınıflandırmak için kullanılan ve büyük verilere daha hızlı erişim sağlayan Karınca Kolonisi Algoritması (Ant Colony Algorithm-ACA) kullanılmıştır. Aynı zamanda, optimize edilmiş algoritmalarının karmaşıklığı ve verileri geri gönderme hızına dair verimlilik de analiz edilmiştir. Sonuçlar, büyük veri ortamında erişim süresinin ve verimliliğin arttığını göstermektedir. Bu çalışma için Nanjing Belediyesi Toprak ve Kaynak Bürosu tarafından sağlanan 150000 veri dosyası kısa bir süreliğine alınmıştır. Ardından bu veriler, sezgisel erişim algoritması, temel ACA algoritması ve uyarlanabilir ACA erişim algoritması ile teste tabi tutulmuştur. Test sonucunda temel ACA erişim algoritmasının büyük veriye erişim hızının, sezgisel algoritmadan daha güçlü olduğu görülmüştür. Uyarlanabilir ACA erişim

² Java dilinde kullanılan çerçeve türleri. Spring, masaüstünden Web'e uygulama geliştirmek için kullanılır. Hibernate, veri katmanına erişmek için kullanılır ve Struts ise web çerçeveleri için kullanılır.

algoritmasının ön verilere erişim hızı ise, ilk iki algoritmaya göre 20 saniyeden daha az olarak görülmüştür. Bunun yanında 30 saniyeden sonra, uyarlanabilir ACA algoritmasının, ilk iki algoritmadan çok daha fazla veriye ulaştığı belirlenmiştir. Bu nedenle, büyük veri havuzuna daha hızlı erişim için kullanılan uyarlanabilir ACA algoritmasının, diğer algoritmalarından daha verimli bir geri kazanıma sahip olduğu anlaşılmıştır.

5. Belge Yöneticisi ve Arşivcinin Yapay Zeka ve Büyük Veri Konusunda Yapması Gerekenler

Bilgi ve iletişim teknolojileri, günümüz belge yöneticileri, arşivciler ve müze ile kütüphane yöneticileri için önem arz eden bir alandır. Fiziksel bir saklama mantığından dijital doğru evrilen bir dönemde, bu teknolojileri kullanabilmek, bilgi profesyonellerine muhakkak fayda sağlayacaktır (Currall ve Moss, 2008). Şüphesiz yeni becerilerin öğrenilmesi gerektiği, kütüphanecilerin ve arşivcilerin sahip oldukları becerilerden daha geniş bir yelpazede profesyonel becerilerin hayata geçirilmesi beklenmektedir.

Dijital malzemelerin uzun süreli korunması hakkında fazla bilgiye sahip olunmakla birlikte bu alandaki becerilerin zaman içinde ve teknolojiler değiştikçe gelişeceği öngörülmektedir. BİT giderek artan bir şekilde yöneticilerin ve arşivcilerin belgelere erişiminde bir uyum sağlar ve denetim kültürünün tarafsız olmayan ve dar profesyonel kaygıların çok ötesine ulaşan sonuçları anlamasına yol açar. Yakın bir gelecekte; kütüphaneci, arşivci ve belge yöneticilerinin kontrolünde olan süreçler, yapay zeka uygulamaları tarafından yönetilecek ve işsiz kalma tehlikesi baş gösterecektir. Bunun için yapay zeka ve büyük veri kavramları iyice anlaşılması gerekmektedir. Yapay zeka ve büyük veri kavramları, hiç şüphesiz belge yöneticilerine yeni bir rol biçecektir. Belge yöneticisi, geleneksel belge yöneticilerinin aksine analitik süreçte bir ortak çalışma grubunda bulunabilir, verilerin konumu hakkında bilgi sağlayabilir ve oluşturma bağlamlarını, yapılarının ve anlamlarının tarihini ve zincir ilişkileri açıklamada görsel analistin veri anlayışını ve çalışmaya güvenini artırabilir. Dolayısıyla, verilere çevrimiçi erişim devam ettikçe, bu verilerin dokümantasyonunu yönetme ihtiyacı da artacaktır (Currall ve Moss, 2008).

Yapay zeka sadece test verileri üzerinden sonuçlara ulaşmak için büyük veri setlerini eğitim kapsamında kullanmayı değil, aynı zamanda arşivcilerin veya belge yöneticilerinin yeni teknolojiler ortaya koyma ve çeşitli hesaplamalarla makinelerden faydalanması için sunduğu bir fırsat olarak görülmelidir (Rolan vd., 2018).

Bir bilgi ve belge yöneticisi, yüksek miktardaki verilerin işlenmesine şu şekilde katkıda bulunabilir:

- Veri temizleme ve hazırlığının azaltılması,
- Depolama maliyetleri ve risklerinin azaltılması,
- İyi meta verilerle verilerin tanımlanmasının iyileştirilmesi,
- Verilere erişimin iyileştirilmesi,
- Çalışan verimliliğinin artırılması,
- Daha iyi, analitik temelli karar alma imkânının yaratılması (Dale, 2015, s. 30).

Bilgi ve belge yöneticisi ayrıca büyük veri bağlamında şu konuları özümseyebilmelidir (Dale, 2015, s. 31):

Hacim: Veri miktarı arttıkça analiz etme ve veriyi yönetme maliyeti o kadar artar. Bilgi ve belge yöneticisi kullanışlı bir bilgi ve belge yönetimi programı ile ayıklama, imha ve tasfiye kriterlerini uygulayarak nispeten daha nitelikli verilerin analiz edilmesine yardımcı olabilir. Örnek olarak, kopya bilgi ve belgelerin ayıklanması ve böylece veri analitiği uzmanlarının daha yüksek kalitede veri ile ilgilenmeleri verilebilir.

Çeşitlilik: Bilgi ve belge yöneticilerinin veri havuzlarının neler olduğu ve bunların envanterlerine sahip olması analiz uzmanlarının işlerini kolaylaştırabilir. Bilgi ve belge yöneticileri yeni teknolojilerin uygulanmasında önemli bir rol üstlenebilir. Diğer teknoloji ekipleriyle işbirliği yapılması, veri oluşturma ve saklama ile ilgili sorunların asgariye indirgenmesi anlamına gelecektir.

Hız: Yüksek hızdaki veri toplama sensörlerinin yaygınlaşması büyük veriye katkıda bulunur ancak yeterli bilgi yönetiminden mahrum kalınması durumunda, bu aynı zamanda bazı risklerin ortaya çıkmasına neden olacaktır. Etkili ve güncel bilgi ve belge yönetimi politikaları ile hızlı veri kontrolsüz veriye dönüşmekten kurtulabilir. Bu politikalar:

- Veri yapısının ve saklama biçiminin verilerin değerini arttırdığından ve koruduğundan emin olmak,
- Verilerin yaşam döngüsü içinde düzenli ve etkin bir biçimde yönetildiğinden emin olmak,
- Veri bilimcisinin veriler hakkında şüphe duymasının önüne geçilmesi, veri manipülasyonu ve sahteciliği gibi kavramları düşünmeden gönül rahatlığı ile iş yapılmasını sağlayacak olan, bilgi ve belge yöneticileridir.

6. Sonuç ve Değerlendirme

Büyük veri ve yapay zeka kavramları son yıllarda gittikçe artan bir şekilde üzerinde çalışılan konular arasında yer almaktadır. Karmaşık birtakım hesaplamalar ve formüller vasıtasıyla yapılan işlemler sonucunda çok büyük miktarlardaki veriler arasından; çeşitli eğilimler, durum analizleri, iyileştirme çalışmaları vb. yapılarak insan hayatına doğrudan etki edecek yenilikler üretilmiştir. Bu teknolojilerin etki ettiği alanlardan birisi de belge yönetimi ve arşiv disiplini. Yapılan çalışmalarda belgelerin içerik, form ve ilişkisel özelliklerinin korunmasının yanında hacim, çeşitlilik ve erişim hızı bağlamında geliştirmelerin olduğu görülmektedir. Bu alandaki çalışmalar, genellikle büyük miktardaki belgelerin seçimi, değerlendirilmesi, ayıklama ve imhasının yapay zeka teknolojileri aracılığıyla yapılmasını üzerine yoğunlaşırken (Rolan vd., 2018) bazı çalışmalar da arşivlenmiş ve tarihi değere sahip olan çok sayıda belgenin içerik analizi yapılarak çevrimiçi ortamında kullanıcıların erişimine sunulmasını içermektedir (Cox vd., 2018). Bunun yanında yakın bir gelecekte tasarlanması muhtemel bir “akıllı arşiv” kavramı, özelleştirilebilir algoritmalar ile desteklenerek bilinen tüm süreçleri farklı bir boyuta taşıyacaktır. Bu projeksiyon ile ilgili olarak bilgisayar bilimleri ile arşiv biliminin ortak bileşenleri üzerine çalışılan ve nistepen yeni bit alan olan Hesaplamalı Arşiv Bilimi (Computational Archival Science-CAS) literatürde kendine yer bulmuştur. Bu yeni kavramda, son zamanlarda oluşan yöntem ve teknolojilerin büyük veri ve yapay zeka teknolojileri etrafında birleştirilmesi ve sonuçlarının arşiv uygulamaları ve yeni analiz biçimleri ile birlikte arşivlerin tarihsel, sosyal, bilimsel ve kültürel bağlamda karşılıklı ilişkileri araştırılmaktadır. Endüstri 4.0’ın sunmuş olduğu büyük veri, yapay zeka vb. teknolojiler ile Hesaplamalı Arşiv Bilimi (HAB); belge ve arşiv yönetimi alanında çalışan uzmanların Elektronik Belge Yönetim Sistemi (EBYS), Yönetim Bilişim Sistemleri (YBS) ve veri bilimi alanlarında yetkin olmaları ve kendilerini geliştirmeleri konusunda fırsat sunduğu bir gerçek olarak karşımıza çıkmaktadır. Bu yeni teknolojilerin benimsemesi, bilgi merkezlerinin yönetim süreçlerinde verimlilik ve etkililiğini geliştirme anlamında farklı bir vizyon getirecektir. Belgelerin -özellikle elektronik belgelerin- oluşturulması, belirli bir döngüye dahil edilmesi, ayıklama işleminin yapılması ve ardından saklanmasına veya imhasına karar verilmesi süreçlerini iyileştirecek veya yeniden yorumlanmasını sağlayacak olan bu teknolojilerin bilgi ve belge yönetimi süreçlerinde uygulanması, insan kaynaklı hataların oluşmasını engelleyecek ve belge yönetimi ve arşiv süreçlerinin daha etkin bir işlerlik kazanmasına yardımcı olacaktır.

Kaynakça

- Aktan, E. (2018). Büyük veri: Uygulama alanları, analitiği ve güvenlik boyutu. *Bilgi Yönetimi Dergisi*, 1(1), 1-22.
- Arlitsch, K. ve Newell, B. (2017). Thriving in the age of accelerations: A brief look at the societal effects of artificial intelligence and the opportunities for libraries. *Journal of Library Administration*, 57(7), 789-798.
- Boyd, D. ve Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- Cox, M. ve Ellsworth, D. (1997). Application-Controlled Demand Paging for Out-of-core Visualization. Erişim Adresi: <https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf>

- Cox, R., Shah, S., Frederick, W., Nelson, T., Thomas, W., Jansen, G., Dibert, N., Kurtz, M. ve Marciano, R. (2018). A case study in creating transparency in using cultural big data: The legacy of slavery project. *IEEE Big Data 2018: 3rd CAS Workshop*. Seattle: Washington. Erişim Adresi: https://dcicblog.umd.edu/cas/wp-content/uploads/sites/13/2018/12/12.Cox_-2.pdf
- Currall, J. ve Moss, M. (2008). We are archivists, but are we OK?. *Records Management Journal*, 18(1), 69-91.
- Çiçek, N. (2009). *Modern Belgelerin Diplomatığı*. İstanbul: Derlem Yayınları.
- Dale, K. L. (2015). RIM's role in harnessing the power of big data. *Information Management*, 2015, 29-32
- Dollar, C. M. (1993). Archivists and Record Managers in the Information Age. *Archivaria*, 36, 37-52.
- Emery, P. (2003). Document and records management: Understanding the differences and embracing integration. *ZyLAB Technologies White Paper*, September 2003, 1-8.
- Frey, C. B. ve Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?" *Technological Forecasting and Social Change*, 114, 254-280.
- Government Office for Science. (2014). The Internet of Things: Making the Most of the Second Digital Revolution. Erişim Adresi: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/409774/14-1230_internet-of-things-review.pdf
- Gramatica, R. ve Pickering, R. (2017). Start-up story yewno: an ai-driven path to a knowledge-based future. *Insights the UKSG Journal*, 30(2), 107-111.
- Grewal, S. (2014). A critical conceptual analysis of definitions of artificial intelligence as applicable to computer engineering. *IOSR Journal of Computer Engineering*, 6(2), 9-13.
- Gutierrez, D. (2017). Above the Trend Line-Your Industry Rumor Central for 4/10/2017. Erişim Adresi: <https://insidebigdata.com/2017/04/10/trend-line-industry-rumor-central-4102017/>
- Harari, Y. N. (2017). *The Rise of the Useless Class*. Erişim Adresi: [http:// ideas.ted.com/the-rise-of-the-useless-class/](http://ideas.ted.com/the-rise-of-the-useless-class/)
- IBM. (2010). The Smart Archive Strategy From IBM. IBM Software Industry Solutions.
- Janssen, M. ve Hoven, J. (2015). Big and open linked data (bold) in government: A challenge to transparency and privacy. *Government Information Quarterly*, 32, 363-368.
- Jie, S., Linqian, D., S. Jing, S. ve Huan, H. (2016). Management of University Endowment Fund Archives Under the Background of Big Data. Erişim Adresi: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7976552>
- Kandur, H. (2006). Elektronik Belge Yönetimi Sistem Kriterleri Referans Modeli v.2.0. Ankara: Devlet Arşivleri Genel Müdürlüğü, Cilt No: 29.
- Lee, M., Zhang, Y., Chen, S., Spencer, E., Cruz, J. D., Hong, H. ve Marciano, R. (2017). Heuristics for Assessing Computational Archival Science (CAS) Research - The Case of the Human Face of Big Data Project. Erişim Adresi: <https://ieeexplore.ieee.org/document/8258179>
- Levi, A. S. (2013). Humanities big data. 2013 IEEE International Conference on Big Data. Erişim Adresi: <https://pdfs.semanticscholar.org/b302/6e202008a321fa2e15b7668331be9f3cc8cf.pdf>
- Mahon, D. (2015). Focus on the active. *Business Intelligence Journal*, 19(4), 40-47.
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M. ve Conrad, M. (2018). Archival records and training in the age of big data. *Advances in Librarianship*, 44(B), 179-199.
- McDonald, J. ve Léveillé, J. V. (2014). Whither the retention schedule in the era of big data and open data?. *Records Management Journal*, 24(2), 99-121.

- Moses, R. P. (2005). *A Glossary of Archival and Records Terminology*. Chicago: The Society of American Archivists.
- Newell, B. C. (2014). Technopolicing, surveillance, and citizen oversight: a neorepublican theory of liberty and information control. *Government Information Quarterly*, 31(3), 421-431.
- Prescott, A. (2013). Bibliographic Records as Humanities Big Data. 55-58. Erişim Adresi: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6691670>
- Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoukova, T. ve Stuart, K. (2018). More human than human? artificial intelligence in the archive, *Archives and Manuscripts*, 47(2), 179-203.
- Sakya, K. T. (2016). Big Data: Understanding Big Data. Erişim Adresi: <https://arxiv.org/abs/1601.04602>
- Shallcross, M. (2016). Appraising Digital Archives with Archivematica. 3272-3276. Erişim Adresi: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7840985>
- Shepherd, E. J. ve Yeo, G. (2003). *Managing Records: A Handbook of Principles and Practice*. London: Facet Publishing.
- Thompson, N., Ravindran, R., ve Nicosia, S. (2015). Government data does not mean data governance: lessons learned from a public sector application audit. *Government Information Quarterly*, 32(3), 316-322.
- Uribe, L. M. (2018). Digital Archives as Big Data. *Mathematical Population Studies*, 26(2), 69-79.
- Ünal, M. A. ve Özdemirci, F. (2017). EBYS (e-BEYAS) ve e-arşiv sistemlerinde/uygulamalarında yapay zeka yaklaşımı. F. Özdemirci ve Z. Akdoğan (Ed.) *Bilgi Sistemleri ve Bilişim Yönetimi Bildiriler Kitabı* içinde (ss.57-63). Ankara: BİL-BEM.
- Yong, M., Liming, L. ve Yongsheng, Q. (2015). Improvement of Big Data Retrieval Algorithm in the Intelligent Archives Management. Erişim Adresi: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7494245>