

Kurumsal Akademik Arşivlerde Üst Veri Kalitesi Üzerine Değerlendirme *Evaluation on Metadata Quality in Institutional Repositories*

Sönmez Çelik*, **Sami Çuhadar****

Sayfa: 191-203

Öz

Veri, ölçülebilir veya gözlemlenebilir gerçekleri temsil eder. Veri, olayları, nesnelere, ürünleri veya onların özelliklerini tanımlamak için kullanılır. Üst veri ise, bir çalışmayla ilgili olarak; çalışmanın başlığı, özeti, anahtar kelimeleri, yazar bilgileri, atıfları gibi verileri ifade eder. Ayrıca, bir bilgi kaynağının türü, dili, kullanım hakkı, konu başlıkları, kalıcı linkleri, otorite dizini gibi kaynağı kullanmayı ve yönetmeyi kolaylaştıran bilgiler de içerir. Üst veriler, çalışmanın doğru şekilde tanıtılması, keşfedilmesi ve erişilebilir hale getirilmesi için kritik öneme sahiptir. Bir akademik kurumda üretilen yayınlara ve bunlarla ilgili kaynaklara ilişkin üst veriler, araştırmalara kolay erişim, veri bütünlüğü, verilerin analizi, keşfi, yapılan bir araştırmanın sonraki araştırmalara temel oluşturması, atıf alma, kurumsal itibar, görünürlük, verilerin korunması ile uzun süreli saklanması gibi sayısız fayda sağlar.

Bu bağlamda üst verilerde kalite çok önemlidir. Kaliteli üst veri; standartlara uygun, doğru, güvenilir, eksiksiz, tutarlı, amaca uygun özelliklere sahip olmalıdır. Bu özellikleriyle üst veri, doğru analizlerin yapılmasına, etkili kararların alınmasına, eğilimlerin belirlenmesine ve sağlam bilgi temelli stratejilerin oluşturulmasına katkı sağlar. Bu da kaynakların doğru kullanımı, alınan kararların isabetli sonuçlar vermesi ve dolayısıyla kurumsal akademik arşiv ile ilgili hizmet kalitesinin artmasını sağlar. Üst veri kalitesi ve erişilebilirliği, araştırmaların ilerlemesini, bilimsel etkileşimin artmasını ve kurumun akademik başarısını destekler.

Raporlamanın doğru yapılabilmesi üst verilerin kalitesi ile doğrudan ilişkilidir. Yöneticilerin ve araştırmacıların üst verilerden anlamlı sonuçlar çıkararak kuruma ya da yapılan bir çalışmaya yön gösterebilmesi verinin kalitesi ve büyüklüğü ile mümkündür. Ayrıca kaliteli üst verinin uluslararası protokol ve sistemlere uygun olarak oluşturulmasında, paylaşılmasında ve görünür hale getirilmesinde kütüphane ve kütüphanecilerin rolü önemlidir.

Betimleme ve örnekleme yöntemlerinin kullanıldığı bu çalışmada, üst verinin temsil ettiği kaynağın keşfedilebilirliği, etki faktörü, standartlara uygunluğu, bilimsel iletişime katkısı gibi konulardan hareketle, veri kalitesi üzerinde durulmuştur. Kaliteli üst verilerle, akademik bir kurumda farklı metriklerle ileri düzey raporlama yapılabilir ve bu üst veriler kurumsal stratejik kararlara destek verebilir. Öte yandan üst veri yapısı araştırma ve performans değerlendirme bilgi sisteminin de ana omurgasını oluşturur. Akademik kurumlar açısından oldukça önemli olan performans değerlendirme bu çalışmada kaliteli üst veri yaklaşımıyla bütüncül olarak ele alınmaya çalışılmıştır. Kaliteli üst veri örnekleri aracılığıyla akademik kurumlarda ne tür değerlendirmeler yapılabileceğine dair genel bir bakış açısı verilmeye çalışılmıştır.

Çalışma çerçevesinde Türkiye'deki kurumsal akademik arşivlerdeki üst verilerin kalitesi, bu verilerin raporlamada ve veri madenciliğinde kullanılabilirliği değerlendirilmiştir. Öte yandan bu üst verilerden ne gibi çalışmalar yapılabileceği konusunda öngörüle bulunulmuştur. Türkiye'deki kurumsal akademik arşiv çalışmaları ile ilgili faaliyetler dünyadaki gelişmeler ile paralel olarak ve çoğu zaman uluslararası projelerin paydaşları olarak gerçekleştirilmiştir. Anadolu Üniversite Kütüphaneleri Konsorsiyumu, Yükseköğretim Kurulu Başkanlığı ile Ulusal Akademik Ağ ve Bilgi Merkezinin girişimleri ile yapılan çalışmalar sonucunda 2023 yılı itibarıyla Türkiye'de 157 adet kurumsal akademik arşiv bulunmaktadır. Çalışma kapsamında Türkiye'deki kurumsal akademik arşivden rastgele seçilen 17 tanesine ait 850 kayıt

* Daire Başkanı, İbn Haldun Üniversitesi Kütüphanesi, sonmez.celik@ihu.edu.tr, ORCID ID: 0000-0002-3973-4664

** Direktör, İstanbul Bilgi Üniversitesi, Kütüphane ve e-Kaynaklar, sami.cuhadar@bilgi.edu.tr, ORCID ID: 0000-0003-0602-7928

incelenmiştir. Kurumsal akademik arşivlerdeki üst verilerin kaliteli, orta düzey ve zayıf olmak üzere üç grupta değerlendirilebileceği sonucuna varılmıştır.

Yapılan çalışma sonucunda kaliteli verilerin doğru raporlama ve öngörü için kurumlara yardımcı olabileceği, orta kalitedeki verilerin bir kısım raporlamada kullanılmasının mümkün olmadığı ve zayıf düzeydeki verilerin raporlama ve temsil etme bakımından kurumlar için olumsuz sonuçlarının olabileceği kanaatine varılmıştır. Kendi alanında özgün olan çalışma, bu alanda çalışan araştırmacılar ve kurumlar için faydalı bilgiler içermektedir.

Anahtar Sözcükler: Üst veri; veri yönetimi; kurumsal arşivler; üst veri kalitesi; raporlama

Abstract

Data represents measurable or observable facts. Data is used to describe events, objects, products or their characteristics. Metadata refers to data related to a work, such as the title, abstract, keywords, author information, citations. It also includes information that makes it easier to use and manage an information source, such as its type, language, right of use, subject headings, permalinks, authority index. Metadata is critical for proper promotion, discovery and accessibility of the work. Metadata on publications and related resources produced in an academic institution provide numerous benefits such as easy access to research, data integrity, data analysis, discovery, the basis of a research for further research, citation, institutional reputation, visibility, data protection and long-term storage.

In this context, the quality of metadata is very important. High-quality metadata should be in compliance with standards, accurate, reliable, complete, consistent and fit for purpose. With these features, metadata contributes to accurate analysis, effective decision-making, identification of trends and the creation of sound information-based strategies. This, in turn, leads to the correct use of resources, accurate results of the decisions taken, and thus increases the quality of service related to the institutional repository. Metadata quality and accessibility supports the advancement of research, increased scholarly interaction and the academic success of the institution.

Accurate reporting is directly related to the quality of metadata. The ability of administrators and researchers to draw meaningful conclusions from metadata to guide an organization or a study is possible with the quality and size of the data. In addition, the role of libraries and librarians is essential in creating, sharing and making visible quality metadata in accordance with international protocols and systems.

This study, which uses the descriptive method, focuses on data quality based on issues represented by metadata such as discoverability of the source, impact factor, compliance with standards and contribution to scientific communication. With high-quality metadata, advanced reporting with different metrics can be done in an academic institution and these metadata can support institutional strategic decisions. On the other hand, the metadata structure also forms the basis of the research and performance evaluation information system. In this study, performance evaluation, which is very important for academic institutions, is tried to be handled holistically with a quality metadata approach. Through examples of quality metadata, it is tried to give a general perspective on what kind of evaluations can be made in academic institutions.

Within the framework of the study, the quality of metadata in institutional repositories in Turkey and the usability of these data in reporting and data mining were evaluated. On the other hand, predictions have been made on what kind of studies can be carried out from these metadata. The activities related to institutional repositories in Turkey have been carried out in parallel with the developments in the world and often as stakeholders of international projects. As a result of the studies carried out with the initiatives of the Anatolian University Libraries Consortium, the Presidency of the Council of Higher Education and the National Academic Network and Information Center, there are more than 140 institutional repositories in Turkey as of 2023. Within the scope of the study, randomly selected records from institutional repositories in Turkey were analyzed. It was concluded that metadata in institutional repositories can be evaluated in three groups: quality, medium level and weak.

As a result of the study, it was concluded that high quality data can help institutions for accurate reporting and forecasting, medium quality data cannot be used in some reporting, and poor quality data may have

negative consequences for institutions in terms of reporting and representation. The study, which is unique in its field, contains useful information for researchers and institutions working in this field.

Keywords: Metadata; data management; institutional repositories; metadata quality; reporting

Giriş

Kurumsal Akademik Arşiv (KAA), bir kurum tarafından doğrudan ve dolaylı olarak yayınlanan; makale, bildiri, kitap, tez, rapor, araştırma verisi gibi tüm araştırma çıktılarını uluslararası standartlarda arşivleyen, yönetilmesini sağlayan, akademik performansın ölçülmesini sağlayan, açık erişime altyapı sağlayan, yayınları uzun süreli saklayan ve içeriği telif haklarına uygun erişim sunan sistemdir. Akademik performans sisteminin temeli olan KAA'lar, açık erişime altyapı sağlar. Öte yandan yayınların görünürlüklerini ve erişilebilirliğini artırır. Bu yolla araştırmanın etki faktörünün yükseltilmesine katkı sağlar.

Bilimsel araştırma verilerini yaymak yükseköğretim kurumlarının görevlerindedir. Öğretim elemanlarının ve araştırmacıların bilimsel yayınlarının üniversitelerde ve Yükseköğretim Kurulunda derlenmesi ve arşivlenmesi 2547 sayılı Yükseköğretim Kanunu'nun emridir. Kanunun 42/c maddesi ile yayınların derlenmesi, 42/d maddesi ile arşivlenmesi ve 12/c maddesi ile bilim verilerinin söz, yazı ve diğer araçlarla yayılması istenmektedir (Yükseköğretim Kanunu, 1981).

Akademik çıktıların uzun süreli saklandığı KAA'lar, kuruma ait yayınların uluslararası standartlarda arşivlenmesini sağlar ve bağlı olduğu kurumu veri sağlayıcı kurum kategorisine taşır. Açık erişim; bilginin hızlı bir şekilde daha çok kişiye engelsiz ulaşmasını, yayınlara kolay erişim ayrıcalığının sağlanmasını, daha önce aynı konuda yapılan çalışmaların belirlenmesini, yayınlara daha çok atıf yapılmasının sağlanmasını, araştırmacıların ve kurumların saygınlıklarının artmasını sağlar. Öte yandan açık erişim, finansal ve yasal engelleri ortadan kaldırarak bilimsel bilgiyi özgürleştirir, adil kullanım hakkı sağlar.

Hızla yükselen maliyetler, bilimsel iletişimin geniş platformlara taşınması, araştırma etkisinin artışına ve elektronik ortamdaki mirasın korunmasına duyulan ihtiyaç, açık erişimin ortaya çıkış nedenlerini oluşturmaktadır. KAA'lar finansal ve yasal engellerin kaldırılmasına aracılık ederek bilimsel bilgiyi özgürleştirmeye, adil kullanım hakkı sağlamaya, araştırma sonucunun tanınırlığını ve görünürlüğünü artırarak bilimsel etkisini yükseltmeye, atıf oranlarının artışına etkisi ve araştırmacıların literatüre daha geniş alandan bakmasını sağlamaya aracılık etmesi sistemin gerekliliğini ortaya koymaktadır (Tang, Yu, ve Bever, 2017).

Türkiye'de kurumsal akademik arşiv yazılımı olarak DSpace yoğun olarak kullanılmaktadır (ANKOS, 2023). Bu durum ülke olarak kurumsal akademik arşivler konusunda birlikte hareket etme ve standart yapıda üst veri oluşturma fırsatını da beraberinde getirmektedir. Uluslararası düzeyde kabul görmüş, standartlara uygun bu sistem ile harmanlama sistemlerine 12 ayrı yapıda veri sağlanabilmektedir.

TÜBİTAK ULAKBİM öncülüğünde Türkiye Akademik Arşivi adlı Harman sistemi KAA'ların içeriğini merkezi olarak harmanlaması önemli bir kilometre taşıdır. Harman sistemi kurumları sistemde yer almak, yayın sayılarını artırmak, standartlara göre kayıt girme gibi konularda teşvik etmektedir. Ancak bu sistemin kabul şartları arasında arşivlerin içeriğine yönelik bir bilgilendirme bulunmamaktadır. Kabul koşulları arasında en azından KAA tanımına atıfta yapılmalı ve KAA'ların tanıma uygun arşivlerin harmanlanacağı belirtilmelidir (Harman, 2023).

Türkiye'deki tüm akademik kurumlarda KAA'ların kurulması, ulusal ve uluslararası harmanlama sistemlerinde harmanlanması, Türkiye adresli yayınların uluslararası sistemlerde erişilebilir hale gelmesi için çok önemlidir. KAA'lar; bölüm, fakülte, üniversite genelinde; ulusal harmanlama sistemi ise ülke genelinde üretilen akademik çıktıların listesinin oluşması, bunlara tek ara yüzden erişim sağlanmasına imkân sağlamaktadır. KAA'lar;

- akademik performans değerlendirme sistemlerinin temel yapısının oluşturulmasına,
- kurum yayınlarının hızlı bir şekilde araştırmacılara engelsiz ulaştırılmasına,
- yeni çalışmalar için kurum yayınlarına kolay erişim ayrıcalığının sağlanmasına,
- daha önce aynı konuda neler yapıldığının kolayca belirlenmesine,
- yayınlara daha çok atıf yapılmasının sağlanmasına,

- kültürel mirasın gelecek kuşaklara aktarılmasına,
- kurumun sıralama (ranking) sistemlerinde yukarılara çıkmasına,
- araştırmacıların ve kurumun saygınlığının artmasına

katkı sağlar.

KAA'lar, öğretim üyelerinin, üniversite personelinin ve öğrencilerinin araştırma ve bilimsel çıktılarını toplar. KAA'lardaki üst veriler hem insanlar hem de makineler tarafından oluşturulan üst verilerden oluşur. KAA'daki üst veriler; yazarlar, KAA sorumluları ya da belli kaynaklardan profesyoneller tarafından toplu olarak derlenip arşivlenir. Farklı yollarla arşivlenen yayınlarda üst veri hataları ya da eksikleri ortaya çıkmaktadır. Hataların türleri arşivler arasında farklılık gösterdiği gibi aynı KAA içinde de farklılık göstermektedir. Hataların büyük çoğunluğu, veri girişlerinin profesyonel kişilerce yapılmaması ve bu konudaki rehberlerin eksikliğinden kaynaklanmaktadır. Üst veri hatalarının nedenlerinin kapsamlı olarak belirlenmesi, KAA'lardaki mevcut hataların düzeltilmesi ve gelecekte bu hataların azaltılmasına büyük katkı sağladığı söylenebilir.

Bu çalışmada, KAA'lara üst veri girişlerinden hatalarının azaltılması, eski kayıtlardaki üst veri hatalarının geriye doğru düzeltilmesine yönelik olarak iyileştirmenin önemi üzerinde durmuştur. Kaliteli üst veri oluşturmak için bu konuda ilgili personelin eğitiminin yanında OpenAIRE'in veri zenginleştirme gibi ilgili araçlarının da kullanımının gerekli olduğu görüşündedir. Öte yandan üst veri kalitesine yönelik teorik konuların ele alınması, kaliteli üst verilerle yapılacak analizler aracılığı ile anlamlı projeksiyonlar yapılması, araştırmalarda anlamlı sonuçlara ulaşılması yolu ile bilime ve topluma katkı sunulması amaçlanmaktadır.

Literatür Taraması

KAA'larda üst veri kalitesi ve değerlendirmesi konusunda yabancı literatürde az sayıda çalışma (Ochoa ve Duval 2009; Piazzini, 2022; Jeannette ve Stokes, 2019; Kurtz, 2010; Sandy ve Dykas, 2016; Palavitsinis, Manouselis ve Sánchez-Alonso, 2017; Osman, Yanti Idaya ve Abrizah, 2023; Yuji Park ve Tosaka, 2010; Park, 2009; Palavitsinis, Manouselis ve Sanchez-Alonso, 2014) vardır. Türkiye'deki KAA'lardaki üst veri kalitesi üzerine yapılmış doğrudan bir çalışma tespit edilememiştir. Türkiye'de KAA'larla ilgili çalışmalar daha çok KAA'ların önemi, yaygınlaştırılması ve teknik açıdan ilgililerin bilgilendirilmesi konularındadır. KAA'larla ilgili çalışmalar yeterli olmamakla birlikte KAA'lardaki üst veri kalitesi ve bu verilerle yapılacak analizler üzerinde çalışma olmaması büyük bir boşluktur. Çünkü KAA'lardaki yayın sayıları her geçen gün hızla artmasına rağmen, üst veri kalitesi konusundaki çalışmanın yeterli olmadığı söylenebilir. Bu çalışmada, KAA'lardaki üst verilerde bulunan eksik ve hataları tespit etmeye çalışarak bu konudaki boşluğun giderilmesine katkıda bulunmanın yanında önemini vurgulamak amaçlanmıştır.

Çalışmanın Kapsam, Sınır, Evren ve Yöntemi

Bu çalışmanın kapsamı, konu ile ilgili literatür araştırması yapmak ve Türkiye'deki üniversitelerde bulunan kurumsal akademik arşivleri belirlemek, arşivlerde bulunan üst verileri incelemek ve bu veriler üzerine değerlendirmelerde bulunmaktır.

KAA üst veri kalitesini hakkında temel bilgiler verilmesi, Türkiye'de 157 üniversitede bulunan ve belirli standartları sağlayan 101 KAA'nın %10.82'sini oluşturan 17 KAA'daki belirli sayıda kaydın üst verisinin kalite açısından incelenmesi ve değerlendirilmesi çalışmanın sınırlılığını oluşturmaktadır. Türkiye'de bulunan 208 yükseköğretim kurumu araştırmanın örneklem evrenini oluşturmaktadır.

Bu çalışmada, betimleme ve örnekleme yöntemi kullanılmıştır. Betimleme yöntemi, olgular arasındaki uygunluk, benzerlikleri, daha önceki olaylar ve koşullar arasındaki etkileşimi dikkate alarak açıklamayı hedefler (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz ve Demirel, 2022, 7. s.). Betimleme yöntemi, "bu nedir" sorusuna yanıt aramaya yöneliktir. Örnekleme ise "çalışma evreni için araştırma evreninden, onu temsil edecek sayıda birey veya objelerin oluşturduğu bir parçanın seçilmesidir" (Aziz, 2022, 249. s.). Çalışmalarda araştırılan hedef kitlenin tamamını incelemek mümkün olmayabilir. Bu durumlarda hedef kitleyi temsil edebilecek belirli bir parça ele alınarak araştırılan hedef kitlenin özellikleri belirlenmeye çalışılabilir ve bir sonuca varılabilir (Seyidoğlu, 2009, 41. s.). Bilimsel çalışmalarda

kullanılan bu yöntem örnekleme yöntemi adıyla ifade edilir. Bu yöntemler ile elde edilen veriler değerlendirilmiş, hata payının en az olmasına özen gösterilmiştir.

Veri ve Veri Yönetimi

Veri yönetimi, yaşam döngüsü boyunca veri varlıklarının değerini sağlayan, kontrol eden, koruyan ve artıran planların, politikaların, programların ve uygulamaların geliştirilmesi, yürütülmesi ve denetlenmesidir. Veri yönetimi, veri varlıklarının üzerinde yetki ve kontrolünün planlanması yanı sıra izleme ve uygulama sağlanmasını da kapsar. Veri yönetiminin detayları şu şekilde tarif edilebilir (DAMA International, 2017):

Veri Mimarisi: Bir ihtiyaca göre oluşturulan sistemde veri ihtiyaçlarının belirlenmesi ve bu ihtiyaçları karşılayacak ana planların tasarlanması ve sürdürülmesi.

Veri Modelleme ve Tasarımı: Veri gereksinimlerinin belirlenmesi, analiz edilmesi ve kapsamının belirlenmesi ile bu veri gereksinimlerinin bir veri modelinde temsil edilmesi.

Veri Depolama ve İşlemleri: Arşivlenen verilerin kalitesini en üst düzeye çıkaracak şekilde tasarlanması, uygulanması ve desteklenmesi.

Veri Güvenliği: Verinin uygun şekilde kimlik doğrulamasını, yetkilendirmesini, erişimini ve denetlenmesini sağlamak için güvenlik politikalarının ve prosedürlerinin planlanması, geliştirilmesi ve yürütülmesi.

Veri Uyumu ve Birlikte Çalışabilirlik: Verilerin arşivler, uygulamalar ve kuruluşlar içinde ve arasında paylaşılması ve harmanlanması.

Belge ve İçerik Yönetimi: İlişkisel veri tabanlarının dışında arşivlenen verilerin bulunması, depolanması, erişimi ve kullanımının kontrol edilmesi.

Referans ve Ana Veri Yönetimi: Kurumsal hedeflerin karşılanması, risklerin azaltılması, daha yüksek kalite sağlanması ve veri uyum maliyetlerinin azaltılması için paylaşılan verilerin yönetilmesi.

Veri Arşivi ve İş Zekâsı: Karar destek verilerinin sağlanması ve raporlanması, sorgulanması ve analizle ilgilenenlerin desteklenmesi için planlanması, uygulanması ve kontrol süreçleri.

Üst Veri Yönetimi: Yüksek kaliteli, uyumlu üst verilere erişimi mümkün kılan planlama, uygulama ve kontrol faaliyetleri süreci.

Veri Kalitesi Yönetimi: Verilerin kullanıma uygun olmasını ve araştırmacıların ihtiyaçlarını karşılamasını sağlamak için verilere kalite yönetimi tekniklerini uygulayan planlama, uygulama ve kontrol faaliyetleri süreci.

Kurumsal Akademik Arşivlerde Üst Veri

Üst veri, bilgi kaynaklarının derlenmesi, kimliklendirilmesi, kullanımı ve yönetimi ile ilgili yapılandırılmış bilgiler topluluğudur. Kurumsal akademik arşivler üst verilerini en azından Open Archives Initiative (OAI) standartlarına uygun olarak Dublin Core (DC) biçiminde sağlamalıdır. DC biçimi en çok kullanılan üst veri biçimlerinden biridir ve 1995 yılında Online Computer Library Center (OCLC) sponsorluğundaki bir çalışmada oluşturulmuştur. DC bir doküman tanımlama standardı olarak oluşturulmuştur ve bilgi kaynaklarına erişim sağlamak üzere 15 temel üst veri ögesi ve ek alanlardan oluşan bir kümedir. Tüm alanlar isteğe bağlı ve tekrarlanabilir özelliğe sahiptir. Temel ve ek DC alanları aşağıda verilmiştir (Kurtz, 2010):

- | | |
|-----------------------------|------------------------|
| 1) Başlık | 9) Biçim |
| 2) Yazar / Oluşturan | 10) Kaynak Tanımlayıcı |
| 3) Konu / Anahtar Kelimeler | 11) Kaynak |
| 4) Tanım / Açıklama | 12) Dil |
| 5) Yayıncı | 13) İlgililik |
| 6) Katkı sağlayan | 14) Kapsam |
| 7) Tarih | 15) Haklar |
| 8) Yayın Türü | |

Ek DC alanları:

- 1) Hedef Kitle
- 2) Kayıt Bilgisi
- 3) Hak Sahibi

DC, alan içeriklerinin daraltılması ve özelleştirilmesine imkân sağlar. Alanların sınırlandırılması gerekli değildir; ancak yerel kullanımlar için standart olmayan alanların kullanılmasına da izin verir.

Standart ve kaliteli üst veri akademik içerikli sistemlerin konuşabilirliği için çok önemlidir. Üst veri, bilgilerin nitel olarak doğru kullanımına işaret eder. KAA sistemlerinde standart DC alanlarının kullanılmasıyla OAI uyumlu veri sağlayıcı ile servis sağlayıcı arasında üst veri iletişimi sağlanabilmektedir. KAA, her türlü bilimsel çıktıyı temsil eden elektronik kaynağı içeren donanım, yazılım, veri ve prosedürden oluşan bir sistemdir. Bu nedenle üst verileri detaylı ve kurallara uygun olarak oluşturmak gerekir.

Kurumsal akademik arşivlere veri girişinde standart biçim ve kalite için DRIVER Rehberi 2.0'da belirtilen usullerin uygulanması gerekir. Böylece en azından üst verilerdeki asgari gereksinimlerin karşılanması sağlanmış olur (DRIVER, 2008).

KAA Sistemlerine Veri Giriş Standartları

DRIVER Rehberi KAA'lara veri girişini detaylı bir şekilde tarif eder; ayrıca KAA veri tabanlarının taşınması gereken teknik özellikler ve alt yapısı için de önemli bir kaynak olma özelliğini taşır (DRIVER, 2008 ve Çelik, 2016). DRIVER Rehberi 2.0'ın Türkçeye çevrili 2014 yılında yapılmıştır ve elektronik ortamda yayınlanmıştır (DRIVER, 2008).

DRIVER Rehberi 2.0 esas alınarak, "Kurumsal Akademik Arşiv Sistemine Kayıt Giriş Rehberi" adıyla bir rehber yayımlanmıştır (Çelik, 2016). Bu rehber, Türkiye'de yaygın olarak kullanılan DSpace yazılımı örneği üzerinden hazırlanmıştır. Rehberin amacı, Türkiye'deki kurumsal akademik arşiv sistemlerine kayıt girişi yapanlara yardımcı olmak ve kayıtların standartlara uygun olarak oluşturulmasına katkı sağlamaktır. KAA'lardaki kayıtların nitelikli ve standartlara uygun olması, araştırmacıların işini kolaylaştırması, harmanlama sistemlerinin içeriğinin belli bir düzende olması ve uluslararası iş birlikleri bakımından gereklidir.

Rehberde, KAA'larda zorunlu, isteğe bağlı ve önerilen seçenekleriyle yoğun olarak kullanılan alanların tanımı, kullanım talimatı, kullanımına ilişkin örnekler, kullanım seçeneği, DC etiketi ve varsa alanla ilgili uyarılar belirtilmiştir.

Kurumsal Akademik Arşivlerde Üst Veri Kalitesi

Başta DSpace olmak üzere KAA yazılımlarının amacı, bir akademik kurumda akademik çıktıları belli standartlarda arşivlemek ve telifine uygun olarak açık erişime sunmak, bilimsel literatüre erişimi kolaylaştırmak, bilgiye erişimde fırsat eşitliği yaratmak, bilginin yeniden kullanımına ortam sağlamak, kurumların ve araştırmacıların görünürlüğünü sağlamak, akademik performans ölçümüne temel oluşturmak ve içeriklerinin arama motorlarıncaya harmanlanabilir olmasını sağlamaktır. KAA yazılımlarının bağlı bulunduğu kurumları veri sağlayıcı pozisyonuna getirmeleri, İnternette görünürlüklerini artırmaları, araştırmaların daha çok atıf almasını sağlayarak etki değerini artırmaları, KAA'ların hızla yaygınlaşmasını ve doğal olarak içeriklerinin hızla artmasını sağladığı söylenebilir. Ancak KAA'lardaki mevcut verilerin kalitesinin aynı oranda arttığı söylenemez. Özellikle eski verilerin farklı platformlardan derlenip düzenlenmeden aktarılması, orijinal kaynaktaki veri kalitesi ile KAA'lara yansımaktadır. Bazı kaynaklardaki eski verilerin yeterli bilgiye sahip olmadan toplu aktarımı içeriğin hızla artmasını sağlarken veri kalitesinin ilgili kaynaktaki kadar olmasının ötesine geçemedi. Bu çalışma çerçevesinde yapılan analizlerde bazı KAA'lardaki veriler zorunlu veri alanlarını bile karşılamayacak yapıda gerçekleştiği gözlemlendi.

Kütüphane katalog kayıtlarında olduğu gibi KAA'lardaki kaliteli üst veri için profesyonel veri uzmanlarının sorumluluğunda iş akış sürecinin yürütülmesinin bir zorunluluk olduğu kanaatindeyiz. 2006 yılından itibaren Türkiye'de KAA çalışmaları yapılmakta ve Eylül 2023 itibarıyla ANKOS Rehberi kayıtlarına göre üniversitelerimizde 157, diğer kurumlarda da 5 olmak üzere toplam 162 adet KAA mevcuttur (ANKOS, 2023). Bu arşivlerden 96 adedi OAI'de, 138 adedi ULAKBİM Harman'da,

113 adedi OpenDOAR'da, 89 adedi OpenAIRE'de, 104 adedi ROAR'da, 86 adedi ROARMAP'de, 93 adedi BASE'de ve 38 adedi CORE'da kayıtlıdır (ANKOS, 2023).

Türkiye'deki KAA'ların bir kısmı yeterli bilgi birikimi ve deneyime sahip olunmadan kurulmuş ve yönetilmeye çalışılmaktadır. Bu KAA'lardaki veriler, kullanılan yazılımın varsayılan veri giriş formları ile yapılmakta ve kalitesi, sağlıklı analiz yapmak için yeterli değildir. Nitelikli arşivlerde yayın türüne göre tasarlanmış ve zenginleştirilmiş veri alanlarından oluşan formlar kullanılmakta ve sınırlı da olsa eğitim almış personel tarafından yönetilmektedir. Buna rağmen kaliteli üst veriye sahip KAA sayısı beklenen düzeyde değildir.

KAA'ların önemini kavrayamamış bazı kurumlarda veri girişinin sayı odaklı olarak yürütüldüğü söylenebilir. Bu da verilerin, veri giriş formlarındaki zorunlu alanların girilmesi ile sınırlı kalmaktadır. Orta düzeyde üst veri kalitesine sahip grupta yer alan KAA'ların verileri de genelde ya profesyonel personelin sayıca yetersiz kalması ya da geriye dönük verilerin, düzenlenmeden, zenginleştirilmeden bulunduğu kaynaktaki yapısıyla toplu aktarılmasından kaynaklıdır. İyi sayılabilecek düzeyde üst veri kalitesine sahip KAA'ların profesyonel kütüphanecilerin eseri olduğu söylenebilir. Çok iyi düzeyde üst veriye sahip KAA'ların ise kurum yönetimi ve kütüphanecilerin iş birliği ile yapılan KAA'lar olduğu söylenebilir.

Mevcut Üst Verilerin Düzenlenmesi ve Üst Veri Zenginleştirme

Zenginleştirme, üst verilerdeki eksiklerin, düzenleme ise üst verilerdeki hata veya standart dışı yazımların DRIVER rehberine uygun olarak elle ya da farklı teknik yöntemlerle kullanılan şemalarda giderilmesidir. Elle düzenleme oldukça zaman alıcıdır ve insan gücüne gereksinim duyulur. Teknik yollarla yapılan düzenlemeler profesyonel insan gücüne gereksinim duymasının yanında teknik açıdan profesyonellik de gerektirir. Dolayısıyla elle veri düzenleme ve zenginleştirme zaman alıcı, teknik yollarla veri düzenleme ise oldukça riskli bir süreçtir. Teknik yollarla yapılacak düzenlemelerde çok hassas olunmalı ve her türlü olasılık dikkate alınmalıdır. Bu yöntemle yapılan düzenlemeler üst verilerin zenginleştirilerek standart yapıya kavuşması, tekilleştirme için altyapı oluşturulması, otorite alanlarının oluşturulması, KAA içeriğinin harmanlanması, sağlıklı analiz yapılması gibi durumlarda hızlı ve etkili sonuç verir.

KAA'lara Toplu Üst Veri Aktarımı

KAA'lara toplu üst veri aktarımı seçeneği KAA yazılımlarının büyük çoğunluğunda mevcuttur. Ancak KAA'lara yönelik toplu veri aktarımına yönelik yeterli vaka analizi ve teknik doküman bulmak oldukça zordur. Toplu aktarımlardaki basit düzeydeki hatalar bile veri girmekten daha çok zaman alıcı bir sonuç ortaya çıkarabilir. Toplu veri aktarımı öncesi Excel gibi mevcut programlar ve veri düzenlemeye yönelik geliştirilen yazılımlardan destek alınması gerekir. Her yayının üst verisi KAA'ya bir kopya olarak aktarılmalıdır. Oysa bir yayın, TR Dizin, Web of Science, Scopus, PubMed, DergiPark, Kütüphane Otomasyonu gibi kaynaklarda ayrı ayrı kayıtlı olabilmektedir. Her bir sistemdeki yayının toplu olarak aktarılması durumunda KAA'da aynı yayının birden çok üst verisi yanında, her bir veri alanı içeriği tekrar ettiği kayıt sayısı kadar farklı yazım türleri oluşturabilmektedir. Bu nedenle her bir ortamdaki kayıtlar derlenmeli, en zengin kayıt seçilmeli, diğer kayıtlarda olduğu halde seçilen kayıta olmayan veri alanları seçilen kayda eklenerek kayıt zenginleştirilmeli ve tekilleştirme sağlandıktan sonra KAA'ya aktarılmalıdır.

Farklı platformlardan üst veri derleme, dönüştürme, düzenleme, temizleme, zenginleştirme, tekilleştirme ve KAA'ya aktarım süreci veri uzmanlığı yanında alanda iyi düzeyde bilgi sahibi olmayı gerektirir. Bu alanda çalışan profesyonel KAA yöneticileri üst veri çeşitliliği ve düzenleme zorluklarının farkındadır. İyi bir raporlama ve analiz için kaliteli üst veriye ihtiyaç duyulduğunun farkındadırlar ve KAA'larını bu çerçevede yönetmeye çalışırlar. Temelde profesyonel KAA yöneticilerinin analiz yaklaşımı, üst verilerin başlangıçta amaçlanmayan yeni bir bağlamda kullanılması olarak tanımlanan üst verilerin başka bir süreç için yeniden kullanılmasına odaklanmalarıdır. Üst veri iş akış süreci her ne kadar belli şablonlar çerçevesinde yapılsa da tam olarak oluşturulmuş olmaz ve ihtiyaca göre geliştirilmeleri gerekir. Yazılımların ve şablonların çeşitliliği üst veri dönüştürme, düzenleme, zenginleştirme, aktarım gibi zorlukları da beraberinde getirir. Önemli olan her bir veri parçasının standartlara uygun olarak bir şablonda ayrı alanlarda veya aynı alanda olsa bile belli etiketlerle

etiketlenerek depolanmasıdır. Standartlara uygun yapıdaki verilerle yapılacak çalışmalarda veri kaybı yaşanması da söz konusu olmaz.

Türkiye’deki KAA’larda İş Akış Süreçleri

Birçok ülkede olduğu gibi Türkiye’deki KAA girişimlerinde üst veriler, kurumdaki alt yapı, personel ve yönetimin beklentileri çerçevesinde çeşitli düzeylerde gerçekleşmektedir. Akademik kurumlarda görünürlük ve akademik performans ölçümlerine yönelik çalışmalar, KAA’ların akademik performans ölçümleri için temel oluşturduğunun farkına varılmasına büyük katkı sağlamıştır. Bununla birlikte doğru analizler için zengin verilere ihtiyaç duyulduğu gözlenmiştir. Bu çerçevede KAA yöneticileri veri giriş formlarının detaylandırılması suretiyle daha kaliteli veri girmeye yönelmişler ve ayrıca eski verileri de zenginleştirme yoluna gitmeye başladığı bu çalışma kapsamında yapılan incelemelerde gözlemlenmiştir.

KAA’ların işlevselliği, veri giriş formlarının detaylı tasarlanmış olması üst veri oluşturma ve kalitesi bakımından etkilidir. Türkiye’de KAA yazılımı olarak %96,8 civarında DSpace kullanılmaktadır (ANKOS, 2023). DSpace’in varsayılan kurulumunda sadece geleneksel veri giriş formu mevcuttur ve basit düzeyde tasarlanmıştır. Varsayılan bu veri giriş formu bilimsel iletişim, raporlama ve güncel eğilimleri karşılayabilecek yapıda değildir. Ancak DSpace’de yayın türüne göre veri giriş formları tasarlanabilmekte ve ihtiyaç duyulduğu kadar alan veri tabanına eklenebilmektedir. Üst veri girişleri de en az bir editör yetkisine sahip kişinin onayına müteakip KAA’da depolanmaktadır. DSpace’deki iş akış süreci kaliteli veri oluşturulması öngörüsü ile oldukça iyi düzeyde tasarlanmıştır (Çelik ve öte, 2013).

Üst Veri Kalitesi Göstergeleri ve Ölçümleri

“Bir bilgi kaynağını tanımlayan, açıklayan, konumlandırılan veya başka bir şekilde almayı, kullanmayı veya yönetmeyi kolaylaştıran yapılandırılmış bilgiler, üst veri olarak adlandırılmakta; kütüphane ortamında, üst veriler yaygın olarak, dijital veya dijital olmayan herhangi bir nesne türüne uygulanan herhangi bir resmi kaynak tanımlama şemasına işaret etmektedir.” (Özel, 2022: 702-703. ss.).

Üst veriler, temsil ettiği kaynağın etkili bir şekilde bulunmasına, temsil ettiği kaynağı anlamaya, güvenmeye ve kullanmaya yardımcı olur. Üst veri kalitesini ölçmek oldukça zordur. Üst veri alanları konusunda genel kabul görmüş şemalar, bu şemalarda bir yayına yönelik kullanılması zorunlu alanlar belirlenmiştir. Ancak veri girişinde şemalardaki zorunlu olmayan alanlar ile kuruma ve yayınlara yönelik özel verilerin de kayda dâhil edilmesi veri kalitesini üst seviyeye çıkarması bakımından önemlidir. Bununla birlikte üst veri kalitesinin değerlendirilmesinde, temsil ettiği kaynağın keşfi, kullanımı, güncelliği, orijinliliği, tamlığı, doğruluğu, tutarlılığı seçenekleri de oldukça önemlidir.

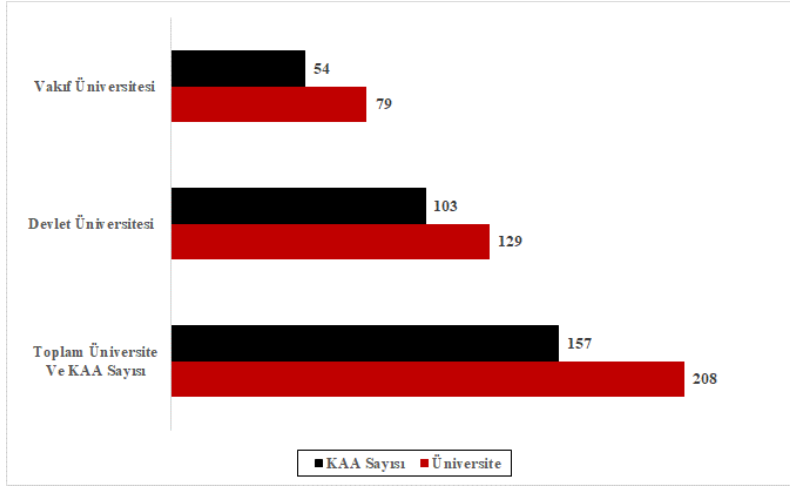
KAA’lardaki veri giriş formlarının detaylı olarak tasarlanması üst veri kalitesini etkileyen önemli bir husus olmakla birlikte, KAA’lara veri girişi yapanların bir yayına yönelik bilgileri eksiksiz ve doğru olarak girmesi de önemlidir. KAA’lara veri girişi yapan profesyonel kütüphaneciler, veri girişinde DRIVER rehberine ilave olarak kataloglama kurallarına dikkat ederek genelde düzenli ve detaylı veri girmektedir. Veri girişindeki temel sorunlar, KAA’lara veri girişlerinin profesyonel olmayan kişilere yaptırılması, veri giriş formlarının detaylı ve kaynağa özel tanımlanmaması, KAA’ların yazar ve kurumlar için akademik performans ölçümleri için temel bir kaynak olduğunun yerince anlaşılması sayılabilir. Üst verilerde bulunan temel sorunlar ise hatalı veriler, yanlış öğeler, eksik bilgi, tutarsız veriler olarak söylenebilir. Üst veri kalitesi “tamlık”, “doğruluk”, “tutarlılık” göstergeleri ile değerlendirilebilir (Bellini, 2013):

- Tamlık, bir kaynağı temsil eden tüm bilgilerin üst veri şemasındaki alanlara eklenmesidir. Üst verilerde bazı alanların genellikle boş bırakılması, bu alanların gerekli olmadığı veya ihtiyaç duyulmadığı anlamına gelmez. Örneğin; kaynaktan mevcut olduğu halde çeviren, derleyen, sponsor, kurum bilgisi gibi alanların kayda eklenmemesi.
- Doğruluk, serbest metin alanlarında yazım hatasının bulunmaması ve alanlardaki verilerin önceden belirlenmiş standartlara uygun olarak oluşturulmasıdır. Örneğin; kaydın standartlara uygun dosya türü, dil, yazım biçimi ve benzeri nesnel bilgileri içermesi.
- Tutarlılık, temsil ettiği kaynakla uyumlu üst veri veya güvenilir üst veri olarak ifade edilebilir. Bir üst veri kaydında, gözden kaçan bir tutarlılık kontrolünün sonuçları birçok alanı etkileyebilir. Örneğin; üst verinin temsil ettiği bilgilerin kaynak bilgilerinden farklı olması.

“Doğruluk” değeri “tamlik” değerinden daha küçüktür çünkü bu süreçte analiz edilen alan sayısı, tüm üst veri alanlarının dikkate alındığı “tamlik” değerlendirmesi sırasında değerlendirilen alanların sayısından daha azdır. Aynı husus “tutarlılık” değeri için “doğruluk” değeri açısından da geçerlidir. Bunun nedeni, bazı alanlarda “doğruluk” veya “tutarlılık” boyutunun otomatik bir süreçle değerlendirmenin zor olmasıdır (Bellini, 2013).

Türkiye’deki KAA’ların Üst Veri Örnekleme

Türkiye’deki KAA’ların üst veri örnekleme geçmeden önce üniversite sayısı, KAA sayısı ve bunların devlet ve vakıf üniversiteleri bazında tespit edilmesi önemlidir. Bu kapsamda konu ile ilgili detaylı bilgiler Şekil 1’de verilmiştir.



Şekil 1. Türkiye’de KAA’lar ve Üniversiteler (YÖK, 2023; ANKOS, 2023)

Şekil 1’de görüldüğü üzere Türkiye’de 129’u devlet ve 79’u vakıf olmak üzere toplam 208 yükseköğretim kurumu bulunmaktadır Devlet üniversitelerinin %79,84’ü (103 üniversite) ve vakıf üniversitelerinin %68,35’inin (54 üniversite) KAA’sı olduğu görülmektedir. Genel toplamda ise üniversitelerin %75,48’i (157 üniversite) KAA’ya sahiptir.

Türkiye’de KAA’lardaki koleksiyon sayısı, kaynakların “Handle” numarasının olup olmaması, sistemlerin “OpenDOAR” ve OpenAIRE’de kayıt durumları da bu çalışma kapsamında incelenmiştir. Elde edilen bulgular Tablo. 1’de sunulmuştur.

Tablo 1: Türkiye’deki Kurumsal Akademik Arşivlerin Üyelik Durumları ve Kayıt Sayısı

KAA Sayısı	Kayıt	ANKOS AE Rehberine Kayıtlı Üniv. Sayısı	“Handle” Kayıtlı KAA Sayısı	OpenAIRE Kayıtlı KAA. Sayısı	OpenDOAR Kayıtlı KAA Sayısı
1-1.000		25	18	7	9
1.001-5.000		68	63	42	55
5.001-10.000		26	23	16	19
10.001-20.000		13	12	12	12
20.001-40.000		9	8	3	6
40.001-80.000		7	6	4	4
80.000 >		9	8	5	8
Toplam		157	138	89	113

Kaynak: (Harman, 2023; ANKOS, 2023)

Tablo 1’de görüldüğü üzere KAA’ların %43,4 (68 üniversite) 1.000 ile 5.000 arasında kaynağa sahiptir. 20.000’in altında kaynağa sahip olanların oranı ise %75,80 (119 üniversite) olduğu görülmektedir. KAA’ların %86,24 (138 üniversite) ile büyük çoğunluğu Handle kullanmaktadır. “OpenAIRE”e kayıt olma oranı %56,69 (89 üniversite) iken “OpenDOAR”a kayıtlı olma oranı %61,07 (113 kurum) olduğu görülmektedir.

Türkiye'deki KAA'ların Üst Veri İncelemesi

Türkiye'de yukarıda da belirtildiği gibi 157 adet kurumsal akademik arşiv vardır. Bu arşivlerin üst verilerinin kalite açısından değerlendirilmesi için DC alanları temel alınarak bir değerlendirme yapılmıştır. Değerlendirilmek üzere seçilen KAA'lar için ön koşul KAA'nın Handle üyeliğinin olması, OpenAIRE ile TÜBİTAK ULAKBİM Harman arşivlerinde harmanlanması ve OpenDOAR ile ANKOS rehberlerinde kayıtlı olmasıdır. Ayrıca değerlendirme için her arşivde en az bin adet kayıt içermesi de ön koşula dâhil edilmiştir. Tüm bu koşulları taşıyan 101 adet KAA belirlenmiş ve çalışma bu kurumlar ile devam edilmiştir. Çalışmayı yönetebilmek ve örneklem oluşturmak için 101 KAA'dan rastgele 17 KAA seçilmiştir. Bunların 9 adedi devlet, 8 adedi vakıf üniversitesi KAA'sıdır. Seçilen ve incelenen arşivlerin (17 KAA) toplam arşive (101 KAA) oranı %16.83'tür ve Türkiye'deki KAA'ların genelini temsil edeceği düşünülmektedir.

Her bir arşivden 50'şer adet olmak üzere toplamdan 850 adet kayıt derlenmiştir. Kayıtlara ulaşmak için KAA'larda İngilizce ve Türkçe dillerinde yazılmış yayınlara erişmek için "COVID" anahtar kelimesi kullanılmıştır. Bu anahtar kelime ile farklı dillerdeki yayınlara erişmek ve ortak alanları daha doğru tespit etmek hedeflenmiştir. Erişilen kayıtlar Microsoft Excel'e aktarılmıştır. Daha sonra her bir kayıt DC alanlarına göre tek tek incelenmiş Tablo 2'de belirtilen genel sonuçları detaylı olarak sunulmuştur.

Tablo 2: Türkiye'deki KAA'lardaki Üst Veri Kalitesi

DC Alanları	İncelenen Kayıt Sayısı	Tamlık (%)	Doğruluk (%)	Tutarlılık (%)	Ortalama (%)
Başlık	850	100.00	76.60	93.00	89.87
Yazar	850	92.80	83.00	94.00	89.93
Katkı Sağlayan	850	98.00	80.00	83.00	87.00
Konu	850	85.64	76.00	74.00	78.55
Açıklama	850	5.00	4.00	7.00	5.33
Yayıncı	850	97.00	68.00	68.00	77.67
Tarih	850	98.00	98.00	89.00	95.00
Tür	850	96.00	92.30	91.00	93.10
Dil	850	100.00	77.00	92.00	89.67
Kaynak	850	21.00	40.90	78.00	46.63
Haklar	850	87.70	87.70	68.00	81.13
Kayıt Bilgisi	850	100.00	100.00	100.00	100.00
Hak Sahibi	850	0.00	0.00	0.00	0.00
Hedef Kitle	850	0.00	0.00	0.00	0.00
Ortalama	850	70.08	63.11	66.93	66.71

Tablo 2'de görüldüğü üzere kayıtların tamlık durumları ortalama %70.08 iken tutarlılık ortalaması %66.71 ve doğruluk ortalaması ise %66.93 olarak gerçekleşmiştir. DC alanlarından yayınlara erişim için kayıt bilgisi (%100), tür (%93.10) ve tarih (%95) verileri en doğru biçimde girilirken; hak sahibi (%0), hedef kitle (%0) ve açıklama (%5.33) alanlarının en az sisteme girildiği gözlenmiştir.

Bu çalışmanın amacına uygun olarak KAA'lardaki kayıtlar üst veri kalitesi açısından incelenmiştir. Bu bağlamda bir kayıt için zorunlu DC alanlarının %33'ü girilmiş ise bu kayıt üst verisi için "zayıf", eğer %66'ı girilmiş ise kayıt üst verisi "orta" ve yine aynı kaydın %67'den fazla alanlarının verisi girilmiş ise bu kaydın üst verisi iyi veya kaliteli olarak varsayılmıştır. 17 üniversitedeki KAA'dan 850 kaydın tamlık bağlamında ortalamasına bakıldığı zaman %70.08 ile üst kayıt bilgilerinin iyi olduğu sonucuna varılmıştır. Yine aynı yöntemle kayıtların doğruluk (%63.11) ve tutarlılık (%66.93) ortalamasına bakıldığı zaman kayıtların üst veri kalitesinin orta düzeyde olduğu sonucuna varılmıştır. Kayıtların tamlık, doğruluk ve tutarlılık genel ortalamasına göre ise %66.71 ile orta düzeyde olduğu görülmektedir.

Veri giriş şemalarında DC alanları dışında zorunlu olmayan ve kayıtlar için önemli bilgiler içeren üst veri alanları vardır. Bu üst veriler yayınlara erişim, yayın hakkında bilgi edinme gibi hususlarda önemi yüksek olan DOI, Cilt, Sayı, Kategori, Kurum, ORCID ve Abstract gibi birçok ek alandan oluşur. 17 KAA'nın 850 kaydı esas alınarak DC'de zorunlu olmayan söz konusu alanlar da incelenmiştir. Zorunlu

olmayan bu alanlara ortalama üst veri girişi %35 olarak tespit edilmiştir. Bu kapsamda zorunlu olmayan alanda üst veri girişlerinin “zayıf” çok yakın olduğu tespit edilmiştir.

Sonuç ve Değerlendirme

Üst veri kalitesi önemli bir terimdir ve üst verinin kurumlar tarafından hangi amaçla kullanılacağı üst veri yapısının belirlenmesinde etkilidir. Kurumların akademik hafıza oluşturarak kendi verilerine kolayca erişme, raporlama, internette görünürlük ve keşfedilebilirlik beklentileri KAA'larda üst veri yapısının belirlenmesinde etkilidir. KAA'lardaki üst veri yapısının belirlenmesinde farklı platformlardan derlenen verilerin yapısı da etkili olmaktadır. Çünkü derlenen üst verilerdeki alanlar KAA'da mevcut değilse bu alanları KAA'da yansıtılması için veri tabanında olmayan alanların eklenmesine vesile olur.

KAA'lara veri girişi belli standartlar çerçevesinde yapılmasına rağmen yerel uygulamalarda farklılıklar doğal olarak ortaya çıkmaktadır. Yerel ya da kuruma yönelik uygulamalar standart uygulamaları etkileyecek şekilde olmadığı sürece genel veri yapısına zarar vermeyeceği gibi ve harmanlama sürecini de olumsuz etkilemez. KAA yöneticileri yerel uygulamaları genelde ana şablona uygun olarak yaparlar ve takip ettikleri belli başlı arşive ve kılavuzlara da dikkat etmeye çalışırlar. Özellikle DRIVER ve OpenAIRE kılavuzları bu süreçte oldukça faydalı olmaktadır. Veri girişlerinde belli kılavuzların takip edilmesi üst verilerin birlikte çalışabilirliğini sağlamaktadır.

KAA içeriklerinin arama motorlarında yer alması günümüzde farkındalığın artmasını sağlayan önemli bir seçenektir. Bilişim teknolojileri alanındaki farklı uygulamalar bilimsel iletişimi daha güçlü hale getirmektedir. Dolayısıyla zenginleştirilmiş üst veriler de bu iletişimin anahtarı durumundadır. Üst veri ile uğraşan kişilerin, üst veri kalitesinin yanında farklı analiz ve rapor beklentilerini öngörerek üst verileri nasıl zenginleştireceklerinin yollarını da aramalıdır.

Türkiye'deki KAA'ların içeriği her geçen gün hızla artmaktadır. Bir yandan da kurumların ve araştırmacıların görünürlük, analiz, raporlama ve performansa ilişkin ölçüm beklentileri farklılaşmaktadır. KAA sistemlerinin değişen ve gelişen teknolojiye uyumunun yanında farklı sistemlerle entegrasyonu da önemli hale gelmektedir. KAA'ların beklenti ve uygulamalara yanıt verebilmesi için iş akış süreçlerinin daha denetimli hale gelmesi ve içeriklerindeki üst verilerin zenginleştirilmesi ihtiyacı ortaya çıkmıştır.

Türkiye'de üniversitelerin büyük çoğunluğunun (%75.48) KAA'ya sahip olduğu görülmektedir. Diğer taraftan KAA'ların Handle kullanımı, Harman, OpenDOAR, OpenAIRE gibi ulusal ve uluslararası sistemlere kayıtlı olması ve entegrasyonun geliştirilmesi gerekmektedir. Bu durumun bilgiye erişim, yayınların görünürlüğü, veri aktarımı gibi birçok açıdan önemlidir.

Türkiye'de bulunan 157 KAA'ların %75.80'ni 20.000 altında kayda sahiptir. KAA'lar sayı olarak artmasının yanında içerik olarak da geliştirilmesi gerekmektedir.

Çalışmasının temel çıkış noktası KAA'larda üst veri kalitesini tespit etmektir. Yapılan örneklem çalışmasında KAA'lardaki kayıtların tamlık, doğruluk ve tutarlılık ortalaması orta düzey kayıt olarak çıkmıştır. Bu kapsamda KAA veri girişlerinin geliştirilmesi gerektiği söylenebilir. Yine KAA kayıtlarının kaliteli getirilmesi için mevcut iş akış süreçlerinin ve formalarının beklentilere göre elden geçirilmesi, KAA'larda mevcut üst verilerin elden geçirilerek anlamlı hale getirilmesi faydalı olacaktır.

Bu çalışmada KAA'lardaki üst veri kalitesinin bilimsel ve yönetsel alandaki etkisini ve kaliteli üst veri için atılması gereken adımları paylaşılmaya çalışılmıştır. Ayrıca bu çalışma, kütüphanecilere, kullanıcılara, araştırmacılara ve bu alana ilgi duyan herkese katkı sağlayabilecek özgün ve orijinal bir çalışmadır. Bundan sonra bu alanda yapılacak diğer bilimsel çalışmalar için de bir temel oluşturacaktır.

Kaynakça

Anadolu Üniversite Kütüphaneleri Konsorsiyumu [ANKOS]. (2023). Türkiye'deki kurumsal akademik arşivler. <https://docs.google.com/spreadsheets/d/1k1UMDJPBlzOKEGk6PRYttw6lHeJ3V3nDm dLNhD6gEk/edit#gid=2107335516> adresinden erişildi.

Ayla Stein, Kelly J. Applegate & Seth Robbins (2017). Achieving and maintaining metadata quality: Toward a sustainable workflow for the IDEALS institutional repository. *Cataloging & Classification Quarterly*, 55:7-8, 644-666, DOI: 10.1080/01639374.2017.1358786

- Aziz, A. (2022). *Sosyal bilimlerde araştırma yöntemleri ve teknikleri* (14. bs.). Ankara: Nobel Yayıncılık.
- Bellini, E., Nesi, P. (2013). Metadata quality assessment tool for open access cultural heritage institutional repositories. In: Nesi, P., Santucci, R. (eds) *Information Technologies for Performing Arts, Media Access, and Entertainment. ECLAP 2013. Lecture Notes in Computer Science*, vol 7990. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40050-6_9
- Bulut, B., Çelik, S., Gürdal, G., Holt, İ., Kafalı Can, G., Madran, O ve Tonta, Y. (2020). Türkiye’de açık erişim ve açık bilim. *ÜNAK2020 Akademik Yayıncılık, Açık Bilim ve Kütüphane(ciler Sempozyumu, 1-2 Ekim 2020*, Yaşar Üniversitesi, İzmir. Erişim adresi: <https://hdl.handle.net/20.500.12154/1262>
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2022). *Eğitimde bilimsel araştırma yöntemleri* (33. bs.). Ankara: Pegem Akademi Yayıncılık.
- Çelik, S. (2016). Kurumsal akademik arşiv sistemine kayıt giriş rehberi. Erişim adresi: <https://www.sonmezcelik.net/2016/10/kurumsal-akademik-arsiv-sistemine-kayt.html>
- Çelik, S., Çelik, R. (2019). Sürdürülebilir bir kurumsal akademik arşiv yönetimi: Doğu Üniversitesi akademik arşiv sistemi deneyimi. *Yükseköğretim Dergisi*. 9(1), 91-102. Erişim adresi: <https://doi.org/10.2399/yod.18.020>
- Çelik, S., Gürdal, G. Ketten, B., Kutlutürk, L., Türkfidanı, A., (2013). Açık erişim ve DSpace kurumsal arşiv yazılımı. *Akademik Bilişim - 2013 Konferansı, Akdeniz Üniversitesi Enformatik Bölüm Başkanlığı 23-25 Ocak 2013*, Antalya.
- DAMA International (2017). *DAMA-DMBOK: Data management body of knowledge*: 2nd ed., Technics Publications
- DRIVER Rehberi 2.0 : İçerik sağlayıcılar için rehber – OAI-PMH ile metinsel bilgi kaynaklarının keşfi.* B. Yazıcı (Çev.), S. Çelik, G. Güneş, G. Gürdal (Gözden Geçiren ve Yay. Haz.), İstanbul. Erişim adresi: https://ankos.org.tr/wp-content/uploads/2018/10/DRIVER-Rehberi_19_08_2014_tr.pdf
- Harman (2023). Harman: Türkiye akademik arşivi. Erişim adresi: <https://harman.ulakbim.gov.tr/harvestingManual>
- Jeannette Ho, & Charity Kay Stokes. (2019). Core metadata element recommendations for institutional repositories at Texas A&M University Libraries. *Journal of Library Metadata*, 19, 187–214.
- Kurtz, M. (2010). Dublin Core, DSpace, and a Brief Analysis of three university repositories. *Information Technology and Libraries*, 29(1), 40–46. <https://doi.org/10.6017/ital.v29i1.3157>
- Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital repositories. *International Journal on Digital Libraries*, 10(2–3), 67–91. <https://doi.org/10.1007/s00799-009-0054-4>
- Osman, R., Yanti Idaya, A. M. K., & Abrizah, A. (2023). Metadata matters: evaluating the quality of Electronic Theses and Dissertations (ETDs) descriptions in Malaysian institutional repositories. *Malaysian Journal of Library & Information Science*, 28(1), 109–125. <https://doi.org/10.22452/mjlis.vol28no1.7>
- Özel, N. (2022) Bilginin düzenlenmesinde yeni eğilimler: Anlamsal web ve bağlı veri. Erol Yılmaz ve Leyla Kanık (Eds.). *Bilgi Merkezlerinde Yönetim-II* içinde (ss.689-716)
- Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2014). Metadata quality in digital repositories: Empirical results from the cross-domain transfer of a quality assurance process. *Journal of the Association for Information Science and Technology*, 65(6), 1202-1216–1216. <https://doi.org/10.1002/asi.23045>
- Palavitsinis, N., Manouselis, N., & Sánchez-Alonso, S. (2017). Metadata and quality in digital repositories and libraries from 1995 to 2015: A literature analysis and classification. *International*

- Information and Library Review*, 49(3), 176-186-186.
<https://doi.org/10.1080/10572317.2016.1278194>
- Park, J.-R. (2009). Metadata quality in digital repositories: A survey of the current state of the art. *Cataloging and Classification Quarterly*, 47(3-4), 213-228-228.
<https://doi.org/10.1080/01639370902737240>
- Piazzini, T. (2022). Bibliographic control and institutional repositories: welcome to the jungle. *JLIS.It*, 13(1), 132-142-142. <https://doi.org/10.4403/jlis.it-12717>
- Sandy, H. M., & Dykas, F. (2016). High-quality metadata and repository staffing: Perceptions of United States-Based OpenDOAR participants. *Cataloging and Classification Quarterly*, 54(2), 101-116-116. <https://doi.org/10.1080/01639374.2015.1116480>
- Seyidođlu, H. (2009). *Bilimsel araştırma ve yazma el kitabı* (10. bs.). İstanbul: Güzem Can.
- Tang, M., Yu, F.-H., & Bever, J. D. (2017). Open access increases citations of papers in ecology. *Ecosphere*, 8(7). <https://doi.org/10.1002/ecs2.1887>
- Yuji Tosaka, & Jung-ran Park. (2010). Metadata quality control in digital repositories and collections: Criteria, semantics, and mechanisms. *Cataloging & Classification Quarterly*, 48, 696-715.
- Yükseköğretim Kanunu* (1981). T.C. Resmi Gazete. (22354, 4 Kasım 1981).
- Yükseköğretim Kurulu [YÖK]. (2023). *YÖKAKADEMİK*. <https://akademik.yok.gov.tr/AkademikArama/view/universityListview.jsp> adresinden erişildi.