



Açık Bilim ve Açık Veri: COAR'ın (Açık Erişim Bilgi Depoları Konfederasyonu) Gelecek Nesil Kurumsal Bilgi Depoları

Open Science and Open Data: COAR (Confederation of Open Access Repositories) 's the Next Generation Institutional Repositories

Senem ACAR

University of North Carolina at Greensboro, raca_menes@yahoo.com

Öz

Bu çalışmada, açık erişim akademik çıktılarından açık veriye doğru evrilen veri-bilgi-bilim üretim alışkanlıklarının kurumsal bilgi deposu kavramına yansımaları incelenirken, açık erişimin savunucusu olan Açık Erişim Bilgi Depoları Konfederasyonu (Confederation of Open Access Repositories - COAR) çatısı altında faaliyet gösteren “Üstveri Kalitesi” Çalışma Grubunun hazırladığı inceleme, araştırma ve analizler doğrultusunda, “Yeni Nesil Kurumsal Bilgi Depoları”nda bulunması gereken nitelikler ortaya konmuştur. Literatüre FAIR ilkeleri olarak geçen bu standartlar çerçevesinde büyük ölçekli ve karmaşık açık verinin ve araştırma çıktılarının nasıl yönetilebileceğinin incelenmesinin ardından, bu ilkelerin uygulanabileceği akademik disiplinlere yönelik veri yönetim sistemleri incelenmiş, olası uygulamalar için örnekler verilmiştir. Bu çerçevede akademik yayın çıktılarından açık erişimle sunulduğu kurumsal bilgi depolarının yanı sıra e-bilim ve Bilim 2.0 gibi kavramlarla araştırma verisinin de depolandığı değişen siber altyapı ihtiyaçları incelenmiş, bu ihtiyaçları karşılayabilecek, özellikle araştırma verisi üretilirken aynı zamanda bu tür büyük ölçekli veriyi işlemeyi öngören ve ‘Dataverse’ e evrilen yapay zekâ ile güçlendirilmiş bilişim ekosistemleri değerlendirilerek bunların hangi FAIR standardını karşıladığı analiz edilmiştir. Ayrıca, COAR Üstveri Kalitesi Çalışma Grubunun faaliyetleri kapsamında yürütülen çalışmalarda, yalnızca kütüphanecilik ve kurumsal bilgi depoları açısından değil, araştırmacı perspektifinden doğabilecek olası ihtiyaçlara ilişkin öngörüler de özetlenmiş; bu doğrultuda yeni nesil kurumsal bilgi depoları için çeşitli önerilerde bulunulmuştur.

Abstract

This study examines the reflections of evolving data-information-science production habits, transitioning from open access academic outputs to open data, on the concept of institutional repositories. It focuses on the analysis, research, and review prepared collaboratively by the "Metadata Quality" working group under Confederation of Open Access Repositories (COAR), an advocate of open access, which identifies the essential features required for "Next-Generation Institutional Repositories." Within the framework of the standards known as the FAIR principles in the literature, the study explores how large-scale and complex open data and research outputs can be managed. It then examines data management systems tailored to academic disciplines where these principles can be applied, providing examples of possible implementations. In this context, the study investigates the changing cyber-infrastructure needs that support not only institutional repositories offering open access to academic publications but also research data storage under concepts like e-Science and Science 2.0. It evaluates AI-powered computing ecosystems, which are evolving into dataverse platforms and are designed to process large-scale data during research data production. The alignment of these ecosystems with specific FAIR standards is analyzed.

Bilgi Yönetimi Dergisi

Cilt: 8 Sayı: 1 Yıl: 2025

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



Hakemli Makaleler

İnceleme Makalesi

Makale Bilgisi

Gönderildiği tarih: 23.09.2024

Kabul tarihi: 07.04.2025

Yayınlanma tarihi: 30.06.2025

Article Info

Date submitted: 23.09.2024

Date accepted: 07.04.2025

Date published: 30.06.2025

Anahtar Sözcükler

Gelecek Nesil Kurumsal Bilgi Depoları, Üstveri Kalitesi, Bilim 2.0

Keywords

Next Generation Institutional Repositories, Metadata Quality, Science 2.0

DOI numarası

10.33721/by.1554995

ORCID

0000-0001-9374-9018



Additionally, the discussions within the scope of COAR Metadata Quality working group activities are summarized, not only from the perspective of librarianship and institutional repositories but also from the viewpoint of researchers' potential needs; recommendations are provided for next-generation institutional repositories.

1. Giriş

Yapay zekâ teknolojilerinin giderek daha önemli bir hâle geldiği günümüzde üniversiteler ve araştırma enstitülerindeki kurumsal bilgi depolarında toplanan büyük ölçekli ve karmaşık verilerin, araştırma çıktılarının etkin bir şekilde işlenip sunulabilmesi; bulunabilirliği ve uluslararası ortak veri tabanlarında harmanlanabilmesi için üstveri standartlarının uygulanması ile verilerin temiz ve işlenebilir olması büyük önem taşımaktadır. Dünya genelinde kurumsal bilgi depoları için açık erişim prensiplerine uygun bir dijital varlık yönetim sistemi olan DSpace yaygın olarak kullanılsa da kurumlarda farklılıklar gösteren üstveri şablonları, bilimsel çıktıların ortak kullanılan büyük diital veri tabanlarında veri harmanlamasının ve işlenmesinin önünde bir engel teşkil etmektedir. Standartlaşmamış üstveri şablonlarının ve üstveri tanımlamalarının yanı sıra, bu tanımlamalarda birbirinden farklı dillerin kullanılması da ortak veri işlemenin önünde ikinci bir engel oluşturmaktadır. Makine öğrenmesi uygulayan yapay zekâ sistemlerinin veri işlemeye yönelik çalışmaları ise üstveri standartlaşmasını daha da hayati hâle getirmektedir. Kurumsal bilgi depolarında standartlaşma ve ortak bir pratiğin yerleşmesi için uluslararası platformların önemi yadsınamaz. Özellikle bilimsel verilerin ve akademik çıktıların açık erişim olarak yaygınlaştırılmasında önemli bir rol oynayan COAR, üstveri standartlaştırmasında “Üstveri Kalitesi” Çalışma Grubuyla bu alanda çalışmalarını sürdürmektedir. Bu bağlamda bu Çalışma Grubunun çıktılarıyla literature kazandırılan *Findability/Bulunabilirlik, Access/Erişim, Interoperability/Birlikte çalışabilirlik, Resuability/Yeniden kullanılabilirlik* (FAIR) ilkeleri belirlenmiş, gelecekte kurumsal bilgi depolarında olması gereken nitelikler için bir yol haritası çizilmiştir. Bu makalede, hem yapay zekâ ile gerçek zamanlı veri işlenebilecek sistemler araştırılıp incelenmiş, hem de COAR’ın Çalışma Grubunda belirlenen standart ilkeler sunulmuştur. Türkiye’deki üniversite kütüphanelerinin bünyesinde, YÖK’ün çıkardığı rehber ve yönergeler göre oluşturulan kurumsal bilgi depolarına rehber olması amacıyla, bu üstveri standartları ilkelerine makalede yer verilmiştir.

Geleneksel kütüphanecilik ve kurumsal bilgi deposu sistemleri, 15-20 yıl önce geliştirilen ve ihtiyaç duyulan işlevleri desteklemeyen hizmet anlayışlarını sürdürmekte, son yıllarda araştırmacıların ve onların beklentilerinin hızlı değişimine ayak uydurmakta zorlanmaktadır. Her ne kadar açık bilimi destekleyen COAR gibi uluslararası girişimler çaba gösterse de ticari yayıncıların egemen olduğu bir sistem, açık bilimi ve araştırmaları yaymak için ideal bir ortam gibi görünmemektedir. Bu tür ticari yayıncıların açık erişimden çok kâr amacına öncelik verdiği, akademik başarı kriterinde bu yağmacı yayıncıların süreli yayınlarında makalelerin basılmasını şart koştuğu bir ortamda bilimin demokratik bir şekilde yayılmasına engel olan bir kısır döngü söz konusu olmaktadır. Akademik performansın araştırmacıları geleneksel yayıncılık alanlarında yayınlamaya zorlayan sistemde oluşturulan teşvikler ve fonlar sebebiyle bu sorunlar devam etmekte ve kurumların, kurumsal bilgi depoları yoluyla oluşturulan açık erişim kanalları adına yenilik yapma çabalarını büyük ölçüde baltalamaktadır. Yeni nesil kütüphanecilik ise bu anlamda yeni nesil kurumsal bilgi deposu politikaları geliştirerek, daha büyük hızla ve daha fazla verinin üretildiği bir dönemde, yayıncıların hegemonyasını açık ve şeffaf bilgi paylaşımı ekseninde kırmayı amaçlamaktadır. *e-bilim* içinde veri-bilgi-bilim üretme yöntemlerinin değişmesiyle birlikte büyük verinin bir yandan üretilirken bir yandan da işlenmesi ve paylaşılması ihtiyacının artması, bu büyük verinin işlenmesi için teknolojik anlamda ulusal ölçekli siber altyapının geliştirilmesinin gerekliliği, üretilen verinin küratörlüğünün belirli bir uzmanlık ve yetkinlik gerektirmesi, yeni nesil araştırmacının değişen araştırma alışkanlıkları ve ihtiyaçları, kütüphanecilik ve kurumsal bilgi deposu anlayışında da yeni vizyonlar geliştirme ihtiyacını doğurmaktadır. Yeni nesil kurumsal bilgi depoları önermesi, bu anlamda alternatif bir vizyon sunmaktadır: kurumsal bilgi depoları, bilimsel iletişim için küresel ağa bağlı bir altyapının temeli olarak konumlandırmak, bunun üzerine katma değerli hizmet unsurlarının (özgün araştırma çıktılarının ve dijital arşivlerin dışa açılması) eklenmesi sayesinde daha fazla araştırma çıktısına malzeme sağlanması yönünde bir girişim olarak görülebilir.

2. e-Bilim

Yeni nesil kurumsal bilgi depoları ve veri yönetimini gerektiren koşulları oluşturan en önemli etken özellikle son 15-20 yılda araştırma verisinin büyüklüğünün artması, buna orantılı olarak da veri üretme maliyetlerinin düşmesidir. Veri üretme hızının artması, maliyetlerin düşmesiyle orantılı olmuştur. Diğer bir deyişle daha çok veri daha kısa zamanda üretilirken, bu verinin üretiminin maliyetinin azalması daha çok verinin üretilmesini sağlamış, bu olgu da büyük veriye ve onun bir yandan üretilirken bir yandan da işlenmesi ve analiz edilmesi ihtiyacına doğru evrilmiştir. Örnek vermek gerekirse bütün bir insan genomunun ilk tamamlanan dizisi olan İnsan Genom Projesi, 1990'daki tasarımıyla 2003 yılında tamamlanmasına kadar yaklaşık 13 yıl sürmüştür. Bu geçen sürede 2 milyar sterlinden fazla bir meblağ harcanmıştır. Kısa bir süre öncesine kadar, tüm bir genomu sıralamak – yani sarmaldaki 3 milyar deoksiribonükleik asit (DNA) harfinin tamamının sırasını belirlemek- yıllar almışken bugün, yeni nesil dizileme aynı sıralamayı 24 saatte bin sterlinden daha az bir fiyata yapabilmektedir (Chivers, 2018).

Bir yandan düşük maliyetlerle yoğun veri üretimi söz konusuken, bir yandan da geleneksel bilimin araştırma yapma ve veriyi kullanma alışkanlıklarının da buna bağlı olarak değiştiği görülmektedir. Cambridge'de bir insan biyolojisi araştırma merkezi olarak faaliyetlerini sürdüren Babraham Enstitüsünde DNA'nın bağışıklık sistemiyle ve özellikle enfeksiyona karşı savunmak için üretilen antikorlarla olan ilişkisi üzerine çalışan bir gruba liderlik eden bilim insanı Dr. Anne Corcoran, bu değişimden bahsetmektedir. Kendi deyimiyle, kariyerinin ilk dönemlerinde geleneksel bir biyolog olarak ıslak laboratuvarlarda pipetlerle, Petri kaplarıyla, koruyucu gözlükleriyle pratik becerilerini geliştirirken, tezgahlarda cam eşyalarla deneyler yaptığını belirtmektedir. Corcoran'ın ifade ettiği gibi onun alanında, 15 yıl önceki doktora öğrencilerini tamamen ıslak laboratuvarlarda çalışmak üzere ekibe aldıklarını, günümüzde ise ileride bilim insanı olacak doktora öğrencisi adaylarında aradıkları ilk özelliğin, karmaşık biyoinformatik analizlerle baş etmelerini sağlayacak bilgisayar yazılımı kullanım becerisine sahip olmalarının olduğunu ifade etmektedir. Corcoran ayrıca, 'günümüzde bilim insanı olmak için istatistikçi hatta programcı olmanız, karmaşık algoritmalarla çalışabilmeniz gerekmektedir' şeklinde bir yorumda bulunmaktadır. Buradan hareketle Corcoran örneğinde olduğu gibi birçok kıdemli bilim insanı açısından, eğitim aldıkları dönemde var olmayan teknikleri/teknolojileri anlama ve denetleme ihtiyacı doğmuştur. Bu ihtiyaca rağmen üniversiteler, modern bilimsel araştırma yöntemlerinin ihtiyaç duyduğu becerileri öğretmeyen pek çok derece ile bilim insanı yetiştirmektedir. Ancak büyük ve çığır açan bilimsel keşifler, 10 yıl öncesinde mümkün olmayan büyük buluşlar artık yeni tekniklerle, özellikle de yeni nesil muazzam büyüklükte, baş döndürücü derecede büyük veriler ve onların eş zamanlı analizleriyle gerçekleşmektedir (Chivers, 2018).

Şüphesiz veri analizinin baskın olduğu araştırma alanlarını sadece sayısal bilimlere sınırlamak doğru olmaz. Günümüzde dijital beşeri bilimler alanında da yüksek kapasiteli bilgisayar teknolojisine ve altyapısına ihtiyaç duyulmaktadır. Beşeri ve sosyal bilimlerden büyük veri kullanımına örnek, Danimarka'da yürütülen "sosyal doku" çalışmalarıdır. Bu çalışmaların kaynağı, sosyal ağların önemi hakkında bir dizi soruyu yanıtlamayı amaçlayan, anket ve kullanım istatistiği verilerinin bir araya getirilip analizinin yapıldığı disiplinlerarası bir araştırma projesidir. Projenin önemli bir kısmı Sensible Danimarka Teknik Üniversitesi (DTU) deneyidir. Bu deneyde DTU, bin (1000) yeni öğrenciye tüm sosyal etkileşimlerini kaydeden bir akıllı telefon vermiş ve bu projenin temelini oluşturan bir dizi deneysel veri seti sağlamıştır. Öğrencilere yöneltilen ankete ek olarak, hedefin bir kısmının 'büyük veri araştırması' olarak bilinen bağlamda - içsel etik ve metodolojik konulara vurgu yaparak - anket perspektifleri oluşturmak olduğu göz önüne alındığında, deney kendi başına bir araştırma nesnesi haline gelmiştir. Bu çalışma, bir anlamda vatandaş bilimi (citizen science) veya topluluk tabanlı bilim (community science) olarak bilinen, iştirak edenler açısından "bilimsel araştırmaya halkın katılımı" şeklinde belirlenen katılımcı eylem araştırmasının sonucudur (Danish e-infrastructure Cooperation, 2014).

Çalışmada kullanılan yeni dijital teknolojiler bir araç olmuş, akıllı telefonlar bu bağlamda farklı iletişim türlerinde sosyal ağları haritalamak ve davranış özelliklerini anlamak amacıyla "sosyometre" olarak bilinen ölçüm araçları kullanılmıştır. Buna göre her akıllı telefon, tüm sosyal etkinlikleri kaydeden özel bir uygulama ile donatılmıştır. E-posta adresi kullanımı, telefonlaşma, mesajlaşma ve çevrim içi sosyal aktivite, bu alandaki davranışları incelemek amacıyla analiz edilmiştir. Her akıllı telefon günde 50 ila 100 MB arasında veri toplamıştır. Bin öğrencinin akıllı telefon kullanımında, günde 50-100 GB'lık veri

elde edilmektedir. Bu şekilde sürekli artan veri hacimlerini yalnızca süper hesaplama yardımı ile analiz etmek mümkündür. Toplanan bu veriler gerçek zamanlı olarak büyüdükçe, geliştikçe ve karmaşıklaştıkça sosyal ağlar üzerinden davranış analizlerinde yeni perspektifler oluşturmıştır. Proje, sosyal ağları ayırt etmek için yeni araçlar geliştirirken *-hem nicelik hem de niteliksel olarak-* davranış ve bilgilerin ağlar içinde iletilip aktarılmasını ve nasıl iletiltiğini analiz etmeyi amaçlamaktadır: Sosyal ağların kullanımını araştırırken bir yandan da toplanan veriler sağlık, hastalıkların yayılması, uyku düzenleri ve daha fazlası hakkında fikir vermiştir. Veriler, bu nedenle bir dizi varsayımın ampirik doğrulanması olarak daha geniş bir akademik çevre için de ilginç hâle gelmiştir. Bu çalışmanın olası zorluklarından biri, hassas veri içeriklerine sahip olmasıdır. Bu tür projeler, global anlamda hassas verilere dayalı araştırmalarda etik ve özel hayatın mahremiyetine ilişkin hususların nasıl ele alındığının analiz edilmesine katkıda bulunabilir. Proje, bilişim teknolojileri üzerinden sosyal bilimlerin gelişmekte olan alanını teşvik etmeyi amaçlamıştır (Danish e-infrastructure Cooperation, 2014).

Bu şekilde yöntemler uygulanan projeleri Burton ve Appleford, *e-Araştırma* veya *eAraştırma* (*e-Research* veya *eResearch*) olarak tanımlamıştır (2009, s. 2). Mevcut ve yeni araştırma biçimlerini desteklemek için bilgi teknolojisinin kullanımını yaygınlaştırmıştır. e-Araştırma, Science, Technology, Engineering, Mathematics (Bilim, Teknoloji, Mühendislik, Matematik [STEM]) alanlarında kurulan siber altyapı uygulamalarını ve dijital beşeri bilimler Humanities, arts, and social sciences (Beşeri bilimler, sanat ve sosyal bilimler [HASS]) alanları dâhil olmak üzere tüm araştırma alanlarını kapsayacak şekilde e-bilim içinde genişletilmiştir (Burton ve Appleford, 2009, s. 2).

Yeni araştırma ve araştırma verilerini işleme ihtiyaçlarını oluşturan en önemli etken, geleneksel bilim yapma alışkanlıklarının ve bilimsel araştırma uygulamalarının büyük ölçüde değişmiş olmasıdır. Geleneksel bilim, bilim tarihinde iki farklı felsefi yaklaşımın temsilcisidir; bir paradigma değişimi olarak bilimsel dallardaki iki felsefi eğilime üçüncü bir dalın eklenmesinin gerekliliğini tartışmaktadır. “Bilimsellik” ilkeleri bilim tarihi ve felsefesi içinde incelenirken, Robert Boyle henüz 1660'larda bilimsel yayıncılıkta nelerin ön planda olduğunu ve tartışıldığını dile getirmiştir (Bishop ve Gill, 2020, s. 79-80). Açık bilim ve açık veri kavramlarını desteklemek açısından günümüzde de önemli olan kriterler olarak bağımsız doğrulama adına şüphecilik, şeffaflık ve tekrarlanabilirlik prensipleri vurgulanmıştır. “Bilimsellik” kavramı daha sonra tımdengelimli ve deneysel yaklaşımlar olmak üzere iki ana özelliğe ayrılmıştır. Günümüzde ise bilimsel yöntemde ön plana çıkan, bilişim teknolojisi yaklaşımının bir ürünü olarak Victoria Stodden'ın savunuculuğunu yaptığı ve sonuç çıkarma konusundaki hesaplama ve veri analizi adımlarıdır (Pitkin, 2011). Son 20 yılda insanların yüksek performanslı hesaplama ve simülasyondaki değişiklikleri nasıl ele alacakları konusunda yoğun çaba harcanmaktadır.

Bu yeni anlayışta modern bilimsel keşif süreci nasıl gerçekleşir? Modern bilimsel keşif süreci, matematiksel kanıta ve modellemeye dayanması sebebiyle süreçler ve sonuçlar yeniden yürütüldüğünde deterministik olarak verinin aynı şekilde yeniden üretilmesini, ara sonuçların incelemesini ve anlaşılmasını gerektirir (Syed, Ghanem ve Guo, 2007, s. 169). Bu durumda, veri kaynağının basitçe modellenmesi yeterli değildir. Yeni keşifleri destekleyen kanıtlar sağlamak için verilerin analizinden elde edilen hipotezlerin ve sonuçların kökenini modellemek gerekir. Bilim insanlarına, bir bilimsel araştırma kapsamında keşiflerin evrimini belgelemek ve izlemek için verilerin, ara sonuçların ve nihai sonuçların evrimini izlemelerine yardımcı olmak için bilimsel iş akışları önerilmiş ve geliştirilmiştir. Bu süreçler ve iş akışları, araştırma ve gözlem sırasında üretilen büyük verinin üretimi sırasında analizini de içerir. Bu şekilde verinin yeniden kullanımı, üretimi sırasında gerçekleşir (Syed, Ghanem ve Guo, 2007, s. 174).

Bilim insanlarının çalışma şeklini tamamen değiştiren bu unsur, e-bilim başlığı altında değerlendirilebilir. e-Bilim, yüksek düzeyde dağıtılmış ağ ortamlarında gerçekleştirilen ve bilişim teknolojisini yoğun olarak kullanan bilimdir. Bu terim, ilk kez 1999 yılında Birleşik Krallık Bilim ve Teknoloji Ofisi Genel Direktörü John Taylor tarafından, Kasım 2000'de başlayan büyük bir finansman girişimini tanımlamak için kullanılmıştır. e-Bilim, o zamandan beri daha geniş bir şekilde, bilimsel süreç içinde bilgisayar teknolojisini, modern bilimsel araştırma girişimi için uygulanması olarak tanımlanmıştır. Bunlar, veri modellemesi ve analizi, elektronik ya da dijitalleştirilmiş laboratuvar defterleri, ham ve uyumlu veri setleri, yayınların taslak versiyonları, ön baskılar ve basılı ve/veya elektronik yayınları içermektedir. Jim Gray, "veri-yoğunluklu bilim" veya "e-bilim"i, bilimin "dördüncü

paradigması" (deneysel, teorik, hesaplamalı ve üretilirken aynı zamanda analiz edilen veriye dayalı) olarak düşünmüştür ve "bilimle ilgili her şeyin bilginin etkisi nedeniyle değiştiğini" belirtmiştir (Tansley ve Tolle, 2009, s. 177). e-Bilim, bilimsel yöntemin her iki temel ayağında da devrim yaratmaktadır: Bu, dijital büyük veri yoluyla ampirik araştırma ve özellikle bilgisayar simülasyon modeli oluşturma yoluyla gerçekleştirilen bilimsel teori alanıdır.

Elbette ki sözü edilen “veri yoğunluklu bilim” olan e-bilimin gerektirdiği şartlara adapte olacak bilim insanlarının yetiştirilmesi zorunluluğu aşıkardır. Günümüzün şartlarında artık geleneksel laboratuvar deneyleri yapmak yetkin bilim insanı olmak için gereken beceri donanımı olarak tek başına yeterli değildir. e-Bilimde bilişim teknolojisinin ürettiği büyük ölçekli ve karmaşık veri setlerini analiz edip işlemek için kodlama yapacak, güncel yazılım dillerine hâkim ve makine öğrenimi konusunda becerileri olan “yeni nesil” bilim insanlarına ihtiyaç duyulmaktadır. İster fen bilimleri, ister beşeri ve sosyal bilimler olsun tüm araştırma ve akademik disiplin alanları uygulayıcılarından, akademisyenlerinden teknolojik donanımları ve yetkinlikleri gerektirmektedir. Örneğin sanatsal görsel tasarım projelerinde yapılan veri haritalama işlemi, yazılım, web tasarımı becerileri, veri seti işleme ve makine öğrenimi becerilerini geliştirmeyi şart hâline getirmiştir. Her ne kadar bu tür teknolojik projeler, içinde bilgisayar mühendislerinin de olduğu bir grupta takım hâlinde gerçekleştirilse de bir tasarımcının, projesinin hayata geçirilmesi için teknik ve teknolojik açıdan tasarıma etki edecek olasılıkları, kısıtlamaları ve yenilikçilik/gelişim alanlarını tahmin edebilecek, projeye süpervizörlük edecek seviyede yazılım, donanım ve bilişim bilgisine sahip olması zorunlu hâle gelmiştir.

e-Bilim’in diğer bir özelliği, yüksek düzeyde dağıtılmış ağ ortamlarında gerçekleştirilen hesaplamaya yoğunlaşmasıdır. Dağıtık bilişim teknolojisi gerektiren dağıtık hesaplama (distributed computing) ile muazzam veri kümeleri ile çalışan işletim ve ağ sistemlerini kullanır. Dağıtık (distributed) bir bilişim sistemi, bileşenleri farklı ağa bağlı bilgisayarlarda bulunan verileri ve komutları birbirlerine ileterek eylemleri koordine eder. Dağıtık bilgi işlemcilik, bir yazılım sisteminin bileşenlerinin verimliliğini ve performansını artırmak için birden çok bilgisayar arasında paylaşıldığı, veri madenciliğine dayanan bir modeldir. En dar tanımlara göre, dağıtık hesaplama, sınırlı bir coğrafi alandaki bilgisayarlar arasında paylaşılan bileşenlere sahip programlarla donatılmıştır. Bu sisteme göre, bileşenleri oluşturan unsurlar, ortak bir hedefe ulaşmak için birbirleriyle etkileşim hâlinde ve eşgüdümlü olarak çalışır. Bu şekilde geliştirilen ağ sisteminde işlenen veri, yüksek enerji harcayan, büyük ölçekli soğutma ve depolama ihtiyacı nedeniyle sürdürülebilir ve çevre dostu olmayan tek bir süper bilgisayar yerine, birbirinden farklı yerlerde olup veri işlemenin farklı sistemlere dağıtılmasıyla gerçekleşen bu yöntemin üç önemli özelliği ve zorluğu vardır: Bileşenlerin eşzamanlılığı, küresel bir saatin olmaması ve bileşenlerin bağımsız arızası durumunda oluşabilecek aksamalar (Tanenbaum ve Steen, 2002, s. 17-24). Bu sistemler yüksek kapasiteli teknik altyapıları kullanan paydaşlar arasında dağıtık iş birliğini mümkün kılan teknolojileri gerektirmektedir.

e-Bilim’in özü, ya işlevsel olarak birbirine bağlanmış mevcut verilerin ya da yeni oluşturulan *-nispeten küçük-* veri kümeleriyle kombinasyonu için yeni analiz olanakları sağlamasıdır. Bu anlamdaki en önemli adım ise makine bazlı çalışmada verilerde makine destekli “örüntü tanıma” veya “örüntü tanımlama” özelliğini kullanmasıdır. Bunu, bazı kalıpları veya şablonları rasyonalize etmek ve test edilebilir hipotezleri belirlemek için konformasyon analizi içeren insan bazlı bir çalışma takip eder. Bu çalışmalar doğası gereği döngüsel bir süreçtir. Bu tür büyük, dağınık ve değişken veri kümelerinin hesaplamalı analizi, bir e-bilim işleminde çok önemli bir unsurdur. Bilgisayar hesaplamaları, modeller ve veri görselleştirmeleri dâhil olmak üzere mevcut büyük hacimli verilerin işlenmesi, daha önce erişilemeyen karmaşık sorunların çözümüne ve anlaşılmasına olanak tanır. e-bilim geleneksel olarak parçacık fiziği ve biyoinformatik gibi doğa bilimleri disiplinlerinde kullanılmıştır. e-bilim’in diğer bağlamlarda uygulama alanları artmakta ve daha büyük ham verinin daha kısa sürede analiz edilmesiyle yeni araştırma sonuçları elde etme potansiyeli çok büyük olarak değerlendirilmektedir. (Danmark Forsknings- og Innovationsstyrelsen ve Osgood, 2011).

Diğer yandan üretilen büyük verinin çeşitli sürümleri ve versiyon kayıtlarının tutularak bilimsel iş akışlarına dâhil edilmeleri gerekir. Bağlantılar ve bunların arşivlenmiş kopyaları vasıtasıyla web üzerinde oluşturulmuş büyük verinin Memento özelliği ile versiyonlarının tutulması ve hangi

versiyonları olduğunun görüntülenmesi sağlanabilmektedir¹ (Seneca, 2010). Ayrıca e-bilim’de nihai sonuçların iş birliği yapan gruplar tarafından yayımlanmasından, ham verilerin ön deney sonuçlarının ve ilgili bilgilerin “kamuya açık” paylaşımını içeren daha şeffaf bir yaklaşımı olan *Bilim 2.0* anlayışı gelişmiştir. Bilim 2.0’nin temel prensiplerinden olan “kamuya açık” paylaşımın en temel motivasyonu, vatandaşların vergileriyle fonlanan araştırmaların kamu yararlılığını ve bilgi şeffaflığını korumak istemektir. Bilim 2.0’nin amacı bir yandan da ortaklaşa çalışan ilgili taraflar arasındaki iletişimi, iş birliğini kolaylaştıran araçları sağlamaktır. Bu yaklaşımın potansiyeli şöyle açıklanabilir: Bilimsel keşif süresini kısaltmak, daha önce yapılan çalışmaları tekrarlama riskini azaltarak özgün araştırmalara olanak sağlamak, akademik yayıncılık ve akran değerlendirmesi ile ilgili sorunların üstesinden gelmek, zaman ve maliyet engellerini ortadan kaldırarak yeni bilgi üretme sürecini hızlandırmaktır. Ancak araştırma verisinin “açık” olmasındaki en önemli ve hayati mesele ham veri üzerine yapılan çıktı değerlendirmelerinin ve yorumların subjektif olabileceği, belirli bir görüşü desteklemek için ham verinin manipüle edilebileceği ve hatta bir tezi savunmak için çarpıtılabileceği olasılığıdır. Bilimsel yayınların dayandırıldığı ham araştırma verisinin açık veri olarak genel görüşe arz edilmesi, şeffaflığı destekleyeceği gibi aynı zamanda bilimsel etik ilkelerin gözetilmesinde, bu ilkelere uyulmasında önemli rol oynamaktadır.

Son yıllardaki eğilim, Bilim 2.0 olarak adlandırılabilir ağ teknolojilerinin mümkün kıldığı bilgi paylaşımını ve iş birliğini kullanan yeni bir yaklaşımdır. Açık araştırma ve açık bilim hareketlerine benzeyen bu yaklaşım, büyük ölçüde Web 2.0 teknolojilerinden esinlenmiştir. Bilim 2.0, bilim insanları arasında artan iş birliğinin faydalarını vurgulamaktadır. Bilim 2.0 bulguları, ham verileri ve "yeni ortaya çıkan teorileri" çevrimiçi paylaşmak için wiki'ler, bloglar ve video günlükleri gibi birçok ortak araç kullanır. Bilim 2.0, makaleler ve araştırma fikirleri ve kısmi çözümlerle ilgili açıklık ve paylaşımdan yararlanmaktadır. Bilim 2.0 daha işbirlikçi yaklaşımlarla, genellikle Bilim 1.0 olarak adlandırılan geleneksel bilim yapma yöntemleri arasında bir tezat oluşturmaktadır ve yeni bilim biçimlerinin Web 2.0 teknolojileriyle çalışacağını savunmaktadır (Moe-Behrens, 2012, s. 37).

e-Bilim, aynı zamanda bilimsel çalışmaların veri biçiminde bir depoda toplanmasını, işlenmesini ve kullanılmasını içeren bir araştırma yöntemi olması sebebiyle yeni nesil kurumsal bilgi depoları, onların e-bilim’deki aktif bilim üretiminin “kolaylaştırıcısı” olmalarını gerektirir. Bu anlamda yeni nesil kurumsal bilgi depolarının veri toplama, üretme, dijitalleştirme ve daha sonra depolama için daha verimli yöntemler geliştirmek, tüm profesyonel ve akademik alanlarda araştırmayı desteklemek adına daha etkin bir rol almaları gerekir. Yeni nesil kurumsal bilgi depolarının beslenebileceği diğer bir unsur ise araştırma verilerinin birden fazla paydaşın birden fazla kurum içinde ürettiği şebeke bilişim veya dağıtımli hesaplama (grid computing) özelliğidir. Bu, ortak bir hedefe ulaşmak için çeşitli birimlere/kurumlara/coğrafyalara dağıtılmış bilgisayar kaynaklarının kullanılmasıdır. Dağıtımli hesaplama, küme hesaplama gibi geleneksel yüksek performanslı hesaplama sistemlerinden farklıdır. Bunun sebebi dağıtık sistemde çalışan bilgisayarların her bir düğüm kümesinin farklı bir görevi/uygulamayı gerçekleştirmeye odaklanmış olmasıdır. Dağıtık bilişim ile şebeke bilgisayarları ayrıca yoğun paralel işlemciler, yüksek başarımlı vektör işlemciler ve öbek bilgisayarların oluşturduğu süper bilgisayar sistemindeki çalışma mantığından ve küme bilgisayarlarından daha heterojen ve coğrafi olarak dağınık (dolayısıyla fiziksel olarak birbirine bağlı olmama özelliği taşır) olma eğilimindedir.

e-Bilim’e yönelik araştırma faaliyetlerinin çoğu, bilimsel keşfi desteklemek için bilişim teknolojileri araçları ve bunları destekleyen siber altyapıların (cyber infrastructure) geliştirilmesine odaklanmıştır. Çünkü gelişmiş veri toplama, veri depolama, veri yönetimi, veri entegrasyonu, veri madenciliği, veri görselleştirme ve tek bir kurumun kapsamı dışında internet üzerinden dağıtılan diğer bilişim ve bilgi işleme hizmetlerini destekleyen araştırma ortamını sağlamak için güçlü bir siber altyapıyı gerekli kılmaktadır. Bilimsel alanda, siber altyapı, yeni bilimsel teorilerin ve bilginin türetilmesini sağlamak amacıyla laboratuvarları, verileri, bilgisayarları ve insanları verimli bir şekilde bağlama sorununa teknolojik ve sosyolojik bir çözüm üretmeyi amaçlamaktadır.

¹Memento, şu anda aktif olmasa da Web’de arşivlenen içeriği daha kolay keşfedilebilir hâle getirmeyi amaçlayan, Amerika Birleşik Devletleri Ulusal Dijital Bilgi Altyapısı ve Koruma Programı (NDIIPP) destekli bir projedir. Memento, insanların artan sayıdaki web arşivlerini bilmelerini ve aradıkları kaynağın daha eski bir sürümünü hangi arşivde tutabileceğini tahmin etmek yerine, arşivlenmiş içeriği, arayanın zaten bildiği orijinal URL aracılığıyla keşfedilebilir h+ale getirmeyi önerir. Esasen Memento, kullanıcıların herhangi bir web sayfasını geçmişte belirli bir tarihte görüldüğü gibi görüntülenmesine izin verme girişimidir (Seneca, 2010).

Yazılımların karmaşıklığı ve büyük verinin işlenmesi/analizi için arka planda desteklemesi gereken altyapısal gereksinimleri nedeniyle e-bilim projeleri genellikle araştırma laboratuvarları, büyük üniversiteler veya hükümetler tarafından yönetilen ve geliştirilen büyük ekipleri içerir. Open Science Grid (OGS) konsorsiyumu gibi birlikler, arayüzleri dünya çapında coğrafi olarak geniş bölgelere dağıtılmış depolama önbelleklerini birden çok merkezden dağıtık hesaplama üzerinde birleştirmektedir. OSG, çoğu Linux üzerinde çalıştırılarak dağıtılan, toplamda 43.000'den fazla işlemciye sahip 25 binden fazla bilgisayardan oluşmaktadır. 42 üniversite dâhil 72 kurum, şebekeye kaynak sağlayan konsorsiyum üyesidir. Şebekede, Amerika Birleşik Devletleri (ABD) ve Brezilya'yı kapsayan 90 farklı bilişim ve depolama merkezi vardır. Open Science Grid'den yararlanan bilim alanları şunları içerir:

1. Astrofizik, yerçekimi fiziği, yüksek enerji fiziği, nötrino fiziği, nükleer fizik,
2. Moleküler dinamik, malzeme bilimi, malzeme mühendisliği, bilgisayar bilimi, bilgisayar mühendisliği, nanoteknoloji,
3. Yapısal biyoloji, hesaplamalı biyoloji, genomik, proteomik, tıptır.

Bu konsorsiyumlara diğer örnekler Open Science Grid'le ortak çalışan TeraGrid, LHC Computing Grid, the European Grid Infrastructure ve Extreme Science and Engineering Discovery Environment (XSEDE) gibi birliklerdir. Örnek e-bilim altyapıları, Avrupa Şebeke Altyapısı (European Grid Infrastructure), Açık Bilim Şebekesi (the Open Science Grid) ve Nordic DataGrid Tesisi (Nordic DataGrid Facility) dâhil olmak üzere çeşitli ortaklardan oluşan bir federasyon olan Worldwide LHC Computing Grid'i içerir. Bunlar, ABD'de Purdue, Wisconsin-Madison, Clemson, Nebraska-Lincoln, FermiGrid (FNAL), SUNY-Buffalo ve Oklahoma; Brezilya'da "UNESP Birleşik Krallık e-bilim" programı; Birleşik Krallık'ta ise Glasgow ve Edinburgh Üniversiteleri tarafından yönetilen "Ulusal e-bilim Merkezi (NeSC)" dâhil olmak üzere çok çeşitli kaynaklar, merkezler ve kişilerden oluşmaktadır.

e-Bilim ekosisteminde, araştırma verilerinin ve yöntemlerinin optimum verimlilikte kullanımını sağlama zorluğu, çok sayıda paydaş için karmaşık bir sorundur. Verilerini ve yorumlarını paylaşmak isteyen araştırmacılar; veri analizi ve işleme fonksiyonları sağlayan hizmetlerini, yazılımlarını ve araç oluşturmaları sunan profesyonel veri yayıncıları; özel ve kamu olmak üzere, finans kuruluşları; keşfi ve buluşları ilerletmek için çıktıyı bütünleştiren ve analiz eden bir veri bilimi topluluğu madenciliği, uygun veri yönetimleriyle giderek daha fazla ilgilenmektedir. Büyük, birbirine bağlı veri kümelerinde anlamlı kalıpları keşfetmek için hesaplamalı analiz, hızla rutin bir araştırma etkinliği hâline gelmektedir. Bilgi keşfi için ana alt katman olarak ve bu e-bilim süreçlerinin sorunsuz ve sürdürülebilir bir şekilde işlemesi için makine tarafından okunabilir verileri sağlamak, e-bilim'in en büyük zorluklarından biridir. Günümüzde yeni araştırma yöntemleri, büyük veri birikimi ve altyapı gereklilikleri söz konusuysa kütüphane hizmetleri, kütüphaneciler ve kurumsal bilgi depoları da bu durumdan etkilenmektedir.

3. Yeni Nesil Kurumsal Bilgi Depoları

Son zamanlarda yaşanan dijital dönüşümle birlikte artık kütüphane faaliyetlerinin odağı, yerel olarak bir araya getirilmiş bir koleksiyonun tedariki ve düzenlenmesinden daha farklı faaliyetlere doğru kaymaya başlamıştır. Araştırma, bilgi/veri üretme alışkanlıklarının değişmesi ile üretilen bilginin/verinin himayesi, Dempsey tarafından "*içten dışa kütüphane*" (Dahl, 2018, s. 17) olarak tanımlanan yeni bir kütüphanecilik ve kurumsal bilgi deposu anlayışını da beraberinde getirmektedir. Peki gelecekte yeni araştırmacı profiline uygun, hangi hizmetleri ve rolleri üstlenen, ne tür bir kütüphane hizmetleri öngörülmektedir?

Artık temel kütüphane hizmetleri anlamında kaynak sağlayıcı olarak kütüphane koleksiyonlarının hem dijital hem de basılı format şeklindeki karma malzemeyle değişen modelle karşı karşıyayız. Bu karma malzeme bir kayda erişim haklarıyla ilgili olarak kütüphanenin oynadığı rol ve erişim/kullanım politikaları hakkında sorunlar çıkarmaktadır. İçten dışa kütüphanede üretkenlik süreci, araştırma ve öğrenme sürecinde olduğu gibi dijital bir ortamda gerçekleşir. Kütüphaneler, kendi oluşturdukları kaynaklar olarak kurumsal yaratımların (araştırma verileri, ön baskılar, bilimsel profiller, akademik profiller, dijitalleştirilmiş özel koleksiyonlar) üretilmesini, seçilmesini ve keşfedilebilirliğini gün geçtikçe daha fazla desteklemektedir. Son yıllarda kütüphaneler koleksiyonlarını sadece kendi akademik topluluklarının kullanımı için bir araya getirmekle kalmamakta (dıştan içe model), bir yandan da kullanıcı ihtiyaçları çerçevesinde bir araya getirilmiş ve bir misyon içinde koordineli bir yerel, harici ve iş birliğine dayalı karma bir hizmet erişimi sağlamaktadır (Dempsey, 2015a, 52). İçten dışa kütüphane

modelinde kaynakların ve yerel-harici-iş birlikçi hizmetlerin bir karışımının kullanıcı ihtiyaçları çerçevesinde bir ağ mantığına göre düzenlenmesi, “kolaylaştırılmış koleksiyon” olarak tanımlanmaktadır (Dempsey, 2013, s. 5-6). Bu şekilde, araştırma ve öğrenme ihtiyaçlarını materyalleri yerel olarak bir araya getirip bir derleme oluşturarak değil, kaynakları mevcut en iyi yollarla başka kurumların/organizasyonların hizmetleriyle karşılama amaçlanmaktadır. Buradaki anlayış, parayla satın alınan ve fiziksel olarak depolanan ve “sahip olunan koleksiyon” modelinin şu anlayışa evrilmesidir: Araştırmacıyı Google scholar’daki profillere, bedava erişilen e-kitap kaynaklarına (HathiTrust gibi katalogdaki “ücretsiz” e-kitap kaynakları), paylaşılan dijital koleksiyonlara, açık erişim ve web kaynaklarına yönlendiren ve bu şekilde araştırma-öğrenme ihtiyaçlarını en iyi şekilde karşılayan kolaylaştırılmış bir “dış koleksiyon”. “Sahiplik”ten “sağlayıcı”lık rolüne geçişle yaşadığı dönüşüm, kütüphanenin işlevi ve ne yaptığı hakkında düşünme biçimindeki en önemli değişimdir (Dempsey, Malpas ve Lavoie, 2014, s. 397-398).

Geleneksel kütüphanecilikte iki formül, içten dış kütüphane ve kolaylaştırılmış kaynak sağlama, kaynak ihtiyacına yönelik geçici bir çözüm gibi durmaktadır. Dıştan içe koleksiyon anlayışına sahip olan kütüphaneler de artık sahip olmadıkları veya lisans vermedikleri birçok ağ kaynağını belge tedariki yöntemleriyle sağlamaktadır. Bu tedarik etme modeli sadece satın alınan değil, açık erişim politikalarıyla kurum içinde üretilen entelektüel birikimin paylaşımını da kapsamaktadır. Kurumsal bilgi depolarıyla birlikte akademik kütüphane koleksiyonlarının büyük bir kısmı lisanslı olduğu için bu durum kütüphaneyi yerel bir sahiplikten, bilimsel iletişim politikası ile birlikte Creative Commons gibi bir lisanslama modelini benimsemeye itmektedir. Paylaşılan koleksiyon formülünde ise ne kadar büyük olursa olsun, tamamen yerel bir koleksiyon, potansiyel olarak nadir ve tek olan kaynaklar veya geniş kapsamlı kaynaklar olsa bile giderek daha kısmi görünmektedir. Dolayısıyla kütüphanelerde yerel koleksiyonları daha geniş bir ağ bağlamına yerleştirmek konusunda daha da güçlenen bir eğilim bulunmaktadır (Wilkin, 2015, s. 240).

Üniversite ve kütüphane “sağlayıcılık” rolüyle içten dış modelde, kurumun yarattığı entelektüel kaynakları destekler ve bunları hem iç hem de dış kullanıcıya açar. Kurumun benzersiz entelektüel ürünleri arasında arşivler, özel koleksiyonlar veya yeni oluşturulmuş araştırma-öğrenme materyalleri (e-yayınlar, araştırma verileri, eğitim yazılımları, dijital bilimsel kaynaklar, vb.) ile araştırmacı profilleri gibi kaynaklar bulunmaktadır. Genellikle amaç, bu materyalleri kurum dışındaki potansiyel kullanıcılarla paylaşmaktır (Dempsey, 2015b, s. 53).

Yeni nesil kütüphanecilik ve kurumsal bilgi deposu anlayışına göre içten dış kütüphane modelinde kurumsal kaynakların oluşturulmasına, kürasyonun yapılmasına ve kaynakların keşfedilebilirliğine destek unsurları şu alanları kapsamaktadır: Araştırma verisi yönetimi ve araştırma bilgisi yönetimi (diğer adıyla araştırma verilerinin küratörlüğü), son on yılda büyük üniversiteler ve kütüphaneler için önemli bir sorun olarak ortaya çıkmıştır. Bu alandaki gelişmelerin arkasında, fon sağlayıcı kuruluşların talepleri ve verilerin yeniden kullanılabilirliğine yönelik artan ihtiyaç gibi çeşitli motivasyonlar yer almaktadır. Oldukça etkin bir ilgi topluluğu ve gelişmekte olan en iyi uygulama standartlarıyla birlikte, kütüphaneler kampüs içindeki çok paydaşlı faaliyetlerin potansiyel ortakları hâline gelmiştir. Bu kapsamda sundukları hizmetler, danışmanlıktan veri seçimine kadar geniş bir yelpazede çeşitlilik göstermektedir. Kurumsal itibar yönetimi bağlamında ise -özel koleksiyonlar da dâhil olmak üzere- araştırma ve öğrenme materyallerinin keşfedilebilirliğini artırmaya yönelik çabalar öne çıkmaktadır. Bu çabalar, etkili ve şeffaf kurumsal kimlik paylaşımlarının yanı sıra, araştırmacı profilleri, yayınları, ilgi alanları ve başarılarına yönelik artan ilgiyi de kapsamaktadır. Bu sayede söz konusu alanların daha kolay keşfedilebilir hâle gelmesi hedeflenmektedir.

Birçok üniversite kütüphanesi, bilimsel yayıncılık kapsamında kampüs içi dergi yayıncılığına destek sağlamakta ve üniversite yayınevleriyle çeşitli iş birlikleri yürütmektedir. Bu doğrultuda, bazı kütüphaneler açık erişim yayıncılık modellerini desteklemeyi misyon odaklı bir sorumluluk olarak benimsemektedir. Açık erişime geçiş yeni olmamakla birlikte farklı model ve telif yetkileri nedeniyle veri/bilgi üretenler için karmaşık gelebilir. Kurumlar uyumluluk yönetimi, makale işlem ücretleri vb. gibi konularla ilgilenmektedir. Bu anlamda kütüphanenin merkezi bir danışma ve hizmet rolü vardır. Özel koleksiyonlar ile son zamanlarda kütüphanelerde prestij adına bu koleksiyonların ayırt ediciliğine odaklanma, dikkati özel koleksiyonlara ve arşivlere, bunların araştırma ve öğrenme pratiğindeki rolüne çevirmiştir. Özel koleksiyonların ve arşivlerdeki birincil kaynaklara verilen öncelik ile kütüphanelerde

nicelik ve sayısal veriler (koleksiyon sayıları gibi) yerine nitelik odaklı kütüphane değerlendirmesine yeniden odaklanılmıştır. Diğer bir deyişle özel koleksiyonların ve arşivlerin araştırma ve öğrenme gündemlerine nasıl katkıda bulunduğu daha fazla ilgi gösterilmeye başlanmıştır (Dempsey, Malpas ve Lavoie, 2014, s. 399).

Değişen kütüphanecilik hizmetleri anlayışında şüphesiz kurumsal bilgi depoları ve veri yönetiminde altmış yılı aşkın bir süredir, dijital veri yönetimi ve korunması, araştırmacılar, fon sağlayıcılar, altyapı ve hizmet sağlayıcılar dâhil olmak üzere diğer birçok paydaşın dâhil olduğu kütüphaneler ve arşivler gibi akademik kurumların misyonunun merkezinde yer almaktadır (Mokrane ve Parsons, 2014, s. 89). Bilimsel veri yönetimi, bilim camiasının içinde ve dışında, özellikle çağdaş açık bilim söyleminde giderek artan düzeyde ilgi görmektedir. Kurumlarda 'iyi' veri yönetimi uygulaması konusunda fikir birliği oluşmaya başlasa da araştırma verisi yönetiminde hâlâ yetersiz uygulamalar görülmektedir (Mokrane ve Parsons, 2014, s. 90). Bu konuda yapılabilecekler bir sonraki kısımda yer almaktadır.

4. Akademik Kütüphaneler ve Araştırma Kütüphaneleri: e-Bilim, Veri Kürasyonu ve e-Kütüphaneci

Peki, kütüphanecilik açısından e-bilim ne ifade etmektedir; akademik kütüphaneler ve araştırma kütüphanelerine ne katmıştır? Gelişmiş dijital teknolojilerin bilimsel araştırmalara entegrasyonu ile karakterize edilen e-bilim, akademik kütüphanelerin faaliyetlerini ve hizmetlerini derinden etkilemiştir. Bu evrim, kütüphaneleri modern araştırmanın dinamik ortamını etkili bir şekilde desteklemek için rollerini uyarlamaya ve genişletmeye teşvik etmiştir. e-bilimin akademik ve araştırma kütüphaneleri için katkıları şu şekilde değerlendirilebilir:

4.1. Veri Yönetimi ve Koruma

Dijital veri üretimindeki artış, akademik kütüphaneleri veri yönetimi ve korunmasında temel ortaklar olarak konumlandırmıştır. Kütüphaneciler kapsamlı veri yönetimi planları geliştirmek için araştırmacılarla iş birliği yaparak verilerin güvenli bir şekilde saklanması ve gelecekte kullanım için erişilebilir kalmasını sağlamaktadır. Bu iş birliğiyle bilimsel araştırmanın bütünlüğünü ve tekrarlanabilirliğini artırmak amaçlanmaktadır.

4.2. Siber Altyapının Geliştirilmesi

e-Bilimin taleplerini karşılamak için akademik kütüphaneler, kurumsal bilgi depoları ve dijital kütüphaneler de dâhil olmak üzere güçlü siber altyapıya yatırım yapmaktadır. Bu platformlar, büyük veri kümelerinin ve araştırma çıktılarının depolanmasını, geri getirilmesini ve paylaşılmasını kolaylaştırmakta, böylece işbirlikçi ve ağ bağlantılı bilimi desteklemektedir.

4.3. Eğitim ve Destek Hizmetleri

Araştırma topluluğunun gelişen ihtiyaçlarının farkına varan akademik kütüphaneler, rollerini dijital araçlar, veri okuryazarlığı ve e-araştırma metodolojileri eğitimini içerecek şekilde genişletmiştir. Kütüphaneciler artık araştırmacıların veri yoğun bilimin karmaşıklıklarında gezinmelerine yardımcı olmak ve böylece genel araştırma sürecini geliştirmek için atölye çalışmaları ve kişiselleştirilmiş destek sunmaktadır.

4.4. Açık Bilimin Savunuculuğu

Akademik kütüphaneler ve araştırma kütüphaneleri, araştırmada şeffaflığı ve erişilebilirliği teşvik ederek açık bilim ilkelerinin güçlü savunucuları olmuşlardır. Açık erişimli veri depolarını yöneterek ve araştırmacıların çalışmalarını açık bir şekilde yayınlamalarını destekleyerek bilginin yayılmasını kolaylaştırmaktadırlar. Ayrıca kütüphaneler, bilimsel araştırmalara halkın katılımını genişleterek ve bilgi üretme konusunda daha kapsayıcı bir yaklaşımı teşvik ederek vatandaş bilimi girişimlerinde yer almaktadır.

Diğer bir deyişle, e-bilim akademik ve araştırma kütüphanelerinde önemli bir dönüşümü hızlandırmış ve onları geleneksel bilgi düzenlemenin ötesine geçen yeni roller benimsemeye teşvik etmiştir. Veri yönetimi, altyapı geliştirme, araştırmacı desteği ve açık bilim savunuculuğu konularındaki sorumlulukları benimseyen kütüphaneler, e-bilimin getirdiği ilerlemelere uyum sağlayarak ve bunları

kolaylaştırarak araştırma ekosisteminin ayrılmaz ortakları hâline gelmiştir (Blixrud ve Rambo, 2007, s. 20).

Yeni nesil akademik çıktının, yayının ve araştırma verisinin kürasyonunu yapacak e-bilim kütüphaneciliğinin geliştirilmesi gerekmektedir. Dolayısıyla mesleki açıdan bu alanda yetkinleşecek kütüphanecilere ihtiyaç bulunmaktadır. Tıpkı istatistikten, programlamadan ve kodlamadan anlayan ve uygulayan araştırmacı profili gibi, veriyi anlamlandırmak ve aşağıda açıklanan ihtiyaçlara uygun hâlde kullanıma sunmak için aynı şekilde istatistikten ve veri türlerinden anlayan, programlama ve kodlama yapabilen e-kütüphaneciler yetiştirilmelidir. Araştırmacı olarak e-bilim ve bu alandaki kütüphaneciliğe ilk referansları yapanlardan biri, California Üniversitesi Los Angeles'da (UCLA) Bilgi Çalışmaları Profesörü ve Bölüm Başkanı Christine Borgman, e-bilimi ve ağa bağlı bilgi toplulukları fikrini kütüphanecilik mesleğinde gündeme getiren önemli bir akademisyendir (Borgman, 2004). 2007 yılında, Araştırma Kütüphaneleri Derneği (ARL) e-bilim iş gücü ile e-bilim ve kütüphanecilik hakkındaki raporunu yayımlamıştır. ARL'nin raporu, üye kütüphanelerini, 21. yüzyılda önem kazanacak bu alanı destekleyecek yeni araştırma destek stratejileri geliştirerek e-bilim ile ilgilenen araştırmacılarla ilişki kurmaya teşvik etmiştir (Blixrud and Rambo, 2007, s. 8-10). Tony ve Jessie Hey, kütüphane topluluğu için e-bilimi bir araştırma metodolojisi olarak nitelendirmiştir. Buna göre e-bilim kendi başına yeni bir bilimsel disiplin değildir: e-bilimi, iş birlikçi, ağ bağlantılı bilim olarak tanımlamaktadır (Hey ve Hey, 2006, s. 520).

Akademik kütüphanelerin e-bilim ile uğraşan araştırmacılara destek sağlama konusundaki ilgilerine ek olarak, ABD'de sağlık bilimleri kütüphanesi topluluğu, ağa bağlı araştırma iş birliklerinin bilgi ihtiyaçlarını desteklemek için kütüphaneci pozisyonları oluşturmanın önemli bir savunucusu olmuştur. NYU Sağlık Bilimleri Kütüphanesinin şu anki direktörü ve Washington Üniversitesi Sağlık Bilimleri Kütüphanesinin eski müdürü Neil Rambo, bu terimi kütüphanecilik mesleği açısından kullananlardan biridir. Rambo'nun e-bilim tanımı, bir araştırma ürünü olarak veri oluşturan potansiyel e-bilimin altını çizmektedir: "e-Bilim, ağa bağlı yetenekler ve büyük miktarda veriyi toplama ve saklama pratik olasılığıyla beslenen yeni bir araştırma metodolojisidir." Bu makaleye yanıt olarak, Massachusetts Üniversitesi Tıp Fakültesi Lamar Soutter Kütüphanesi ve New England Bölgesi Ulusal Tıp Kütüphaneleri Ağı, sağlık bilimleri kütüphanelerini mesleki becerileri belirlemek ve e-bilim kütüphanecilerini eğitmek için bir program geliştirmek için iş birliği yapmaya teşvik etmiştir (Martin ve Kafel, 2010, s. 5).

e-Bilim verisi üreten araştırmacıların ihtiyaçlarını karşılamak için geleneksel kütüphanecilik becerilerini yeniden yapılandıran kütüphanecilere bir örnek Witt ve Carlson'un tasarladığı, etkili veri yönetimi ve e-bilim hizmetleri sağlamak için geleneksel referans görüşmesinin bir "veri görüşmesine" uyarlamasıdır. Buna göre, e-bilim'e özgü veri setlerinin korunması amacıyla kaynak ve beklentileri anlamak için gerekli 10 pratik soru hazırlanmıştır. Bu sorular, aynı zamanda e-bilim'e yeni başlayan bir kütüphanecinin ihtiyaç duyduğu bazı eğitim araçlarının ve becerilerin anlaşılmasına da yardımcı olmaktadır. Verinin doğasını anlamak için sorulacak sorular şunlardır (Witt ve Carlson, 2007):

1. Bağlam ve kapsamı anlamak için verilerin hikâyesi nedir?
2. Veriler hangi biçim ve formattadır?
3. Veri kümesinin beklenen ömrü nedir?
4. Veriler nasıl kullanılabilir ve yeniden kullanılabilir?
5. Veri kümesi hangi büyüklüktedir ve büyüme oranı nedir?
6. Verilerin potansiyel kullanıcıları kimlerdir?
7. Verilerin sahibi kimdir?
8. Veri kümesi hassas bilgiler içeriyor mu?
9. Veriler, hangi yayınlardan veya keşiflerden ortaya çıkmıştır?
10. Veriler nasıl erişilebilir hâle getirilmelidir?

Bu açıdan, e-bilim hem deneysel hem de teorik gelenekleri birleştirmeyi hedeflerken bilgisayar simülasyonları olan yapay veriler oluşturabilir ve gerçek zamanlı büyük veriler teorik simülasyon modellerini kalibre etmek için kullanılabilir (Hilbert, 2015, s. 11). e-Bilim kavramsal olarak çok

miktarda hesaplama kaynağı kullanarak internet üzerinden erişilebilen büyük miktarda veriyi analiz ederek yeni bilimsel keşifler yapmak amacıyla bilimsel araştırmalar yürütmeye bilim insanlarını desteklemek için yeni yöntemler geliştirmeye yardımcı olur. Bununla birlikte, değer keşifleri yalnızca hesaplama araçları, bir siber altyapı sağlayarak veya bir sonuç üretmek için önceden tanımlanmış bir dizi adım uygulayarak yapılmaz. Aksine, faaliyetin doğası gereği otomatikleştirilemeyen orijinal, üretken bir yönü olmalıdır. Bu, e-bilim platformlarının bilim yapmanın yeni bir paradigmasını desteklemek için sağlaması gereken özellikleri ve hesaplamalı veri sonuçlarını koruma ve kullanıma sunma gerekliliklerini yerine getirmek için aşağıda tartışılan yeni kuralları tanımlamaya çalışan çeşitli araştırmalara yol açmıştır (Hilbert, 2015, s. 14-15).

5. COAR: Yeni Nesil Kurumsal Bilgi Depoları Üzerine Çalışmalar

Açık erişim açısından kurumsal bilgi depoları camiasında önde gelen konfederasyonlardan biri olan COAR için öncelikli olan stratejik hedef, Nisan 2016'da başlatılan Bilimsel İletişim (Scholarly Communication) Çalışma Grubu kapsamında öncelikle gelecek nesil arşivler için işlevsel ve teknolojik mimarileri belirlemektir. Ayrıca, Yeni Nesil Kurumsal Bilgi Depoları (Next Generation Institutional Repositories) açık içeriğe engelsiz erişime izin veren ve arşivler arası katma değerli hizmetlerin oluşturulmasını teşvik eden küresel bir depo ağı geliştirmeyi hedeflemiştir. Yeni Nesil Kurumsal Bilgi Depoları, COAR'ın, bu alanda yeni ve geliştirilmiş ortak davranışları, protokolleri ve teknolojileri tanımlamak için başlatılmış bir girişimdir.

COAR'ın Yeni Nesil Kurumsal Bilgi Depoları anlayışı kapsamında Üstveri Çalışma Grubunda bu arşivlerin içerik ve veri mimarisindeki kriterlerini belirleyen şu standartlar incelenmiştir:

- Data citation roadmap for scholarly data repositories (Bilimsel veri depoları için veri alıntı yol haritası) (Fenner, Crosas, Grethe, 2019, s. 1-9),
- Core Trust Seal (CTS, 2019),
- FAIR Data Principles (Wilkinson, Dumontier, Jan, Appleton ve Axton, 2016),
- PLOS "Criteria that Matter" (MacCallum, 2019),
- TRUST Principles for Digital Repositories (Lin, 2020),
- COAR Next Generation Repositories Technologies (COAR 2020),
- Plan S (European Science Foundation, 2018).

Kurumsal bilgi depolarını ve araştırma verilerini yöneten uygulayıcılar olarak COAR Üstveri Kalitesi Çalışma Grubu (COAR Metadata Quality Working Group) üyelerinden, 2020 başından itibaren kendi çalıştıkları kurumların ve konsorsiyumların belirlediği kriterler konusunda geri bildirimler ve bunların uygulanabilirlikleri hakkında aşağıdaki başlıklar temel alınarak görüş bildirilmesi istenmiştir. Buna göre Çalışma Grubunun katılımcılarının değerlendirdiği unsurlar şu şekildedir:

- 1) Benimseme kolaylığı ne düzeyde ve temel özelliklerden herhangi birinin uygulanması çok zor mu?
- 2) Prensiplerde boşluklar ve önemli kriterlerde eksiklikler var mı?
- 3) Netlik açısından bu kriterlerden herhangi birinin anlaşılması veya yorumlanması zor mu?
- 4) Bu kriterlerden herhangi biri ilişkisellik açısından uygun mu?

İncelemeler sonucunda yeni nesil arşivlerin temel özellikleri ile olması arzu edilen (ideal) özellikleri, 8 özellik, iki kategori olarak şu şekilde belirlenmiştir. "Temel özellikler" asgari veya olması gereken şartları, "istenen özellikler" ise olması gereken veya ileride geliştirilmesi önerilen özellikler olarak iki alt kategoride değerlendirilmiştir.

5.1. Keşfedilebilirlik

Temel Özellikler: Arşiv, yüksek kaliteli üstverileri destekler (disiplin tabanlı, bölgesel veya genel üstveri şeması olarak: örneğin, Dublin Core). Arşiv, OAI-PMH kullanarak üstverileri harmanlar. Arşivdeki üstveriler, kaynak artık mevcut olmadığı durumlarda bile erişilebilir olmalıdır. Arşiv, kaynağın açılış sayfasına (kaynağın mevcut olmadığı durumlarda bile) alıntı yapılabilir. Kalıcı bir tanımlayıcı (permanent identifier [PID] atar ve ilgili kaynakların harici olarak barındırılan sürümleri barındırır (yayınlanmış makaleye ait bağlantılar gibi). Arşiv, bir arama imkânı veya arşiv üstverilerini sunar ve/veya harici bir keşif hizmeti aracılığıyla indekslenir. Arşiv, bir veya daha fazla disiplin veya

genel kaynak veri tabanlarında listelenir (Re3data, OpenDOAR veya diğer ulusal, bölgesel veya yerel veri tabanı listeleri gibi).

İstenilen Özellikler: Arşiv, uygun olduğu durumlarda birden çok ayrıntı düzeyindeki veriler için kalıcı tanımlayıcıları destekler. Arşiv, makaleler, veriler ve yazılım gibi ilgili malzemeler arasında bağlantı kurar. İçerik uzlaşması seçenekleri hakkında bilgi vermek için http bağlantı başlıkları oluşturulur. Arşiv yazarlar, fon sağlayıcılar, fon programları ve hibeler, kurumlar ve diğer ilgili kuruluşlar için PID'ler için destek sağlar. Üstveriler, standart bir bibliyografik formatta indirilebilir olmalıdır.

5.2. Erişim

Temel Özellikler: Arşivde yayımlanan verilere erişmenin kullanıcıya herhangi bir maliyeti yoktur. Arşiv, kamuya açık olarak belirtilen bir zaman süresince verilere sürekli erişim sağlar.

İstenilen Özellikler: Arşiv, engelliler için de kayıtlara ve belgelere erişimi sağlar. Arşiv, büyük veri kümelerini kullanıcılara aktarmak için bir mekanizma sağlar.

5.3. Yeniden Kullanım

Temel Özellikler: Arşiv, Creative Commons lisanslarının kullanımını destekler. Arşiv, veri atfı kullanımı için gerekli dokümantasyonu ve teknik alt yapıları sunar. Burada atfı yapılabilen PID'ler kullanılabilir. PID kendisine referans yapılabilen bir kalıcı tanımlayıcıdır ve tanımladığı kaynağa özgü bir açılış sayfasına çözümleyen bir URL olarak ifade edilebilir. Açılış sayfası, kaynağı açıklayan makine tarafından okunabilir üstverileri içerir (PID services DOI, ePIC, 2024).

İstenilen Özellikler: Arşiv, tam metin harmanlamayı ve/veya metin ve veri madenciliğini desteklemek için açık programlama arayüzlerine (API'lere) sahiptir. Arşiv, tercih edilen, tescilli olmayan formatların bir listesini sunar. Açılış sayfaları, alıntı için gerekli üstverileri ve ideal olarak, insan tarafından okunabilir (lisans ve PID de dâhil olmak üzere) ve makine tarafından okunabilir biçimde keşfi kolaylaştıran üstverileri içerir. Kaynaklar, makine tarafından okunabilen topluluk standardı formatlarında saklanır.

5.4. Bütünlük ve Özgünlük

Temel Özellikler: Arşiv, kaynakların yetkisiz erişimini önleyen uygulamaların dokümantasyonunu içerir. Arşiv, arşivdeki üstveriler ve kaynaklardaki tüm değişikliklerin bir kaydını tutar. Arşiv, araştırmacıdan alındıktan sonra üstverilerin ve kaynakların sürümlendirilmesini, versiyon kayıtlarının tutulması ve sunulmasını destekler.

İstenilen Özellikler: Arşiv, kaynaktan sorumlu kişi(ler) veya kurum(lar)ın isimleri dâhil olmak üzere içerik sağlayıcı(lar) hakkında bilgi sağlar.

5.5. Kalite Güvencesi

Temel Özellikler: Arşiv, kaynaklara ve üstverilere hangi iyileştirme işlemlerinin uygulandığına ilişkin kayıtlar sağlar.

İstenilen Özellikler: Arşiv, üstverilerin ve verilerin temel seçimini ve küratörlüğünü üstlenir. Arşiv, kaynak ve üstverilere açıklama, yorum veya inceleme yapılmasını destekler.

5.6. Hassas Verilerin Gizliliği (Özellikle İnsan Deneklerin Kullanıldığı Durumlarda)

Temel Özellikler: Arşivin hassas araştırma verilerini topladığı durumlarda, veri sahiplerinin erişimi yalnızca yetkili kullanıcılarla sınırlamasına izin veren mekanizmalara sahiptir.

İstenilen Özellikler: Arşivin hassas araştırma verilerini topladığı durumlarda, verilerin farklı güvenlik gereksinimleri düzeylerine göre katmanlı erişim sağlanır.

5.7. Sürdürülebilirlik ve Koruma

Temel Özellikler: Arşiv (veya arşivi yöneten kuruluş), arşivi yönetmek ve finanse etmek için uzun vadeli bir plana sahiptir. Arşiv, kaynakların uzun vadede yönetileceği süreyi ve koruma uygulamaları ile ilgili belgeleri tanımlayan bir genel sağlama politikasına sahiptir.

İstenilen Özellikler: Arşiv, yaygın olarak kabul gören koruma uygulamalarını benimseyen, korumaya yönelik belgelenmiş bir yaklaşıma sahiptir. Verisini veya araştırma çıktısını kurumsal bilgi deposuna hibe edenle arşiv arasındaki anlaşma, koruma sorumluluklarını yerine getirmek için gerekli tüm eylemleri içermelidir (öğeleri kopyalama, dönüştürme ve saklama hakları gibi). Arşivleme, malzemeyi alımdan yaymaya kadar tanımlanan iş akışlarına göre gerçekleşir.

5.8. Diğer Konular

Temel Özellikler: Arşiv, hak sahiplerine ve kullanıcılara yardımcı olmak için bir iletişim noktası oluşturur. Arşiv, sorulara makul bir süre içinde yanıt verir. Arşiv, arşive kabul edilen içeriğin kapsamı hakkında belgeler sunar.

İstenilen Özellikler: Arşiv, standart bir yöntem kullanarak kullanım bilgilerini toplar ve paylaşır (indirme ve sayfa görüntüleme istatistikleri gibi). İyi desteklenen işletim sistemlerinde ve diğer temel altyapı yazılımlarında depo işlevleri ve web hazırlığı dâhil olmak üzere hizmetlere uygun donanım ve yazılım teknolojilerini kullanır. Emanetçi tarafından yapılan gönderme sistemi hem bireysel içerik oluşturucu yüklemelerini hem de kayıtların toplu yüklemelerini destekler.

COAR Üstveri Kalitesi Çalışma Grubuna katılanların ortak yeni nesil kütüphaneler için COAR'ın "Yeni Nesil Kurumsal Bilgi Depoları" raporu, vizyonu ileriye taşımak için yeni uygulamalar ve teknolojiler için öneriler sunmaktadır. Bu vizyonun önemli bir bileşeni, arşivlerin çok çeşitli araştırma çıktılarına erişim sağlayacağı, bilimsel kayıtlara daha fazla katkı çeşitliliğinin erişilebilir olacağı ve araştırma değerlendirme süreçlerinde resmi olarak tanınacağı koşulları oluşturmaktır. Bu durum, 21. yüzyılda e-bilim ile birlikte geleneksel bilim yapma alışkanlıklarının değişmesiyle ortaya çıkan; kütüphanelerin ve kütüphanecilerin yeni roller üstlendiği, kaynak ve hizmet anlayışının dönüşüme uğradığı bir süreci tanımlar. Bu süreç, siber altyapının güçlendirilmesini gerektiren, ortak ağlar üzerinden yürütülen çok kurumlu ve çok paydaşlı çalışma ortamlarının desteklenmesini içerir. Aynı zamanda, üretilen verinin yalnızca depolanıp paylaşılmasıyla kalmayıp, işlenerek kullanılabilir hâle getirilmesini ve böylece bilgi döngüsüne entegre edilmesini amaçlayan stratejik yaklaşımlarla da büyük ölçüde örtüşmektedir. Bu vizyon, yalnızca yerel kullanıcıları için içerik satın alan kütüphanelerden, kurumlarında üretilen araştırma çıktılarının küratörlüğünü üstlenen; bilimi üreten ve bunu küresel ölçekte açık bilim ilkeleri doğrultusunda paylaşan kütüphanelere doğru bir dönüşümü ifade etmektedir. Ancak bu vizyona ulaşabilmek için, kurumsal açık erişim bilgi depolarında ticari eşdeğerleriyle rekabet edebilecek düzeyde hizmetler sunulmalı; bunun için de yeni işlevler ve teknolojiler benimsenmeli, ayrıca bu yapılar üzerinde standart kullanım ölçümleri, veri kalitesi denetimi ve sosyal ağ entegrasyonu gibi ek hizmetler geliştirilmelidir.

6. Yeni Nesil Kurumsal Bilgi Depolarının İlkeleri: FAIR

Yukarıda belirtilen yeni nesil arşivlerde uygulanan kriterleri içeren standartlar ve çalışmalar içinde en önemlileri şüphesiz FAIR ilkeleri altında özetlenmiştir. FAIR Veri İlkeleri (Wilkinson, Dumontier, Jan, Appleton, Axton, Baak, Blomberg ve Mons, 2019), verilerin insanlar ve makineler tarafından yeniden kullanılabilmesini sağlamak için veri nesnelerinin temel özelliklerini tanımlayarak iyi uygulamayı benimseme ihtiyacını vurgulamaktadır: Bunlar **findable**/bulunabilir, **accessible**/erişilebilir, **interoperable**/birlikte çalışabilir ve **re-usable**/yeniden kullanılabilir olmalıdır.

Ancak, verileri zaman içinde korurken FAIR ilkelerini uygulamak için sürdürülebilir yönetim ve kurumsal çerçeveler, güvenilir altyapı ve topluluk tarafından kabul edilen uygulamaları destekleyen kapsamlı politikalar ile güvenilir dijital arşivler (*Trusted Digital Repositories* [TDR]) gerekmektedir. TDR'ler, hem teknolojiye hem de paydaş gereksinimlerindeki değişikliklere yanıt olarak verileri etkin bir şekilde korumaya yönelik açık sorumluluklarıyla, verilerin değerinin korunmasında önemli bir rol oynamaktadır. Veri yönetiminin sorumluluklarını kabul ettikleri için kullanıcıları tarafından güven konumunda tutulurlar. Bu rolü yerine getirmek için TDR'lerin, hizmet verdikleri topluluklar için zaman

içinde verilere erişimini ve yeniden kullanımını sağlamak için gerekli ve kalıcı yetenekleri göstermesi gerekir. TDR'ler, farklı yeniden kullanılabilirlik düzeyleriyle veri biriktirme ve veri muhafazalarının korunmasını destekler. Bazı durumlarda, makul bir şekilde iyileştirilemeyen veya daha birlikte çalışabilir hâle getirilemeyen düşük kaliteli veriler, kullanıcı topluluğu için hâlâ yüksek değeri koruyabilir ve bu nedenle güvenilir bir yönetim gerektirebilir. Bir TDR, topluluk tarafından kabul edilen kriterleri karşılamaya çalışmalı ve ulaşılan veri kalitesi seviyesini topluluğa bildirmelidir.

Yapılan çalışmalara göre FAIR prensipleri, birçok kurumsal bilgi deposu için uygun veri yönetiminde bir ön şart olarak belirlenmiştir. FAIR Kılavuz İlkeleri kapsamındaki fikirler, makine hareketliliği ve veri yapıları ile semantiğin uyumlaştırılmasına odaklanan Web Kavram Konsorsiyumu (Concept Web Consortium) ortaklarının önceki çalışmalarını yansıtır, birleştirir, geliştirir ve genişletir. Veri Atıf İlkeleri Ortak Bildirisi'ni (JDDCP29) geliştiren bilimsel kuruluşlar, kendi bünyelerindeki açık erişim platformlarını birincil akademik verilere atıf yapılabilir, keşfedilebilir ve yeniden kullanıma açık hâle getirilmesine odaklanmıştır. FAIR İlkeleri ve JDDCP arasındaki benzerlikleri ve örtüşmeleri tanımlama girişimine web sayfasından erişilmektedir (FORCE11, 2019). FAIR İlkeleri ile birlikte “Veri Onay Mührü” (Data Seal of Approval [DSA]) gibi veri üreticilerine verilerinin ve ilgili materyallerinin güvenilir bir şekilde depolanacağı ve yeniden kullanılabilirliği konusunda güvence veren sertifikalara ihtiyaç duyulmaktadır (Leeuw, 2019). DSA, verileri oluşturanların dışındaki kullanıcılar için yeniden kullanılabilir hâle getirirken, öncelikle veri üreticilerinin ve depoların sorumluluklarına ve davranışlarına odaklanır. Öte yandan FAIR ise öncelikle verinin kendisini odağına almaktadır.

FAIR "İlkeleri" aşağıdaki tablo ile gösterilebilir:

Tablo 1

FAIR—Findable, Accessible, Interoperable, Reusable (Bulunabilir, Erişilebilir, Birlikte Çalışabilir, Yeniden Kullanılabilir)²

“F” Bulunabilir olmak için (Findable): F1. Üstverilere küresel olarak benzersiz ve kalıcı bir tanımlayıcı atanır. F2. Veriler zengin üstverilerle açıklanmıştır. (Aşağıda R1 ile tanımlanmıştır) F3. Üstveriler açık ve net bir şekilde tanımladığı verilerin tanımlayıcısını içerir. F4. Üstveriler aranabilir bir kaynağa kaydedilir veya indekslenir.	“A” Erişilebilir olmak için (Accessible): A1. Üstveriler, standartlaştırılmış bir iletişim protokolü kullanılarak tanımlayıcılar tarafından erişilebilir. A1.1 Protokol açık, ücretsiz ve evrensel olarak uygulanabilir. A1.2 Protokol, gerektiğinde bir kimlik doğrulama ve yetkilendirme prosedürüne izin verir. A2. Veriler artık mevcut olmasa bile üstverilere erişilebilir.
“I” Birlikte çalışabilir olmak için (Interoperable): I1. Üstveriler, bilgi temsili için resmi, erişilebilir, paylaşılan ve genel olarak uygulanabilir bir dil kullanır. I2. Üstveri, FAIR ilkelerini takip eden kelimeleri kullanır. I3. Veriler, diğer üstverilere nitelikli referanslar içerir.	“R” Yeniden kullanılabilir olmak için: (Reusable): R1. Veriler ve üstveriler çok sayıda doğru ve alakalı özelliklerle zengin bir şekilde tanımlanmıştır. R1.1. Üstveriler, açık ve erişilebilir bir veri kullanım lisansı ile yayınlanır. R1.2. Üstveriler ayrıntılı kaynakla ilişkilendirilir. R1.3. Üstveriler alanla ilgili topluluk standartlarını karşılar.

² Kaynak olarak Wilkinson, M. D., vd. (2016). *The FAIR guiding principles for scientific data management and stewardship*. Science Data, 3(160018), s. 4 bilgi kullanılmıştır.

6.1. FAIR Özelliklerini Destekleyen ve Türkiye’de Uygulanabilecek Hizmetler

Yeni nesil kurumsal bilgi depoları, yalnızca açık erişimli bilimsel yayınların sunulduğu platformlar olmaktan çıkarak daha kapsamlı bilgi ve veri yönetim sistemlerine dönüşmeye başlamıştır. Artık DSpace gibi platformlar, Bilim 2.0’ın ortaya çıkardığı büyük ölçekli araştırma verileri ve karmaşık veri kümelerini işleyip kullanıma sunma konusunda yetersiz kalmaktadır. Bu nedenle, 2010 yılından itibaren bilimsel veri ve bilgi yönetiminin, araştırma verilerini de kapsayan büyük veri kümelerine yönelik olarak evrildiği gözlemlenmektedir.

Özellikle üretken yapay zekâ teknolojileriyle desteklenmeye başlanan bilgi ve araştırma verisi yönetimi yaklaşımları; verinin üretiminden analizine ve bilim dünyasına sunumuna kadar tüm süreçleri kapsayan, daha donanımlı siber altyapıları destekleyen yeni nesil platformların ortaya çıkmasına neden olmuştur. Bu platformlar içinde öne çıkanlar arasında Dataverse, FAIRDOM, ISA, Open PHACTS, wwPDB ve UniProt gibi örnekler yer almaktadır.

Yapay zekâ destekli bu platformların, yeni nesil kurumsal bilgi depoları kapsamında araştırma verilerini ne ölçüde desteklediği ve dijital varlık yönetimi kapasiteleri açısından FAIR ilkelerini nasıl karşıladıkları aşağıda açıklanmaktadır:

6.1.1. Dataverse

Dataverse (Crosas, 2011), halka açık topluluk kurumsal bilgi depoları veya kurumsal araştırma verisi depolarını desteklemek için pek çok kurumda kurulu açık kaynaklı bir veri depolama yazılımıdır. 60.000’den fazla veri kümesiyle Harvard Dataverse, mevcut Dataverse depolarının en büyüğüdür ve tüm araştırma alanlarından tüm araştırmacılara açıktır. Dataverse, dijital nesne tanımlayıcıyı (Digital Object Identifier [DOI]) veya diğer kalıcı tanımlayıcıları (handle), veri kümesi yayımlandığında herkese açık hâle getirir. Bunun için tümü dizine eklenmiş ve aranabilir üstverilere, veri dosyalarına, veri kümesi şartlarına, feragatlere veya lisanslara ve sürüm bilgilerine erişim sağlayan bir web sayfasına yönlendirme yapılabilir. “Yeniden kullanılabilirlik” adına verilerin/makalelerin üstverilerini, veri dosyalarını ve veri analizlerini anlamak için gereken tamamlayıcı dosyaları (belgeler veya kod gibi) içerir. Gizlilik sorunları nedeniyle veriler kısıtlanırsa veya kaldırılrsa bile üstveriler her zaman herkese açıktır. Bu üstveriler, “Birlikte çalışabilirlik” ve “Yeniden kullanılabilirlik” FAIR ilkelerini kapsamlı bir şekilde destekleyen üç düzeyde sunulmaktadır: 1) DataCite şeması veya Dublincore Temel Koşulları ile eşleşen veri alıntı üstverileri; 2) mümkün olduğunda üstverilerle eşlenen alana özgü üstveriler ve bilimsel bir alanda kullanılan standartlar; 3) tablo veri dosyaları için (sütun düzeyinde üstveriler dâhil) kapsamlı olabilen dosya düzeyinde üstveriler. Son olarak, Dataverse, veri dosyaları kısıtlandığında erişim sağlamak için bir belirteç kullanarak verileri aramak, üstverilere erişmek ve veri dosyalarını indirmek için herkese açık makine erişimli arayüzler sağlar.

6.1.2. Fairdom

FAIRDOM, sistem biyolojisi için bir FAIR veri ve model yönetimi altyapısı üretmek için SEEK (Wolstencroft, 2015) ve openBIS (Bauch, 2011) platformlarını entegre etmektedir. Bireysel araştırma varlıkları (veya veri ve model kümeleri), yayınlanmak üzere DOI’lerle kaydedilebilen benzersiz ve kalıcı HTTP URL’leri ile tanımlanır. Araştırma varlıklarına, bireyler ve/veya bilgisayarlar için uygun çeşitli formatlarda (RDF, XML) web üzerinden erişilebilir. Araştırma varlıkları, topluluk standartları, biçimleri ve ontolojileri kullanılarak zengin üstverilerle açıklanır. Üstveriler, birlikte çalışabilirliği sağlamak için RDF olarak depolanır ve varlıklar yeniden kullanım için indirilebilir.

6.1.3. ISA

ISA (Sansone, 2012) doğa bilimleri veri kümelerinin standartlara uygun olarak toplanmasını, iyileştirilmesini, yönetilmesini ve yeniden kullanılmasını kolaylaştırmak için topluluk odaklı bir üstveri izleme aracıdır. ISA, Nature Scientific Data’nın veri tanımlayıcı makalelerine ve birçok GigaScience veri belgesine aşamalı olarak FAIR yapılandırılmış üstveriler sağlar ve diğer veri kaynaklarının yanı sıra EBI MetaboLights veritabanının temelini oluşturur (González-Beltrán, vd., 2014, s. 5). Özünde, başlangıçta yalnızca tablo şeklinde sunulan, ancak daha sonra RDF tabanlı bir gösterim olarak

geliştirilmiş olan, genel amaçlı, genişletilebilir bir ISA modeli, birlikte çalışabilirlik ve yeniden kullanılabilirlik özelliklerini etkinleştirmek için JSON uygulamalarıyla, yayımlandığında FAIR özelliklerine sahip hale gelir (González-Beltrán, vd., 2015, s. 4).

6.1.4. Open PHACTS

Open PHACTS (Harland, 2012, s. 2-3), ilaç keşfiyle ilgili bilgiler için bir veri entegrasyon platformudur. Platforma erişim, hem insan (HTML) hem de makine tarafından okunabilen (RDF, JSON, XML, CSV, vb.) (Groth, vd., 2014, s. 14) ve FAIR kriterlerinden erişilebilirlik özelliğine sahip, makine tarafından erişilebilir bir arayüz aracılığıyla sağlanır. Arayüz, bir eşleme hizmeti aracılığıyla belirli bir varlık hakkındaki bilgilere erişmek için birden çok URL'nin kullanılmasına izin verir. Bu nedenle, bir kullanıcı, örneğin Chemspider veya DrugBank kaynaklı bilgileri almak için bir ChEMBL URL'si sağlayabilir. Her arama, yanıtında standart bir URL sağlar. Kullanılan tüm veri kaynakları, zengin kaynak ile küresel VoID standardını takip ederek standartlaştırılmış veri kümesi açıklamaları kullanılarak açıklanmıştır. Tüm arayüz özellikleri, bağlantılı veri API özelliğine göre RDF kullanılarak çözümlenmiştir. Ayrıca, veri kümelerinin çoğu ontolojiler üzerinde mutabık kalınan topluluk uygulamaları ışığında tanımlanmıştır.

6.1.5. wwPDB

wwPDB (Berman, Henrick ve Nakamura, 2000, s. 237-240), proteinlerin ve nükleik asitlerin deneysel olarak belirlenen 3B yapıları hakkında bilgi barındıran, özel amaçlı, yoğun bir şekilde küratörlüğü yapılan bir veri arşividir. Tüm wwPDB girişleri bir FTP sunucusunda sabit bir şekilde tutulur ve makine tarafından okunabilir formatlarda (metin ve XML) temsil edilir. İkincisi, Uluslararası Kristalografi Birliği'nin (International Union of Crystallography [IUCr]) bir veri standardı olan Makromoleküler Bilgi Çerçevesi'ne (mmCIF) (Bourne, vd. 1997, s. 571-580) uygun olarak wwPDB tarafından sağlanan üstveriler kullanılarak IUCr için duyarlı makinelerle çalışabilir. WwPDB üstverileri, PubMed ve NCBI Taksonomisi gibi yaygın tanımlayıcılara çapraz referanslar içerir ve bunların wwPDB üstverileri, veri sözlüklerinde ve şema belgelerinde açıklanır. Kimyasal ve yapısal biyoloji alanları ve IUCr veri standard için uygundur. Hem wwPDB verilerini hem de üstverileri yorumlamak için çeşitli yazılım araçları mevcuttur. Her giriş bir DOI ile temsil edilir. DOI, daha fazla sorgulama/yorumlama için özel yazılım gerektiren sıkıştırılmış bir dosyaya çözümlenir. Diğer wwPDB erişim noktaları (Rose, vd., 2015, s. D345-D356), uzun vadede kararlı olma olasılığı yüksek ("Bulunabilir/Findable") URL'ler aracılığıyla wwPDB kayıtlarına erişim sağlar ve tüm veriler ve üstveriler, wwPDB'ye bağlı web sitelerinden biri veya daha fazlası aracılığıyla aranabilir.

6.1.6. UniProt

UniProt (UniProt Consortium, 2015), protein dizisi ve açıklama verileri için kapsamlı bir kaynaktır. Tüm girişler, bir web sayfası, düz metin ve RDF dâhil olmak üzere çeşitli biçimlerde erişim sağlayan sabit bir URL ile benzersiz şekilde tanımlanır. 150'den fazla farklı veri tabanı ile bağlantı kuran her UniProt kaydının kapsamlı bağlantıları vardır ve zengin alıntı yapılmasını sağlar. Bu bağlantılar, RDF gösteriminde makinede işlem görebilir. RDF gösteriminde, UniProt Core Ontology ile verilerin neyi temsil ettiği hakkında ne insanlar ne de makineler için belirsizlik bırakmadan tüm kayıtları açık bir şekilde gösterir ve kayıtların tam otomatik olarak alınmasını ve çapraz referans vermeyi sağlar.

7. COAR Üstveri Kalitesi Çalışma Grubunun Değerlendirmeleri ve FAIR İlkeleri

7.1. Üstveri Kullanımında Standartlaşma: Birlikte Çalışabilirlik ve Bulunabilirlik

COAR Üstveri Kalitesi Çalışma Grubunda dile getirilen konuların başında, kurumların/kütüphane yönetimlerinin kendi kurumsal ihtiyaçlarına göre üstveri alanları oluşturmaları ve bu nedenle de harmanlama durumunda *Birlikte çalışabilirlik* konularında sorunlar olması gelmektedir. Kurumsal bilgi depoları araştırmacının kendi araştırma verisini ve yayını eklemesi durumunda üstveri alanlarının yeterli bir şekilde doldurulmasının, kapsamlı ve doğru olmasının verinin bulunabilirlik kriteri açısından önemli olduğu, o nedenle de araştırmacı/akademisyenin de bilinçlendirilmesinin gerekliliği dile getirilmiştir.

7.2. Veri, Veri Setleri ve Yayınların Bağlantılı Veri (Linked Data) ile İlişkilendirilmesi: Bulunabilirlik ve Erişilebilirlik

Bazen kaliteli bir üstveri şemasına ve alanlarda uygun verilere sahip olmak, arşivdeki öğelerin kolayca bulunabileceği anlamına gelmemektedir. Bir araştırmacı gözünden bakılması ve onların neyi nasıl araştıracağı ve neyi bulmak isteyeceğinin de göz önüne alınması gereklidir. Diğer bir deyişle Bulunabilirlik kriterinin sağlanabilmesi için üstverinin kalitesi ve yeterliliğinin yanı sıra, yayımlanmış bir makale ve ilgili veri kümesi bağlantılarının sağlanması, arka planda bağlantılı veri (linked data) bağlantıların iyi yapılmasıyla ilgili materyallerin de keşfedilebilir olması gerekmektedir. Bazen entelektüel çıktıların ilgili veri tabanlarında veya Google Scholar'da harmanlanması, etki faktörünün çeşitli araçlarla görüntülenmesi ile istatistiki bilgiler de araştırmacının ilgisini çeken bağlantılı verilere dâhil olabilir. O nedenle kurumsal bilgi depolama uygulayıcıları, bir araştırmacının araştırma alışkanlıklarını (araştırma davranışını ve aradıkları nesne) göz önünde bulundurursa ve ilgili öğeleri uygun bir şekilde bağlarsa (araştırma verileri veri setleri, araştırma makalesi, teknik raporlar, vb.), kurumsal bilgi depoları daha etkin çalışabilir ve tam bir Erişilebilirlik özelliği sağlanabilir. Diğer bir deyişle, bir kullanıcı gibi düşünmek ve onların arama yöntemlerini anlamak, düzgün işleyen üstverilerin anahtarıdır. Bu şekilde kurumsal bilgi depoları sadece araştırma çıktısı depolayan ve görüntüleyen alanlar olmaktan çıkmaktadır. Araştırmacı için de kullanışlılığının göz önüne alınması gerekmektedir.

7.2.1. Araştırmacı Profilleri ve İlgili Materyallerin/Yayınların Bulunabilir Olması: Bulunabilirlik/Findability ve Erişilebilirlik/Accessibility

Bir kullanıcı sadece veri veya araştırma çıktısına ulaşmakla yetinmeyebilir. Bir araştırma ögesi bulunduğu, araştırmacının içerikle tam etkileşim kurabilmesi, içeriği derinlemesine anlayabilmesi ve ilgili araştırma çıktıları hakkında daha fazla bilgi edinebilmesi açısından; hangi araştırmacıyla iletişime geçeceğini bilmesi, konuya ilişkin diğer bilgilere ulaşabilmesi veya fiziksel materyalin nerede bulunabileceğini öğrenmesi, bağlam oluşturma açısından büyük önem taşır. Bir topluluk kullanıcısı olarak, araştırmacı tercih ettiği arama motorunda/kurumsal bilgi deposunda/veritabanında diğer arama sonuçlarını görmek isteyebilir. Böylece kendi ihtiyacı olan araştırma çıktılarını kendi arama bağlamında kolayca bulabilir.

7.2.2. Benzersiz Tanımlayıcı Uygulanması: Bulunabilirlik/Findability ve Birlikte Çalışabilirlik/Interoperability

Hem Bulunabilirlik anlamında kurumsal bilgi deposunun içinde bulunabilmesi hem de diğer veri tabanlarında ve Google Scholar'da harmanlanabilmesi için Birlikte çalışabilirlik adına benzersiz tanımlayıcı araçlarının kullanılması önemlidir. Ancak kurumların/kütüphanelerin hangi PID'leri uyguladığı ve bunları nasıl kullandığı konusunda çeşitli sorunlar yaşandığı anlaşılmaktadır. Yazar/akademisyen/araştırmacı tanımlaması için kullanılan ORCID³, Funder ID (fon sağlayıcısı için benzersiz tanımlayıcı), ISNI (International Standard Name Identifier/Uluslararası Standart Ad Tanımlayıcısı) gibi küresel tanımlayıcıların yanı sıra, Türkiye'de YÖKSİS ID gibi tanımlayıcılar kullanılmaktadır. Bunlar bazı kurumsal bilgi depolarında üstveride yazar adıyla birlikte isim otoriteleri olarak, bazılarında ise ayrı bir alanda gösterilmektedir.

7.2.3. Araştırmacının Kullanımı ve Çıktıların Diğer Arama Araçlarında Erişilebilirliği: Bulunabilirlik, Birlikte Çalışabilirlik, Yeniden Kullanılabilirlik

Bir araştırmacının beklentisi, araştırma veri setlerini hem fon sağlayıcıların hem de yayıncıların veri paylaşım politikalarına uygun şekilde saklamaktır. Öte yandan, bir veri analitiği uzmanı açısından beklenti; veriyi daha sonra yeniden kullanmak, başkalarıyla farklı projelerde iş birliği yapabilmek amacıyla ilgili kodları, planları ve istatistiksel analizleri derleyerek bu ihtiyacı karşılamaktır. Aynı zamanda kurumsal bilgi depolarında üstverilerin kütüphane sistemindeki diğer kaynaklar içinde (örn. Katalogda) aranabildiğinden emin olmak istenebilmektedir. Bunu sağlayan tek noktadan aradıklarına ulaştıkları sistemlere yönelmek araştırmacı açısından zaman kazandırıcı bir yöntemdir.

³ Yazar/araştırmacı için bilimsel ve diğer akademik yazarları ve katkıda bulunanları benzersiz bir şekilde tanımlamak için özel olmayan alfanümerik kod.

7.2.4. Telif Hakları ve Diğer Haklar: Yeniden Kullanılabilirlik/Reusability

Bir araştırmacının kurumsal bilgi depolarındaki her türlü çıktının yeniden kullanıp kullanamayacağını belirleyebilmesi için bir öğeyle ilgili izinlerin veya lisans bilgilerinin sağlanması gerekmektedir. Bu bilgi, ticari amaçlarla yeniden kullanımı, öğretimde yeniden kullanmayı, yeniden yayımlama hakkındaki izinleri içermektedir. Bazı ülkelerde yeniden kullanım açık ve net iken, Türkiye’de bu konudaki hükümler ve kanunlar yetersiz kalmaktadır. O nedenle Türkiye’de de Creative Commons gibi uluslararası ölçekte kabul görmüş uygulamaların yasallaşması veya bu konuda daha net hukuki çerçeveler çizilmesine ihtiyaç vardır.

7.3. e-Bilim İçin Altyapıların Oluşturulması: Bulunabilirlik/Findability, Erişilebilirlik/Accessibility, Birlikte Çalışabilirlik/Interoperability, Yeniden Kullanılabilirlik/Reusability

Pahalı ve uzun vadeli planlama gerektiren, büyük boyutlu ve dağıtık bilişim teknolojilerinin kullanıldığı veri işleme, yeniden kullanma, paylaşma, yedekleme ve saklama süreçleri için gerekli siber altyapının sağlanması bir zorunluluk hâline gelmektedir. Bu bağlamda, çok paydaşlı kurumları kapsayan konsorsiyumlarla çalışmak kaçınılmazdır. Ticari, eğitim/araştırma ve kamu kurumları için Open Science Grid benzeri konsorsiyumların oluşturulması ya da mevcut uluslararası konsorsiyumlara katılım yoluyla gerekli altyapılar tesis edilmelidir. Ayrıca, bu faaliyetlerin uluslararası standartlara uygun biçimde yürütülmesini sağlayacak; gerektiğinde danışmanlık ve kalite kontrolü gerçekleştirecek; yurt dışında ve uluslararası platformlarda ülkemizi temsil edebilecek ve resmi muhatap olarak kabul edilecek çok kurumlu, çok paydaşlı bir yapı altında toplanmış temsilci bir kurum ya da enstitü oluşturulması önerilebilir.

8. Sonuç

Yeni nesil kurumsal bilgi depoları, yalnızca açık erişimli bilimsel yayınların sunulduğu platformlar olmaktan çıkarak, daha kapsamlı bilgi ve veri yönetim sistemlerine dönüşmeye başlamıştır. Bu dönüşüm, özellikle Bilim 2.0’ın ortaya çıkardığı büyük ölçekli araştırma verileri ve karmaşık veri kümeleri karşısında, DSpace gibi geleneksel platformların yetersiz kalmasıyla daha da belirgin hâle gelmiştir. 2010’lu yıllardan itibaren bilimsel veri ve bilgi yönetiminin, araştırma verilerini de kapsayan büyük veri kümelerine yönelik olarak evrildiği gözlemlenmektedir. Bu bağlamda, yapay zekâ teknolojileriyle desteklenen bilgi ve araştırma verisi yönetimi yaklaşımları; verinin üretiminden analizine ve bilim dünyasına sunumuna kadar tüm süreçleri kapsayan, daha donanımlı ve sürdürülebilir siber altyapıları destekleyen yeni nesil platformların gelişimini hızlandırmıştır. Bu kapsamda öne çıkan örnekler arasında Dataverse, FAIRDOM, ISA, Open PHACTS, wwPDB ve UniProt gibi platformlar yer almaktadır; bunların FAIR ilkelerini nasıl karşıladıkları analiz edilmiştir.

Araştırma kütüphaneleri geleneksel olarak disiplin veya konu odaklı hizmet vermekte olup, desteklediği sistemler zamanla teknolojik olarak gelişmiş olsa da yukarıda açıklanan e-bilim ortamının ölçeğinde veya karmaşıklığında sistemlere sahip değildir. Üstelik, mali konulardaki kısıtlamalar, akademik ve araştırma kütüphanelerinin teknolojik yeniliklere ve e-bilimin gerektirdiği altyapı sistemlerine ayak uydurmasını zorlaştırmaktadır. Bilimdeki güncel eğilimler giderek daha fazla veri yoğunluklu hâle gelirken, akademik kütüphaneler ne yazık ki genellikle bilimsel verilerden doğrudan sorumlu değildir.

e-Bilim doğası gereği ekip çalışması ile yürütülmekte, kurumsallık ve küresel ölçekte iş birliğini gerektirmektedir. Bu bağlamda verinin işlenmesi, saklanması ve dağıtımı gibi süreçlerin ağ temelli, çok paydaşlı kurumsal yapılar aracılığıyla yürütülmesi daha etkili olmaktadır. Kütüphaneler ise çoğunlukla yerel kuruma bağlı araştırmacılara hizmet vermekte; bu da veri sahipliği, telif hakları gibi konularda ulusal ve uluslararası düzeyde işleyen protokollerin oluşturulmasını zorlaştırmaktadır. Elektronik içerikler için kullanılan lisanslar genellikle tek bir kurumsal toplulukla sınırlı olup, bu lisansları çok kurumlu ortamlara taşımak için gerekli teknolojik ve hukuki altyapı henüz yeterince gelişmemiştir. Dolayısıyla e-bilim, geleneksel araştırma kütüphanesi yapılarıyla çelişen; hukuki bağlayıcılığı olan lisans sorunlarından, katı bürokratik yapılar ve kemikleşmiş kurumsal şemalardan arındırılmış; konsorsiyum yapılarıyla işleyen, çok kurumlu lisansların ve siber altyapıların paylaşılabilirdiği kolektif bir bilimsel anlayış gerektirmektedir. Açık bilim uygulamaları böyle bir açık, esnek ve birlikte çalışabilir bilimsel ekosistem içerisinde sürdürülebilir şekilde hayata geçirilebilir.

Bu çalışmada ele alınan COAR Üstveri Kalitesi Çalışma Grubu girişimleri, konuları araştırmacı ve kullanıcı bakış açısından olduğu kadar, kütüphane hizmetleri ve veri küratörlüğü perspektiflerinden de değerlendirilerek çok yönlü bir katkı sunmuştur. Türkiye’de, uluslararası kalite kontrolü ve standartlara uyum sürecinde; bilimsel veri yayınlama ekosisteminin tüm yönlerini kapsayacak şekilde, daha geniş bir paydaş topluluğu ve iç içe geçmiş vizyonlar etrafında birleşen ortak girişimlere ihtiyaç duyulmaktadır. Önümüzdeki dönemde, e-bilim ve Bilim 2.0’ın bir gerekliliği olarak, ham verinin ve bilimsel çıktıların küresel düzeyde hızlı erişilebilir ve standartlaştırılmış üstveriler aracılığıyla küratörlüğünün yapılması; bu verilerin araştırılabilir formata dönüştürülerek küresel altyapılarla bütünleşik şekilde sunulması zorunlu hâle gelmiştir. COAR Üstveri Kalitesi Çalışma Grubu, bu dönüşümün altyapısını oluşturmak amacıyla FAIR ilkelerini belirlemiş ve bu makalede söz konusu ilkeler doğrultusunda yürütülen faaliyetler bilgi profesyonellerinin dikkatine sunulmuştur. Bugün küresel çapta birçok yükseköğretim kurumu, yapay zekâ destekli kurumsal bilgi deposu sistemlerine geçerken ve araştırma verilerini Dataverse gibi platformlar aracılığıyla açık erişime sunarken, Türkiye’de de benzer bir dönüşümün ivedilikle gerçekleştirilmesi gerekmektedir. Bu kapsamda, bir yandan bilgi yönetimi alanında uzman insan kaynağının yetiştirilmesine yönelik çalışma gruplarının oluşturulması, diğer yandan da güvenilir veri yönetimi adına uluslararası sertifikasyon ve akreditasyon süreçlerine aktif katılım sağlanması, Türkiye’nin bilimsel ekosisteminin çağdaş düzeyde sürdürülebilirliğini sağlayacak stratejik adımlar olarak değerlendirilmektedir.

Etik Standartlar ile Uyumluluk

Çıkar Çatışması: Yazar, herhangi bir çıkar çatışmasının olmadığını beyan eder.

Etik Kurul İzni: Bu çalışma için etik kurul iznine gerek yoktur.

Yazar Katkı Beyanı: Makale tek yazarlıdır.

Finansal Destek: Yoktur.

Kaynakça

- Bauch, A., Adamczyk, I., Buczek, P. vd. (2011). openBIS: A Flexible Framework for Managing and Analyzing Complex Data in Biology Research. *BMC Bioinformatics*, 12, 468. doi.org/10.1186/1471-2105-12-468
- Bell, G., Hey, T. ve Szalay, A. (2009). Computer Science: Beyond the Data Deluge. *Science*, 323(5919), 1297-1298. [doi:10.1126/science.1170411](https://doi.org/10.1126/science.1170411)
- Berman, H. M, vd. (2000). The Protein Data Bank. *Nucleic Acids Res*, 28, 235–242.
- Berman, H., Henrick, K., ve Nakamura, H. (2003). Announcing the Worldwide Protein Data Bank. *Nat. Struct. Biol.* 10, 980–980.
- Bishop, D. ve Eoin Gill. (2020). Robert Boyle on the Importance of Reporting and Replicating Experiments. *The Royal Society of Medicine*, 133(2), 79-83. doi.org/10.1177/0141076820902625
- Blixrud, J., ve Rambo, N. (2007). *Agenda for Developing E-Science in Research Libraries Final Report and Recommendations to the Scholarly Communication Steering Committee, the Public Policies Affecting Research Libraries Steering Committee, and the Research, Teaching, and Learning Steering Committee.* <https://www.arl.org/wp-content/uploads/2007/12/escience-report-final-2007.pdf>
- Bohle, Shannon. (2013, 12 Haziran). What is e-Science and how should it be Managed? *Scientific and Medical Libraries. Scilogs. Nature and Spektrum der Wissenschaft.* <http://www.scilogs.com/scientific-and-medical-libraries/what-is-e-science-and-how-should-it-be-managed/>
- Borgman, Christine L. (2004, 9 Kasım). *E-science, Digital Libraries, and Knowledge Communities.* Public Lecture, Dept. of Computer and Information Sciences. University of Oxford. <https://www.oii.ox.ac.uk/news-events/events/e-science-digital-libraries-and-knowledge-communities/>

- Bourne, P. E., Berman, H. M., McMahon B., Watenpugh, K. D., Westbrook, J.D., Fitzgerald, Paula M.D. (1997). The Macromolecular Crystallographic Information File (mmCIF). *Meth. Enzym*, 277, 571–590. [doi.org/10.1016/S0076-6879\(97\)77032-0](https://doi.org/10.1016/S0076-6879(97)77032-0)
- Burton, O. ve Appleford, S. (2009). Cyberinfrastructure for the Humanities, Arts, and Social Sciences. *EDUCAUSE Center for Applied Research: Research Bulletin*, 1(25), 1-12. <https://library.educause.edu/resources/2009/1/cyberinfrastructure-for-the-humanities-arts-and-social-sciences>
- Chivers, T. (2018, 1 Ekim). How Big Data is Changing Science. *Mosaic. Welcome.org*. <https://mosaicscience.com/story/how-big-data-changing-science-algorithms-research-genomics/>
- COAR. (2020). *COAR Next Generation Repositories Technologies*. <https://ngr.coar-repositories.org/>
- CoreTrustSeal Board (CTS). (2019). *Core Trustworthy Data Repositories Requirements*. CoreTrustSEAL.
- Crosas, M. (2011, Şubat). The Dataverse Network®: An Open-source Application for Sharing, Discovering and Preserving Data. *D-Lib Mag*, 17(1), 2. <http://www.dlib.org/dlib/january11/crosas/01crosas.html>
- Dahl, M. (2018). Inside-out Library Services. *Challenging the “Jacks of all Trades but Masters of None” Librarian Syndrome, Advances in Library Administration and Organization, Vol. 39* içinde (s. 15-34). Emerald Publishing Limited. doi.org/10.1108/S0732-067120180000039003
- Danish e-infrastructure Cooperation (DeiC). (2014). *Big Data in Social Sciences: The Experiment: SensibleDTU- What does a Social Network Look in Real Time?* Escience vidensportal. <https://sunelehmenn.com/sensibleDTU-publications/>
- Danmark Forsknings- og Innovationsstyrelsen (2011). *Dansk Roadmap for Forskningsinfrastruktur 2011*. <https://ufm.dk/publikationer/2011/dansk-roadmap-for-forskningsinfrastruktur-2011>
- Dempsey, L. (2013). The Emergence of the Collective Collection: Analyzing Aggregate Print Library Holdings. L. Dempsey, B. Lavoie, C. Malpas, L.S. Connaway, ve R.C. Schonfeld (Ed.), *Understanding the Collective Collection: Towards a system-wide Perspective on Library Print Collections* içinde (s. 1–8). OCLC Research. <http://www.oclc.org/content/dam/research/publications/library/2013/2013-09intro.pdf>
- Dempsey, L. (2015a). A New Information Management Landscape. N. Allen, ve B. Wilson (Yay. Haz.). *New Roles for the Road Ahead: Essays Commissioned for ACRL's 75th anniversary* içinde (s. 50–55). ACRL.
- Dempsey, L. (2015b). Intra-institutional Boundaries: New Contexts of Collaboration on Campus. N. Allen ve B. Wilson (Ed.), *New Roles for the Road Ahead: Essays Commissioned for ACRL's 75th anniversary* içinde (s. 80–82). ACRL https://www.ala.org/sites/default/files/acrl/content/publications/whitepapers/new_roles_75th.pdf
- Dempsey, L., Malpas C. ve Lavoie B. (2014). Collection Directions: The Evolution of Library Collections and Collecting. *Libraries and the Academy*, 14(3), 393–423. doi.org/10.1353/pla.2014.0013
- European Science Foundation. (2018). Plan S. <https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/>
- Fenner, M., vd. (2019). A Data Citation Roadmap for Scholarly Data Repositories. *Scientific Data*, 6(1), 1-9.
- Force11. (2019). *Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing Version b1.0*. <https://www.force11.org/fairprinciples>

- González-Beltrán, A., Maguire, E., Sansone, S.-A. ve Rocca-Serra, P. (2014). linkedISA: Semantic Representation of ISA-Tab Experimental Metadata. *BMC Bioinformatics* 15 (Suppl 14), S4, 1-9. doi.org/10.1186/1471-2105-15-S14-S4.
- Gonzalez-Beltran, A., vd. (2015). From Peer-reviewed to Peer-reproduced in Scholarly Publishing: The Complementary Roles of Data Models and Workflows in Bioinformatics. *Plos One*, 10(7), 1-20. doi.org/10.1371/journal.pone.0127612
- Groth, P.T., Loizou, A., Gray, A.J.G., Goble, C., Harland ve L. Pettifer, S. (2014). Api-centric Linked Data Integration: The Open Phacts Discovery Platform Case Study. *Ssrn Electronic Journal*, 29, 12-18. doi.org/10.1016/j.websem.2014.03.003
- Harland, L. ve Open PHACTS Consortium. (2012). A Semantic Knowledge Infrastructure for Public and Commercial Drug Discovery Research. *Lecture Notes in Computer Science*, 7603, 1–7. doi.org/10.1007/978-3-642-33876-2_1.
- Hey, T. ve Hey, J. (2006). E-science and its Implications for the Library Community. *Library Hi Tech*. 24(4), 515–528. doi: 10.1108/07378830610715383.
- Hey, T., vd. (2009). *The Fourth Paradigm: Data-intensive Scientific Discovery*. Redmond, Wash: Microsoft Research.
- Hilbert, M. (2015). *e-Science for Digital Development: “ICT4ICT4D”*. Development Informatics, Working Paper Series, The University of Manchester. <https://hummedia.manchester.ac.uk/institutes/gdi/publications/workingpapers/di/di-wp60.pdf>
- https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf
- Kobro-Flatmoen, A. (2012). NextGen VOICES: Results. *Science Magazine*, 335(6064), 36-38. doi:10.1126/science.335.6064.36
- Leeuw, L. de. (2019). *Data Seal of Approval (DSA)*. DANS. doi:10.17026/dans-28z-njxq
- Lin, D., Crabtree, J., Dillo, I. vd. (2020). The TRUST Principles for Digital Repositories. *Science Data*, 7(144), 1-5. doi: 10.1038/s41597-020-0486-7
- MacCallum, C. (2019, 19 Ekim). Data Repository Selection: Criteria that matter. *Open Science*. <https://osf.io/m2bce>
- Martin, E. R. ve Kafel, D. (2010). Response to Neil Rambo's editorial: E-science and Biomedical Libraries. *University of Massachusetts Medical School. Library Publications and Presentations. Paper 106(98)*, 528. doi: [10.3163/1536-5050.98.1.003](https://doi.org/10.3163/1536-5050.98.1.003)
- Moe-Behrens, G. (2012, 6 Ocak). NextGen Speaks. *Science Magazine* 335, 6064, 36-38. /doi/10.1126/science.335.6064.36
- Mokran, M. ve Parsons, M. (2014). Learning from the International Polar Year to Build the Future of Polar Data Management. *Data Science Journal*, 13(15), 88-93. doi.org/10.5334/dsj.IFPDA-15
- PID services (DOI, ePIC). (2024). Swedish National Data Service. <https://snd.se/en/describe-and-share-data/pid-services-doi-epic>
- Pitkin, K. (2011, 25 Kasım). IDCC11 Preview: An interview with Victoria Stodden. DCC. Dcc.ac.uk. 2 Eylül 2022 tarihinde <https://www.dcc.ac.uk/news/idcc11-preview-interview-victoria-stodden> adresinden erişilmiştir.
- Rose, P.W., vd. (2015). The RCSB Protein Data Bank: Views of Structural Biology for Basic and Applied Research and Education. *Nucleic Acids Research*, 42(D1), D345-D356.
- Seneca, T. (2010). *Web Archive Discovery: Memento Implementation Meeting*. California Digital Library. <https://cdlib.org/cdinfo/2010/02/04/web-archive-discovery-memento-implementation-meeting/>

- Syed, J., vd. (2007). Supporting Scientific Discovery Processes in Discovery Net. *Concurrency and Computation: Practice and Experience*, 19(2), 167-179. doi:10.1002/cpe.1049
- Tanenbaum, Andrew S. ve Steen, Maarten van. (2002). *Distributed Systems: Principles and Paradigms*. Upper Saddle River, NJ: Pearson Prentice Hall.
- UniProt Consortium. (2015). UniProt: A Hub for Protein Information. *Nucleic Acids Research*, 43, 204-212. doi:10.1093/nar/gku989
- Wilkin, J. P. (2015). Meanings of the Library Today. A. Crawford (Yay. Haz.) *The meaning of the Library: A Cultural History* içinde (s. 236-253). Princeton and Oxford: Princeton University Press. doi.org/10.1515/9781400865741-014
- Wilkinson, M. D., vd. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Science Data*, 3(160018), 1-9. doi.org/10.1038/sdata.2016.18
- Witt, M. ve Carlson, J. R. (2007). Conducting a Data Interview. Purdue University. *Libraries Research Publications. Paper 81*. https://docs.lib.purdue.edu/lib_research/81
- Wolstencroft, K., vd. (2015). SEEK: A Systems Biology Data and Model Management Platform. *BMC Syst Biol*, 9(33), 1-12. doi: 10.1186/s12918-015-0174-y.