

# Veri Olarak Kültürel Miras Koleksiyonları\*

*Cultural Heritage Collections as Data*

Meltem DİŞLİ\*\* ve Yaşar TONTA\*\*\*

## ÖZ

**Amaç:** Günümüzde dijital insani bilimler araştırmalarında kullanılan ana veri/bilgi kaynakları dijital ortama aktarılmış çözümlü yüksek belgeler ve verilerden oluşan koleksiyonlar ile bu koleksiyonlara ait makinece okunabilir ve işlem yapılabilir üstverilerdir. Bu araştırmanın amacı dijital insani bilimler araştırmacılarının “veri olarak koleksiyonlara” yönelik ihtiyaç ve beklentilerini ve kültürel bellek kurumlarının bu ihtiyaç ve beklentileri ne ölçüde karşıladıklarını saptamaktır.

**Yöntem:** Araştırma kapsamında nitel yöntemler kullanılmış, 10 dijital insani bilimler araştırmacısı ve 16 kültürel bellek kurumu temsilcisi ile yarı yapılandırılmış görüşmeler gerçekleştirilmiştir. Araştırmacılara 20, kurum temsilcilerine 23 soru sorulmuş, tüm görüşme kayıtları çözümlenmiş (transkribe edilmiş) ve MaxQDA nitel veri analiz yazılımı aracılığıyla görüşme kayıtlarının içerik analizi yapılmıştır.

**Bulgular:** Araştırmacılar koleksiyonlara makinece okunabilir ve işlem yapılabilir formatlarda veri setleri halinde erişmeye ihtiyaç duymaktadırlar. Protokol karşılığında veya talep olduğunda koleksiyonlarını bu şekilde paylaşabilecek az sayıda kurum olsa da henüz hiçbir kurum makinece okunabilir formatlarda ve bilişimsel erişime uygun biçimde dijital koleksiyonlarını hizmete açmamıştır. Kurumların koleksiyonlarını dönüştürmek için nitelikli personele, bütçeye, teknik altyapıya, yönetim desteğine, işbirliğine, zamana ve konuya ilişkin bir rehber ihtiyaçları bulunmaktadır.

**Sonuç:** Kurumlar, üstverileri düzenlemek, koleksiyonları dijitalleştirmek ve bu koleksiyonları veri olarak koleksiyon yaklaşımına uygun olarak araştırmacıların erişimine açmak için bu konuda farkındalık oluşturmaları, teknik, idari ve maddi yönlerden desteklenmeli ve nitelikli personel yetiştirmelidir.

**Özgünlük:** Bu çalışma hem araştırmacıların hem de kültürel bellek kurumlarının bakış açılarından Türkiye’de veri olarak koleksiyonlar hizmetine yönelik gereksinimleri, beklentileri ve engelleri saptayan ilk araştırmadır.

**Anahtar Sözcükler:** Veri olarak koleksiyonlar, kültürel miras, dijital insani bilimler, bilişimsel erişim

\* Bu makale Meltem Dişli’nin Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümünde yaptığı doktora tezine dayanmaktadır.

\*\* Hacettepe Üniversitesi, Bilgi ve Belge Yönetimi Bölümü, Ankara, Türkiye. E-posta: dislimeltem@gmail.com

Hacettepe University, Department of Information and Document Management, Ankara, Türkiye. E-mail: dislimeltem@gmail.com

\*\*\* Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü, Ankara, Türkiye. E-posta: yasartonta@gmail.com

Hacettepe University, Department of Information and Document Management, Ankara, Türkiye. E-mail: yasartonta@gmail.com

**ABSTRACT**

**Purpose:** In today's digital humanities research, the main data/information sources used by researchers are collections consisting of digitized documents and data with high resolution, along with machine-readable and actionable metadata associated with these collections. The purpose of this research is to determine the needs and expectations of digital humanities researchers for "collections as data" and to find out to what extent cultural memory institutions meet these needs and expectations.

**Method:** Qualitative methods were used within the scope of the research. Semi-structured interviews were conducted with 10 digital humanities researchers and 16 representatives of cultural memory institutions. The researchers and the representatives of the institutions were asked 20 and 23 questions, respectively, and all interview records were transcribed and their contents were analyzed using MaxQDA, a qualitative data analysis software.

**Findings:** Researchers need access to collections as datasets in machine-readable and machine-actionable formats. While there are a few institutions that can share their collections through protocols or upon request, no institution has yet made their collections available in machine-readable and actionable formats suitable for computational access. To convert their collections, institutions need qualified staff, budget, technical infrastructure, administrative support, cooperation, time, and guidelines.

**Implications:** In order for institutions to organize metadata, digitize collections, and make them accessible to researchers as collections as data, they should create awareness in this regard, receive technical, administrative and financial support, and train qualified personnel.

**Originality:** This study is the first of its kind that identifies the needs, expectations, and barriers regarding the "collections as data" service in Turkey from the perspectives of researchers and cultural memory institutions.

**Keywords:** Collections as data; cultural heritage; digital humanities; computational access.

**Giriş**

Bilgi teknolojilerinde yaşanan gelişmeler her alanda olduğu gibi bilim dünyasında da değişimlere sebep olmuştur. Günümüz bilim dünyasında Jim Gray'in dördüncü bilim paradigması olarak tanımladığı veri yoğun bilim (data-intensive science) önem kazanmıştır (Gray, 2009). İnsani bilimlerde de verinin yoğun kullanımı, dijital verilere erişim imkânlarının artması ve bilgi teknolojilerinden daha fazla yararlanılması yeni bir disiplin olarak dijital insani bilimler disiplinini ortaya çıkarmıştır. Dijital insani bilimler makine öğrenimi, veri bilimi, yapay zekâ alanlarını da içeren bilgisayar tabanlı teknolojilerin insani bilimlere uygulanmasıdır (Berry, 2019). Dijital insani bilimler çalışmalarının temel veri kaynağı, toplumların geçmiş nesillerden gelecek nesillere aktardığı kültürel, sosyal, ekonomik ve evrensel değerlere sahip olan kültürel miras koleksiyonlarıdır (Akça, 2018). Bu çerçevede son yıllarda kültürel bellek koleksiyonlarındaki verilere erişimi kolaylaştırmak ve dijital insani bilimler araştırmacılarının gereksinimlerini karşılamak amacıyla "veri olarak koleksiyonlar" (collections as data) yaklaşımı benimsenmektedir (Wittmann ve diğerleri, 2019). Veri olarak koleksiyonlar (VOK) yaklaşımı dijital koleksiyonları bilişimsel (hesaplamalı) erişime uygun veri setlerine dönüştürmek şeklinde tanımlanmaktadır (Padilla ve diğerleri, 2019). Bilişimsel erişim (computational access) dijital arşivlerdeki koleksiyonları veri/metin madenciliği ve makine öğrenimi gibi tekniklerle analiz etmek ve yorumlamak için indirerek ya da bir uygulama programlama arayüzü (Application Program Interface – API) kullanarak bu koleksiyonlara makinece okunabilir bir şekilde erişmek anlamına gelmektedir (Digital Preservation Coalition, 2022).

Koleksiyonların veri olarak koleksiyonlara dönüştürülmesi dijital insani bilimler, veri bilimi, veriye dayalı gazetecilik gibi birçok disipline hizmet etmektedir (Padilla ve Ridge, 2019). Bu yaklaşım kültürel miras koleksiyonları üzerinde yapılan veri ve metin madenciliği, veri görselleştirme, haritalama, konu modelleme ve ağ analizi gibi uygulamaların çok daha etkin bir şekilde gerçekleştirilmesini sağlamaktadır (Lampert ve Lapworth, 2020). Bu tür uygulamaları ve bilişimsel erişim yöntemlerini kullanan araştırmacıların sayısı her geçen gün artmaktadır. Bu artışa paralel olarak kütüphane, müze, arşiv gibi çok sayıda kültürel bellek kurumu da veri setlerini CSV, XML gibi açık ve yeniden kullanılabilir formatlarda sunmaya başlamıştır (Wittmann ve diğerleri, 2019, s. 49; Candela ve diğerleri, 2022, s. 2). Ancak henüz bu uygulamalar yeterince yaygınlaşmamıştır. Bunun temel nedeni nitelikli personel eksikliği, yerleşik eğilimleri değiştirmenin zorluğu, korumacı yaklaşımlar, bilişimsel erişimi desteklemeyen geleneksel sistemler ve telif hakkı sorunları gibi engellerdir (Candela ve diğerleri, 2022, s. 2). Ayrıca bazı kültürel bellek kurumlarında veri setleri paylaşılsa da verilerin veya üstverilerin (metadata) kalitesi bilgisayar destekli analizler ile iyi sonuçlar elde etmek için yeterli değildir (Candela ve diğerleri, 2022, s. 15).

Türkiye kültürel miras açısından zengin bir ülkedir. Dolayısıyla ülkemizde veya yurt dışında kültürel miras koleksiyonlarımızı temel veri kaynağı olarak kullanan dijital insani bilimler çalışmaları giderek artmaktadır (Karbeyaz ve diğerleri, 2012, 2013; Grallert, 2016; Hadjikyriacou, 2018; Uğur, 2018; Ben-Bassat ve Buessow, 2020; Stanford University, 2021; University of St Andrews, 2021). Ancak araştırmacılar kültürel miras koleksiyonlarının üstverilerine ya da bu koleksiyonlara topluca erişim sağlayamamaktadırlar. Hem kültürel mirasın daha iyi analiz edilmesinde hem de dijital insani bilimler disiplininin gelişmesinde veri olarak koleksiyonlar yaklaşımının uygulanması kilit rol oynamaktadır.

Bu çalışmanın iki amacı vardır: (1) Türkiye'deki dijital insani bilimler araştırmacılarının veri olarak kültürel miras koleksiyonlarına yönelik gereksinim ve beklentilerini belirlemek; ve (2) kültürel bellek kurumlarının bu gereksinim ve beklentileri ne ölçüde karşıladığını saptamak. Bu kapsamda dijital insani bilimler araştırmacıları ve kültürel bellek kurumlarının temsilcileri ile görüşmeler gerçekleştirilmiştir. Görüşmeler sonucunda araştırmacıların en çok veriye erişim, dijital belgelerin çözünürlük (granularity) kalitesi, ücretler, indeksleme hataları ve yetersiz üstveriler açısından zorluklarla karşılaştıkları görülmüştür. Araştırmacıların kurumlardan beklentileri hususunda açık erişim, toplu indirme seçenekleri, makinece okunabilir formatlarda veriye erişim, dijitalleştirme gibi konular öne çıkmaktadır. Kurum temsilcilerinin ise veri olarak koleksiyonlar yaklaşımına olumlu baktıkları ancak henüz koleksiyonlarını bu yaklaşıma uygun biçimde hizmete sunmadıkları sonucuna ulaşılmıştır. Kurumların önünde telif hakları, nitelikli personel eksikliği, bütçe ve teknik yetersizlikler gibi engeller bulunmaktadır. Konuyla ilgili olarak Türkiye'de ilk kez yapılan bu çalışmadan elde edilen bulguların dijital kültürel miras koleksiyonlarını veri olarak koleksiyonlara dönüştürme politikası geliştirmeyi planlayan kurumlar tarafından kullanılabilmesi kanısındayız.

## Literatür Değerlendirmesi

Dijital dönüşüm, “*dijital teknolojileri kullanarak yeni fırsatlar ve değerler yaratma; sosyal yapıları dijital teknolojilerle güçlendirme ve daha verimli hale getirme süreci*” olarak tanımlanmakta ve her alanda hayatımızı etkilemektedir (Bozkurt ve diğerleri, 2021, s. 40, orijinali italik). Yaşanan gelişmelerin etkisi kütüphane, arşiv ve müze gibi kültürel bellek kurumlarında da yeni bir dönem başlatmıştır. Dijital dönüşüm sonucunda bellek kurumları farklı içerik türlerini dijital verilere dönüştürmekte, bu veriler arşivlerde kayıt altına alınmakta ve her yerden erişilebilir hâle getirilmektedir (Lanzolla ve Anderson, 2008, s. 73; Colavizza ve diğerleri, 2021). Kültürel bellek kurumlarında dijitalleşme alanındaki gelişmelere ek olarak dijital veri odaklı akademik çalışmaların her geçen gün artması kurumların veri olarak koleksiyonlar (VOK) yaklaşımını benimsemesine yol açmıştır.

VOK kavramı ilk kez 27 Eylül 2016 tarihinde gerçekleştirilen Amerikan Kongre Kütüphanesinin ev sahipliği yaptığı bir etkinlikte kullanılmıştır (Library of Congress, 2016). Ancak VOK fikrinin oluşması daha eskilere dayanmaktadır. Doğrudan bu kavram kullanılmamış olsa da literatürde koleksiyonların veri olarak kullanılmasını ele alan çeşitli çalışmalar bulunmaktadır (Johnston, 2011; Padilla ve Higgins, 2014; Posner, 2015; Padilla, 2016). Günümüzde üstveriler veya veri olarak koleksiyonlar toplu halde dijital ortamda kolayca paylaşılabilir. Aslında geçmişte de üstverilere toplu şekilde CD-ROM'lar, mikrofişler veya basılı kataloglar aracılığıyla erişim sağlanabilmekteydi. Örneğin, Millî Kütüphane'nin 2001 yılında yayımladığı 1923-1999 yılları arasında Türkiye'deki 4418 dergide yayımlanan 566.627 makalenin üstverilerine toplu olarak erişim imkânı sunan "Cumhuriyet Dönemi Makaleler Bibliyografyası 1923-1999 CDMB" adlı CD-ROM (Tonta ve Al, 2008) veri olarak koleksiyonlara örnek gösterilebilir. Günümüzün VOK uygulamalarını geçmişteki uygulamalardan farklı kılan en önemli özellik üstverilerin veya koleksiyonların çevrimiçi (online) olarak "makinece okunabilir ve makinece üzerinde işlem yapılabilir" (machine-readable and machine-actionable) ve yeniden kullanılabilir formatlarda sunulmasıdır. Makinece okunabilir veri yapısal olmalıdır. Makinece üzerinde işlem yapılabilirlik ise makinelerin ya da bilgisayarların böyle bir yapıya göre programlanabilmesini sağlayacak şekilde yapılandırılmış tutarlı bilgiyi ifade eder. Örneğin, makine bir XML dosyasındaki etiketleri (tags) kolayca ayırt edebilir ama bu etiketlerde kayıtlı bilgilerin ne olduğunu karmaşık metin analizi algoritmaları kullanılmadan anlayamaz. Oysaki aynı dosyadaki etiketler ad, başlık, yayınevi, yıl gibi spesifik ve bir üstveri şeması aracılığıyla (örneğin, Dublin Core) önceden tanımlanmış ve kayıtlı (tekil URL adresi olan) alan adlarıyla değiştirilirse, makinelerin kitapların yazar adlarını listelemesini sağlayan sorgular kolayca oluşturulabilir ve bu koleksiyonlardan yeni bilgiler çıkarılabilir (Miksa ve diğerleri, 2018). Britanya Kütüphanesi'ndeki RDF/XML formatlarında sunulan Britanya Millî Bibliyografyası, Britanya Kütüphanesi basılı müzik eserlerini içeren veri setleri veya Lewis Carroll'un tüm eserlerinin üstverilerinin bir CSV dosyası olarak sunulması makinece okunabilir ve üzerinde işlem yapılabilir koleksiyonlara örnek olarak gösterilebilir (British Library, 2023).<sup>1</sup>

VOK son yıllarda bazı kültürel bellek kurumları için temel faaliyet alanlarından biri hâline gelmiştir (Ames, 2021, s. 1). Giderek daha fazla kültürel bellek kurumu dijital kültürel miras koleksiyonlarını hem bilimsel hem de yaratıcı amaçlar için makinece okunabilir, açık ve yeniden kullanılabilir formatlarda veri setleri olarak sunmaktadır (Lampert ve Lapworth, 2020; Ziegler, 2020, s. 4; Candela ve diğerleri, 2022, s. 2). ABD, İngiltere, Hollanda, Fransa, Avustralya Millî Kütüphaneleri ve Avustralya Millî Arşivi koleksiyonlarını veri olarak koleksiyonlara dönüştüren kurumlardan bazılarıdır. Kültürel bellek kurumlarının yanı sıra HathiTrust ve Europeana gibi dijital kütüphaneler de bu yaklaşımı benimsemiştir (Europeana Pro, 2022; HathiTrust, 2022).

Bellek kurumları çeşitli koleksiyonları farklı biçimlerde veri olarak sunmaktadır. Örneğin, Britanya ve Kongre Kütüphaneleri koleksiyonlarını sıkıştırılmış dosyalar halinde çeşitli dosya formatlarında veri olarak paylaşmaktadır (British Library, 2022a; Library of Congress, 2022). Bu kurumların veri sıkıştırma biçimi olarak zip kullandığı ve koleksiyonlarını XML, RDF, CSV gibi formatlarda sundukları görülmektedir. Avustralya Millî Kütüphanesi koleksiyonlarını bir uygulama programlama arayüzü (API) aracılığıyla kullanıma açmıştır (National Library of Australia, 2022). Protokoller aracılığıyla yazılımların birbiriyle iletişim kurabilmesini sağlayan bu arayüzler, arama sonucu ulaşılan koleksiyonların tek seferde toplu halde indirilebilmesine (bulk download), koleksiyonların başka bir veri tabanına aktarılabilmesine veya koleksiyonlardan görselleştirme yapılabilmesine olanak sağlamaktadır. Avustralya Millî Kütüphanesi'nin uygulama programlama arayüzü (Trove) ile oluşturulan araçlar ve görselleştirmeler bunlara örnek olabilir (National Library of Australia, 2023). Europeana Dijital Kütüphanesinde hem Britanya ve Kongre Kütüphaneleri gibi sıkıştırılmış dosya formatlarında indirilebilir veri setleri sunulmakta hem de OAI-PMH (Açık Arşivler Girişimi Üstveri Harmanlama Protokolü)

<sup>1</sup> Örneğin, Mısır'daki Nefertari Tapınağına ait üstverilerin XML gösterimi için bkz. Turco ve diğerleri (2019, Şekil 4, s. 435).

aracılığıyla verileri toplama imkânı sunulmaktadır (Europeana Pro, 2022). Bu protokol ile Europeana verilerinin tamamı veya belirli veri setleri toplu bir şekilde indirilebilmektedir.

Koleksiyonlarını kendi kurumsal web sayfaları üzerinden veri olarak sunan bu kurumların aksine Carnegie Sanat Müzesi ve Miami Üniversitesi Kütüphanesi gibi bazı kurumlar da koleksiyonlarını GitHub<sup>2</sup> üzerinden paylaşmaktadır (Padilla ve diğerleri, 2019). VOK olarak paylaşılacak koleksiyonlar bellek kurumlarının önceliklerine ve kullanıcıların beklentilerine göre belirlenmektedir (Wittmann ve diğerleri, 2019, s. 59). Örneğin, Amerikan Kongre Kütüphanesinde daha çok bölgesel projeler, haritalar, dergiler ve yıllıklar gibi koleksiyonlar, Britanya Kütüphanesinde tarih ve edebiyat koleksiyonları veri olarak paylaşılmaktadır. Singapur Ulusal Üniversitesi Kütüphanesinde ise Endonezya tiyatro arşivindeki belgeler veri odaklı video işleme çalışmalarına destek olabilecek bir veri setine dönüştürülmüştür (Padilla ve diğerleri, 2019, s. 179).

Kültürel bellek kurumlarında kurulan laboratuvarlar da VOK yaklaşımı için önemli girişimlerdir. Laboratuvarlarda hem koleksiyonlar açık veri setleri halinde paylaşmakta hem de dijital kültürel miras koleksiyonları üzerinde çeşitli analizler yapılmaktadır. Laboratuvarlar aracılığıyla koleksiyonların daha iyi tanıtılması ve araştırmacıların teşvik edilmesi amaçlanmaktadır. Danimarka, Hollanda ve Britanya Milli Kütüphanelerinin laboratuvarları bunların en önemli örneklerindedir (British Library, 2022b; Koninklijke Bibliotheek, 2022; Royal Danish Library, 2022). Kültürel bellek kurumlarından bağımsız bir şekilde kurulan dijital insani bilimler laboratuvarları da bulunmaktadır. Hollanda'daki Dijital İnsani Bilimler Laboratuvarı bilişimsel dilbilim, bilgi erişim, semantik web gibi uzmanlıkları bir araya getirerek insani bilimler araştırmalarını geliştirmeyi amaçlayan kurumlar arası bir araştırma grubudur (Digital Humanities Lab, 2022). Benzer amaçlarla 2022 yılında Türkiye'de de Sabancı Üniversitesinde bir laboratuvar kurulmuştur (Sabancı Üniversitesi Dijital..., 2022). Veri olarak koleksiyonlara ilişkin gerçekleştirilen önemli çalışmalardan biri de GLAM Workbench'tir.<sup>3</sup> GLAM Workbench çoğunlukla Avustralya ve Yeni Zelanda'da bulunan kütüphane, arşiv, müze ve galeri gibi kültürel bellek kurumlarının veri setlerinin bir arada yer aldığı, bu verilerle çalışmaya yardımcı olacak araçlar, eğitimler ve örnekler sunan bir web sayfasıdır (Sherratt, 2022).

Konuyla ilgili çalışmaların bir kısmı da kütüphaneler tarafından gerçekleştirilmektedir. Las Vegas Nevada Üniversitesi Kütüphanesinin koleksiyonlarını veri olarak koleksiyonlara dönüştürmek için gereken üstveriler kütüphaneciler ile bir dijital tarih araştırmacısının işbirliğiyle hazırlanmıştır (Rueda, 2018a, 2018b). Utah Üniversitesi J. Willard Marriott Kütüphanesinde koleksiyonların kullanılabilirliği metin madenciliği, konu modelleme ve CBS gibi çeşitli dijital insani bilimler yöntemleri kullanılarak test edilmiş, kütüphanedeki dijitalleştirme süreçlerine veri olarak koleksiyonlar standart bir uygulama olarak dâhil edilmiş ve Dublin Core<sup>4</sup> üstveri alanları dışında daha ayrıntılı tanımlayıcı üstveriler oluşturulmuştur (Wittmann ve diğerleri, 2019).

Veri olarak koleksiyonlara yönelik en kapsamlı çalışma "Always Already Computational: Collections as Data" projesinin faaliyet ve bulgularını içeren rapordur (Padilla ve diğerleri, 2019). Bu raporda kurumların veri olarak koleksiyonlar uygulamalarını başlatmak için yapabilecekleri 50 unsur listelenmiş, bu alanda çalışan uzmanlar tanıtılmış, kütüphane, arşiv ve müze çalışanlarına yardımcı olacak uygulamalar gösterilmiştir. Raporda veri olarak koleksiyonlar için uygun formatlar, analiz araçları, metin madenciliği, ağ analizi gibi teknikler incelenmiş ve bu teknikler kullanılarak hazırlanan koleksiyonlardan örnekler sunulmuştur. Proje katılımcılarının koleksiyonlarını dönüştürme aşamasındaki deneyimlerine de raporda yer verilmiştir.

2 GitHub: <https://github.com/>

3 "GLAM", Galleries, Libraries, Archives, and Museums'un kısaltmasıdır. Türkçede genellikle KAM (Kütüphaneler, Arşivler ve Müzeler) kısaltması kullanılmaktadır. "GLAM Workbench", KAM Çalışma Tezgâhı ya da KAM Çalışma Masası olarak Türkçeye çevrilebilir.

4 Dublin Core Metadata Initiative: <https://www.dublincore.org/>

Raporda ayrıca “Veri Olarak Koleksiyonlara İlişkin Santa Barbara Bildirgesi” yer almaktadır. Bildirgede veri olarak koleksiyonların amacı, kapsamı, hedef kitlesi ve nasıl yapılandırılması gerektiği gibi konularda 10 ilke belirlenmiştir (The Santa Barbara Statement..., 2018). Aynı projenin devamı niteliğindeki “Collections as Data: State of The Field and Future Directions” toplantısı 25-26 Nisan 2023 tarihlerinde yapılmış (Collections as Data Futures, 2023) ve bir görüş bildirgesi yayımlanmıştır (Chambers ve diğerleri, 2023). Bu toplantıyla aynı adı taşıyan ayrıntılı bir rapor hazırlanması ve 2023 yılı içinde Santa Barbara Bildirgesinin güncellenmesi planlanmaktadır.

Uluslararası GLAM Laboratuvarları Topluluğundan<sup>5</sup> esinlenerek hazırlanan *Open a GLAM Lab* adlı kitap da dijital koleksiyonların veri olarak tanımlanması, değerlendirilmesi, erişime açılması, paylaşılması ve verilerin yeniden kullanılması hakkında bilgiler içeren önemli bir kaynaktır (Mahey ve diğerleri, 2019). Diğer çalışmalarda koleksiyonlarını dönüştürmek isteyen kurumlar için bir metodoloji önerilmiş (Candela ve diğerleri, 2022) ve veri olarak koleksiyonları yayımlama konusunda bir kontrol listesi hazırlanmıştır (Candela ve diğerleri, 2023). Benzeri bir biçimde kültürel miras verileri için de “makine öğrenimi verisi olarak koleksiyonlar” (collections as machine learning data) kontrol listesi oluşturulmuştur (Lee, 2023). Literatürde konu ile ilgili karışımıza çıkan bir diğer kavram da makinece işlem yapılabilir koleksiyonlardır (machine actionable collections). Bu kavram koleksiyonların veri olarak paylaşılmasına ek olarak bu verilerin keşfedilmesi, işlenmesi, yeniden kullanımı için gereken teknolojileri, süreçleri, standartları ve altyapıları da içerir ve yapay zekâ ve makine öğrenimini de besleyen veriler sağlar (Bell ve diğerleri, 2021, s. 68). Bunlar dışında veri olarak koleksiyonlar konusunda birçok proje gerçekleştirilmekte (örneğin, Adalink Project, 2022; Amsterdam Time Machine, 2022; HisGis, 2022; The real face of White Australia, 2022) ve çeşitli akademik araştırmalar yayımlanmaktadır (örneğin, Schmidt, 2012; Sherratt, 2019; UC San Diego Library, 2021; Candela ve diğerleri, 2022).

Ulusal literatürde veri olarak koleksiyonlar hakkında bazı tanıtıcı yayınlar yapılmış olsa da (örneğin, Tonta ve Madran, 2021) konuyla doğrudan ilgili henüz herhangi bir akademik araştırma yapılmamıştır. Doğrudan veri olarak kültürel miras koleksiyonlarına değinilmese de, dijital arkeolojik verilerin yönetimi konusunda hazırlanan bir doktora tezinde kültürel mirasın bir parçası olan arkeolojik verilere erişim ve bu verilerin yeniden kullanımı, üstveriler, standartlar ve bağlı veri gibi konulara yer verilmiştir (Çayırezmez, 2023). Konuyla bağlantılı en önemli girişimler ise yukarıda anılan Sabancı Üniversitesi Dijital İnsani Bilimler Laboratuvarı ile Dijital Osmanlı Çalışmaları platformudur (Digital Ottoman Studies, 2022; Sabancı Üniversitesi Dijital..., 2022). Laboratuvarda metin, görsel ve video gibi çeşitli veri türlerinin bilişimsel erişim yöntemleri ile analiz edildiği çeşitli projeler gerçekleştirilmektedir (Hasan Ağa ve Kardeşi..., 2022; Türkiye’yi Temsil Eden..., 2022). Dijital Osmanlı Çalışmaları platformunda ise Türkiye’deki dijital insani bilimler çalışmalarına katkı sağlayan projeler, makaleler, kitaplar, tezler, araçlar ve etkinlikler paylaşılmaktadır. Bu platform ile dijital insani bilimler projeleri için kişileri ve kurumları bir araya getiren bir dijital ağ oluşturulması amaçlanmaktadır (Digital Ottoman Studies, 2022).

Türkiye’de yapılmış olan dijital insani bilimler araştırmalarının birçoğu kültürel bellek kurumlarının koleksiyonlarına dayanmaktadır (Karbeyaz ve diğerleri, 2012, 2013; Uğur, 2014; Gökmen, 2015; Öner, 2019; Aladağ, 2020; Bekman, 2020; Ekim ve diğerleri, 2021; Şavk ve Doğu, 2021; Yüksek, 2021). Ancak Türkiye’de araştırmacıların veri bulmaları son derece zordur. “Osmanlı Epir’ini Haritalandırmak” adlı projenin yürütücüsü Prof. Ali Yayıoğlu’ya göre dijital tarihin gelişmesinin önündeki en önemli engel veri bulmaktır. Türkiye’de Osmanlıca yazılmış çok zengin belgeler olmasına rağmen projede yararlanılan belgelerin öncelikle veriye dönüştürülmesi gerekmiştir (ANAMED Kütüphanesi, 2021). Türkiye’de dijital insani bilimler araştırmalarının gelişmesi için kültürel bellek kurumlarının koleksiyonlarını veri olarak koleksiyonlara dönüştürmeleri ve bu verileri dijital ortamda araştırmacıların kullanımına açmaları son derecede önemlidir.

5 International GLAM Labs Community. <https://glamlabs.io/>

## Yöntem

Bu çalışmada aşağıdaki araştırma sorularına yanıt aranmıştır:

- 1.AS1: Dijital insani bilimler araştırmacıları veri olarak kültürel miras koleksiyonlarına gereksinim duymakta mıdır?
- 2.AS2: Dijital insani bilimler araştırmacılarının kültürel bellek kurumlarından beklentileri nelerdir?
- 3.AS3: Kültürel bellek kurumları kültürel miras koleksiyonlarını veri olarak koleksiyonlar yaklaşımına uygun biçimde sunmakta mıdır?
- 4.AS4: Kültürel bellek kurumlarının koleksiyonlarını veri olarak koleksiyonlara dönüştürmek için gereksinimleri nelerdir?

Araştırmanın temel hipotezi (H1) “Türkiye’deki kültürel bellek kurumlarının veri olarak kültürel miras koleksiyonlarına yönelik mevcut uygulama ve süreçleri dijital insani bilimler araştırmacılarının gereksinim ve beklentilerini karşılamamaktadır” şeklinde belirlenmiştir.

Araştırma kapsamında veri olarak koleksiyonlara yönelik dijital insani bilimler araştırmacılarının gereksinimlerinin ve kültürel bellek kurumlarının mevcut durumlarının saptanması amacıyla nitel bir araştırma planı hazırlanmıştır. Nitel araştırmalarda öncelik genelleme yapmak değil; olayların, durumların derinlemesine incelenmesi ve anlaşılmasıdır (Yıldırım ve Şimşek, 2018, s. 61). Bu çerçevede hem araştırmacıların gereksinim ve beklentilerinin derinlemesine incelenerek keşfedilmesi hem de kurumların mevcut durumlarının belirlenmesi amacıyla araştırmacılar ve kurumlar ile yarı yapılandırılmış görüşmeler yapılmıştır. Yarı yapılandırılmış görüşmelerde araştırmacı “görüşmenin akışına bağlı olarak değişik yan ya da alt sorularla görüşmenin akışını etkileyebilir ve kişinin yanıtlarını açmasını ve ayrıntılandırmasını sağlayabilir” (Türnüklü, 2000, s. 547). Araştırma kapsamında kültürel bellek kurumu temsilcilerine çalıştıkları kurum ve pozisyon, kurumun sahip olduğu dijital koleksiyonlar, bu koleksiyonların nasıl ve hangi koşullarda paylaşıldığı, koleksiyonların veri olarak paylaşılmasının önündeki engeller ve bu konudaki gereksinimler ile ilgili 23 soru sorulmuştur. Dijital insani bilimler araştırmacılarına ise araştırmacının görevi ve çalışma alanı, araştırmalarda kullanılan veri kaynakları, verilerin nasıl ve nerelerden elde edildiği, araştırmalarda karşılaşılan zorluklar ve bu konuda kültürel bellek kurumlarından beklentileri ile veri olarak koleksiyonlara yönelik görüşleri hakkında 20 soru sorulmuştur.<sup>6</sup>

Görüşmeler için dijital kültürel miras koleksiyonlarını kullanarak ağ analizi, metin madenciliği, haritalama ve benzeri teknikler kullanarak araştırmalar yapan ve araştırma bulgularını ulusal/uluslararası literatürde yayımlayan tarih, edebiyat, bilgisayar mühendisliği gibi farklı bölümlerden amaçlı örnekleme (purposive sampling) tekniğiyle 10 araştırmacı<sup>7</sup> ve Türkiye’de dijital kültürel miras koleksiyonları bulunan kütüphane, arşiv, müze ve araştırma merkezleri arasından 16 kurum<sup>8</sup> seçilmiştir. Araştırmacılar literatür taraması ile dijital insani bilimler alanında yapılan çalışmaların tespit edilmesi ve araştırmacıların profillerinin değerlendirilmesi sonucunda belirlenmiştir. Kurumlar da benzer şekilde kurumların web sayfaları incelenerek dijital insani bilimler araştırmaları için değerli koleksiyonlara sahip ve hâlihazırda koleksiyonlarını dijitalleştirmiş olanlar arasından seçilmiştir. Kurumlar arasında Osmanlı ve Cumhuriyet dönemi arşivlerini bünyesinde barındıran

6 Bkz. Ek 1: Veri olarak kültürel miras koleksiyonları: Kurumlar için görüşme soruları; Ek 2: Veri olarak kültürel miras koleksiyonları: Araştırmacılar için görüşme soruları

7 Araştırmacıların çalıştığı üniversiteler: Bilkent (2), Marmara, Hacettepe, İzmir Ekonomi (2), Kıbrıs Amerikan, Koç, Üsküdar, Yeditepe.

8 Cumhurbaşkanlığı Devlet Arşivleri Başkanlığı, Dijital Sakıp Sabancı Müzesi, Diyanet İşleri Başkanlığı Kütüphanesi, Marmara Üniversitesi Kütüphanesi, Millet Kütüphanesi, Milli Kütüphane, Salt Araştırma, TDV İSAM, Türk Dil Kurumu Kütüphanesi, Türk Tarih Kurumu Kütüphanesi, VEKAM, Kültür Varlıkları ve Müzeler Genel Müdürlüğü, İBB Atatürk Kitaplığı, İngiliz Arkeoloji Enstitüsü Kütüphanesi, İstanbul Üniversitesi Kütüphanesi.

Cumhurbaşkanlığı (CB) Devlet Arşivleri Başkanlığı, kültür varlıklarımız ile müzelerin bağlı olduğu Kültür Varlıkları ve Müzeler Genel Müdürlüğü, Türkiye'nin en büyük kütüphanesi olan CB Millet Kütüphanesi ve Millî Kütüphane yer almaktadır.

Görüşülen dijital insani bilimler araştırmacıları askeri tarih, Osmanlı tarihi, sinema tarihi, iktisat tarihi, kültür tarihi, Türk dili ve edebiyatı, dilbilim, Osmanlı edebiyatı, bilgisayar mühendisliği ve yeni medya alanlarında çalışmaktadır. Kültürel bellek kurumlarından görüşülen temsilcilerin ise 9'u kurum yöneticisi, 2'si kütüphaneci, 2'si arşiv uzmanı, 2'si birim sorumlusu ve 1'i öğretim görevlisidir. Araştırmacıların isimleri ile kurum temsilcilerinin çalıştığı kurumların adları kişisel verilerin korunması amacıyla anonimleştirilmiş ve araştırmacılar A1'den A10'a, kurumlar ise K1'den K16'ya kadar kodlanmıştır.

Görüşmeler 8 Şubat–24 Mayıs 2022 tarihleri arasında Zoom<sup>9</sup> çevrimiçi görüşme platformu kullanılarak yapılmıştır. Görüşmeler ortalama 40 dakika sürmüştür ( $\pm 20$  dak.). Tüm görüşmeler katılımcıların onayları ile kayıt altına alınmıştır. Kurumlardan ikisi Zoom üzerinden görüşmeyi kabul etmemiş, kurum temsilcilerinden biriyle yüz yüze, diğeriyle telefon ile görüşme sağlanmış ve bu görüşmelerin ses kayıtları alınmıştır. Bir araştırmacı da telefonla görüşmeyi tercih etmiş ve görüşme kaydedilmiştir.

Sorulara çoğunlukla ayrıntılı yanıtlar verilmiştir. Anlaşılamayan ya da yeterli yanıt alınamayan sorular farklı biçimlerde sorularak katılımcının yanıtını açması sağlanmıştır. Görüşmelerin ses kayıtları çözümlenerek metne aktarılmıştır. Metne aktarılan görüşmeler MaxQDA yazılımı aracılığıyla düzenlenmiştir. Bu aşamada ses kayıtları ve çözümler yazılıma yüklenmiş ve yazılımın “deşifre modu” özelliği ile yeniden gözden geçirilmiş ve çözümlerde saptanan hatalar düzeltilmiştir.

Daha sonra aynı yazılım ile nitel verilerin analizini kolaylaştırmak, güvenilirliği artırmak ve belirli ölçüde verileri nicelleştirmek amacıyla metinlerde kodlamalar yapılmıştır. Kodlama ile ham veriler kavramsallaştırılarak kategorilere ayrılmış ve görüşmelerde öne çıkan temalar saptanmıştır (Neuman, 2010, s. 663). Kodlama sürecinde öncelikle her bir soru için ayrı bir kod açılmış, bu sorulara verilen yanıtlar belirli temalar altında toplanmıştır. Örneğin, Şekil 2'de görülen araştırmacıların kültürel bellek kurumlarından beklentileri bir ana temayı oluşturmakta, açık erişim, toplu indirme gibi beklentiler ise alt temaları oluşturmaktadır. Açık erişim beklentisi olan her araştırmacıya açık erişim kodu, toplu indirme beklentisi olan her araştırmacıya toplu indirme kodu atanmıştır. Tüm kodlamalar bittikten sonra yazılımın “kod sistemi” içinde bulunan “alt kodların istatistiği” özelliği kullanılarak ana temalar altındaki alt temaların istatistiğini çıkarmak ve bu temaların etiketlendiği cümlelere ulaşmak mümkün olmaktadır. Kodlama sürecinde birbirinden çok farklı yanıtların belirli kategoriler altında toplanarak alt temaların belirlenmesi en zorlu aşama olmuştur. Her bir metinde yeni temalar ortaya çıkmış ve geriye dönerek kodlanmış metinler tekrar bu tema çerçevesinde gözden geçirilmiştir. Bu aşama ilk yazar tarafından yapılmıştır.

Kodlama aşaması bittikten sonra görüşme metinlerinin içerik analizi yapılmış, katılımcıların veri olarak koleksiyonlar özelinde paylaştıkları bilgi ve görüşleri ağırlıklı olarak değerlendirilmiş, bu bilgi ve görüşlerden yararlanılarak araştırma sorularına ilişkin çeşitli çıkarımlar yapılmıştır (Gökçe, 2006, s. 19-21). İçerik analizi sürecinde MaxQDA yazılımının yukarıda bahsedilen kod istatistikleri özelliğinin yanı sıra harita özelliğinden yararlanılmıştır (Şekil 3, 4). Önceden kodlanan temalar ve alt temalar MaxQDA ve Excel aracılığıyla görselleştirilmiş, ayrıca kodlar kullanılarak katılımcıların belirli konudaki görüşlerini içeren nitel bulgulara kolaylıkla erişilip analizde hem nicel hem de nitel verilere yer verilmiştir.

---

9 Zoom. <https://zoom.us/>



## Bulgular ve Yorum

Çalışmanın bu bölümünde veri olarak koleksiyonlara yönelik dijital insani bilimler araştırmacılarının ihtiyaç ve beklentileri ile kültürel bellek kurumlarının bu ihtiyaç ve beklentileri ne ölçüde karşıladığı değerlendirilmektedir.

### *Araştırmacıların Karşılaştıkları Zorluklar ve Kültürel Bellek Kurumlarından Beklentileri*

Araştırmacıların verinin bulunması, toplanması ve analiz edilmesi aşamalarında en çok karşılaştıkları zorluğun erişilebilirlik (N=6)<sup>10</sup> olduğu görülmektedir (Şekil 1). Araştırmacılar elde etmek istedikleri veriye ulaşmak için bazen kuruma gitmek zorunda olduklarını, verilere uzaktan erişemediklerini bildirmişlerdir (A6, A7). Erişilebilirlik ile ilgili bir diğer zorluk ise kütüphane katalogları birbirinden bağımsız olduğundan araştırmacının hangi kaynağın hangi kurumda olduğunu bilme zorunluluğu olmasıdır (A5, A7, A8). Bu konu aynı zamanda verilerin dağınık olması (N=2) ile de ilgilidir. Özellikle süreli yayınlarda bazı sayıların bir kurumda, diğer sayıların farklı bir kurumda olması ve bu konuda herhangi bir yönlendirme yapılmaması araştırmacıları zorlamaktadır (A7, A8). Dijital belgelerin çözünürlüğünün düşük olması (N=5) ve bazı kaynaklara ücretli erişilebilmesi (N=5) karşılaşılan diğer önemli zorluklardandır. Araştırmalarda karşılaşılan bir diğer zorluk ise indeksleme hatalarıdır. Özellikle el yazmaları ve Osmanlı arşivlerinde indekslemede yapılan harf hataları aranan veriye ulaşmayı engellemektedir (A2, A3, A5, A10). Üstverilerde hatalar olması kadar yetersiz üstveriler (N=4) de araştırmaları zorlaştırmaktadır. Üstverilerin yetersiz olması hem verilere erişmeyi zorlaştırmakta hem de yetersiz üstveriler araştırmacıların ihtiyaçlarını karşılamamaktadır (A1, A2). Üstverilerde belli bir standart kullanılmaması da araştırmacılar için sorun yaratmaktadır. Farklı kurumlardan veri setleri indirerek çalışmak isteyen araştırmacılar aynı standartlar kullanılmadığı için öncelikle verilerin düzenlenmesi ile uğraşmak zorunda kalmaktadırlar (A9). Her kurum sisteminin ayrı ayrı keşfedilme ihtiyacı (A5), bir kurumda ücretli olan kaynağın başka bir kurumda ücretsiz olması (A10) sorun yaratmaktadır. Araştırma süreçlerinde verilerin farklı formatlara dönüştürülme ihtiyacı, kullanıcı dostu olmayan arayüzler, açılmayan formatlar, optik karakter tanıma (OCR)<sup>11</sup> olanağının olmaması ve belge indirme sayısında sınır bulunması belirtilen diğer zorluklardan bazılarıdır.

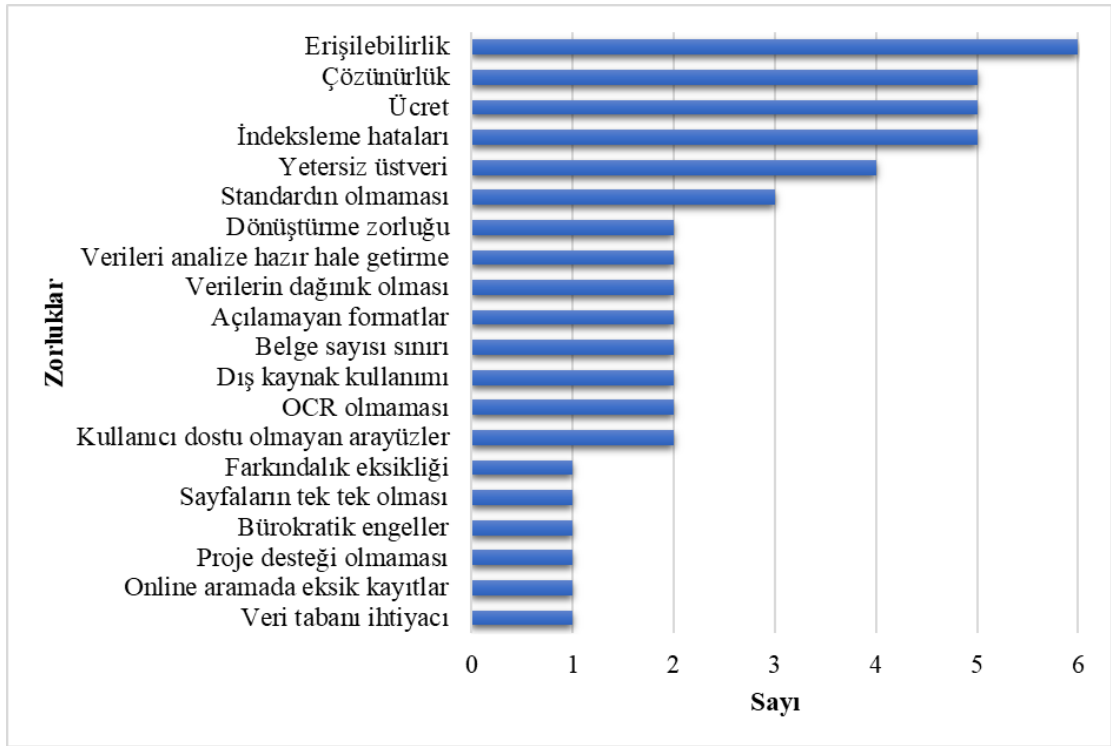
Araştırmacılar en çok (N=6) açık erişime ihtiyaç duymaktadırlar (A1, A5, A7-A10) (Şekil 2). Araştırmacılarından biri açık erişimin önemini şu sözlerle aktarmıştır:

Açık erişim, açık veri, açık fikirlilik. Bu üçü olmazsa dijital tarih, dijital beşeri bilimler imkânsız (A5).

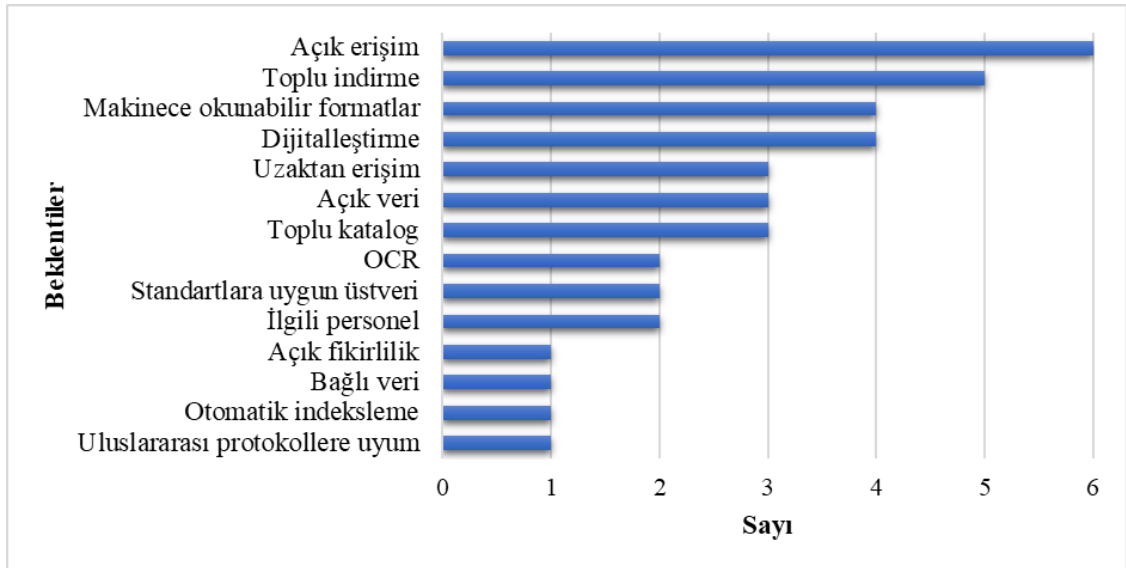
Araştırmacıların kurumlardan beklentilerinde ikinci sırada (N=5) toplu indirme imkânları yer almaktadır (A5, A7-A10). Araştırmacılar koleksiyonlara toplu bir şekilde, tek tıklamayla erişebilmek istemektedirler. Araştırmacıların dijitalleştirme (A2, A3, A5, A8) ve verilerin makinece okunabilir formatlarda sağlanması yönündeki beklentileri de (A5, A8-A10) yüksektir. Makinece okunabilir formatlarda dijitalleştirme için OCR gereklidir (A4-A5).

10 Bu sayı Şekil 1'de erişilebilirlik sorununu vurgulayan katılımcı sayısını göstermektedir. Metnin izleyen kısımlarında şekillerdeki sayılara atıf yapılırken aynı notasyon kullanılmıştır. Spesifik bir araştırmacı tarafından öne sürülen görüşlere araştırmacı kodu ile referans verilmiştir (örneğin, A6).

11 Optical Character Recognition (OCR): Basılı belgelerin ve verilerin makinece düzenlenebilir formatlarda dijital ortama aktarılması.



Şekil 1. Araştırmacıların veri bulma, toplama ve analiz etme aşamalarında karşılaştıkları zorluklar (N=10)



Şekil 2. Dijital insani bilimler araştırmacılarının kültürel bellek kurumlarından beklentileri (N=10)

#### Kültürel Bellek Kurumlarının Dijital Koleksiyonlarına Erişim Yolları

Araştırmacıların beklentileri bazı kurumlar tarafından kısıtlı ölçüde de olsa karşılanmaktadır. Örneğin, görüşülen kurum temsilcilerinden biri bir kullanıcı herhangi bir gazete koleksiyonu ya da bir konuyla ilgili yazma eserler gibi telif kapsamına girmeyen bir koleksiyonun tamamına erişmek isterse bunu sağlayabileceklerini, ayrıca talep edildiğinde koleksiyonlarını makinece okunabilir formatlarda sunabileceklerini söylemiştir (K14). Birçok kurum araştırmacı istekte bulunduğu yüksek çözünürlüklü veya makinece okunabilir formatlarda koleksiyonlarını paylaşabilmektedir. Bazı kurumlar herhangi bir prosedür olmadan istenen formatı araştırmacı

ile paylaşmakta (K9, K12, K13), bazıları ise resmi başvurular uygun bulunursa koleksiyonlarını ücretsiz (K4, K15) ya da belli bir ücret karşılığında (K6) paylaşmaktadır.

Dijital koleksiyonlara erişim sorunu genel olarak ele alındığında kurumların farklı koleksiyonlara yönelik farklı erişim koşulları belirledikleri anlaşılmaktadır. Örneğin, aynı kurumdaki bazı koleksiyonlar tamamen açık bir şekilde tüm kullanıcılarla paylaşılırken, özel koleksiyonlar yalnızca üyelere veya yalnızca kurum içinden erişime açılabilir. Kurumların yarısı (N=8) dijital koleksiyonlarını açık erişim olarak hizmete sunmaktadır (K2-K4, K9, K12-K13, K15-K16). Koleksiyonlarını açık erişimde sunan bu kurumlar bazı koleksiyonlarına ise üyelik ile (K2, K9), ücretli ve/veya istek üzerine (K12, K15) ya da yalnızca kurum içinden (K3, K9) erişim sağlamaktadır.

Yedi kurum koleksiyonlarının bir kısmına erişim için kullanıcılardan ücret almaktadır (K5-K8, K11-K12, K15). Bu kurumlardan birinde mevzuat gereği ücret alındığı belirtilmiştir (K8). Bazı kurumlar ücreti araştırmacının öğrenci/akademisyen olmasına ya da koleksiyonun akademik/ticari amaçla kullanılmasına göre belirlemektedir (K6, K7, K9). Koleksiyonlarını ücretli sunan kurumlardan ikisi filigranlı veya sadece küçük görseller sunarak belgelere kısıtlı bir açık erişim imkânı sağlamaktadır (K6-K7). Bu kurumlar referans verilmeden koleksiyonlarının sosyal medyada paylaşılmasını önlemek amacıyla belgelere filigran eklemektedir.

Üç kurumun (K1, K10, K14) koleksiyonları ise ne açık erişim ne de ücretli olarak erişime açıktır. Bu kurumlar koleksiyonlarının çoğuna üyelikle erişim sağlamaktadır. Bu kurumlardan ikisi (K1, K14) belirli koleksiyonları yalnızca kurum içi erişime açmaktadır. İki kurumda (K10, K14) indirilen ve erişilen belge sayısı sınırlıdır.

Açık ya da ücretli erişim seçeneklerinden bağımsız olarak ele alındığında toplam yedi kurumda üyelik ile erişime açık koleksiyonlar bulunmaktadır (K1-K2, K5, K9-K11, K14). Bazı kurumlarda farklı üyelere farklı erişim izinleri olması (K1) veya istatistiki veri tutma ihtiyacı (K2, K14), indirme sınırlaması (K5, K14) gibi sebeplerle üyelik istenmektedir. Altı kurumda yalnızca kurum içinden erişime açık kaynaklar bulunmaktadır (K1, K3, K5-K6, K9, K14). Altı kurumda ise bazı koleksiyonlar yalnızca araştırmacılar istekte bulunduğu paylaşılmaktadır (K5, K7-K8, K10, K12, K15). Bu kurumlardan üçü koleksiyonların hangi amaçla kullanılacağını açıklayan bir başvuru yapılmasını talep etmektedir (K7, K10, K15).

### ***Araştırmacıların Veri Olarak Koleksiyonlar Konusundaki Görüşleri***

Araştırmacıların veri olarak koleksiyonlara ilişkin görüşleri çoğunlukla olumludur (Şekil 3). Sekiz araştırmacı veri olarak koleksiyonların önemli olduğunu düşünmekte ve bazıları konunun önemini şu sözlerle vurgulamaktadır:

Tüm bir koleksiyona makro ölçekte bakabilmek bambaşka araştırma soruları sağlayacaktır. Bilgi üretimine katkısı açısından çok fark yaratacağını düşünüyorum (A5).

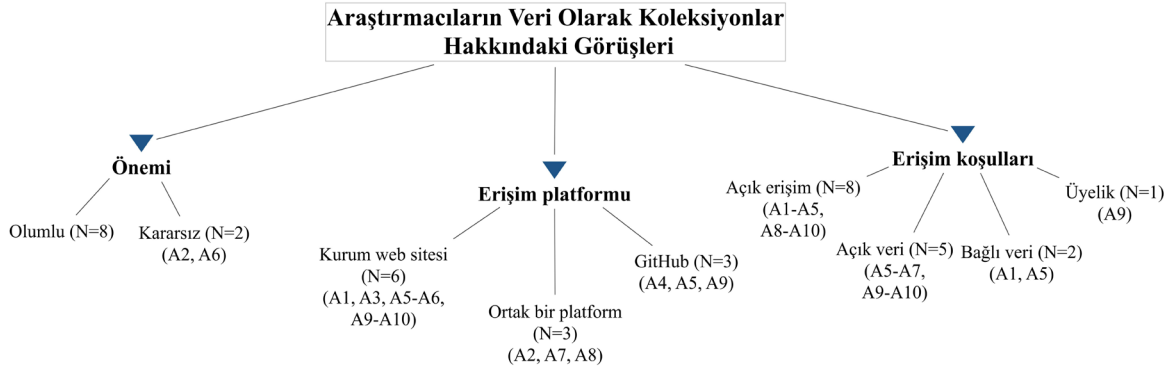
Müthiş bir adım atmış oluruz. Yılda birkaç tane proje yapabiliriz bu durumda (A10).

Araştırma sürecini tamamen otomatize etmemizi sağlamaz ama iş akışımızı çok hızlandırır, emek yoğunluğunu biraz daha azaltıp veri üretimine odaklanabilmemizi sağlar (A1).

Olumsuz görüş bildiren araştırmacı olmasa da iki araştırmacı bu konuda kararsızdır:

Bunu hem bir imkân hem de bir zorluk olarak görüyorum. Makinece okunabilir formatlar herkesin işine yarayacaktır ancak tüm bir koleksiyonu toplu halde indirmek araştırmayı kolaylaştırmak yerine içinden çıkılmaz bir hâle getirebilir diye düşünüyorum (A2).

Muhteşem olur ancak bazı zorlukları var. Bir kaynağın o koleksiyona ait olduğuna veya olmadığına kim karar verecek? Koleksiyonun tam olup olmadığından emin olamam. Konu bazı koleksiyonlar oluşturulursa çok zor (A6).



**Şekil 3.** Araştırmacıların veri olarak koleksiyonlara ilişkin görüşleri (N=10)

*Not:* Araştırmacılar bazı sorularla ilgili olarak birden fazla görüş bildirmişlerdir.

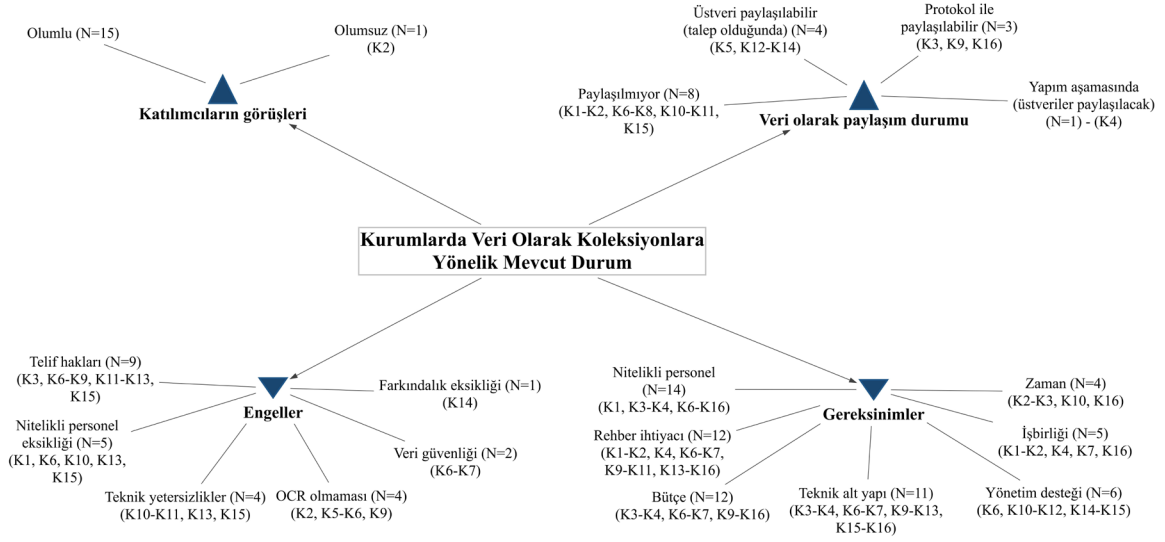
Çoğu araştırmacı (N=6) veri olarak koleksiyonların kurum web sitesi üzerinden erişime açılmasını uygun bulmaktadır. Böylece süreç daha disiplinli ve özenli yürütülebilir (A10). Web sitelerini tek tek gezip kaynak aramak zor olduğundan tüm kültürel bellek kurumları koleksiyonlarını veri olarak paylaşabilecekleri ortak bir platform oluşturmalıdır. Gerekirse Millî Kütüphane bu konuda birleştirici bir rol üstlenebilir (A2). Koleksiyonların veri olarak GitHub üzerinden paylaşılması, hatta bu koleksiyonlara hem GitHub hem de web sitesinden ulaşılabilmesi gerekir (A5, A9).

Araştırmacılar veri olarak koleksiyonların en çok (N=8) açık erişim ve açık veri (N=5) olarak paylaşılmasını istemektedirler. Dahası, bağlı veri (linked data) yaklaşımıyla tüm kültürel bellek kurumlarının koleksiyonlarına erişilebilir ve böylece merkezileşme ihtiyacı ortadan kaldırılabilir (A1, A5). Kamu verisi olmayan koleksiyonlara üyelik ile erişim sağlanabilir (A9).

#### ***Bellek Kurumları Temsilcilerinin Veri Olarak Koleksiyonlar Konusundaki Görüşleri ve Kurumların Mevcut Durumları***

Şekil 4 kurum temsilcilerinin veri olarak koleksiyonlara yönelik görüşlerini, koleksiyonların veri olarak paylaşım durumlarını, veri olarak koleksiyonların önündeki engelleri ve bu konuda kurumların gereksinimlerini göstermektedir.

Görüşme yapılan hemen hemen tüm kurum temsilcileri (N=15) dijital koleksiyonların veri olarak koleksiyonlara dönüştürülmesine olumlu yaklaşmaktadırlar. Sadece bir kurum temsilcisi (K2) kurumun dijital koleksiyonlarının genellikle Arapça veya Osmanlıca olması nedeniyle bu koleksiyonların makinece işlenebilir formatlara dönüştürülmesinin zor olduğunu düşünmektedir. Bazı kurumlar veri olarak koleksiyonlara tamamen olumlu bakmakta (K13, K14), bazıları da kuruma katkısı olacak bir projeye dayalı olarak koleksiyonların paylaşılmasını önemsemektedir (K6, K7). Kurumlardan biri (K12) yasal altyapı oluşturulmasının önemli olduğunu düşünmektedir.



Şekil 4. Kurumların veri olarak koleksiyonlara yönelik mevcut durumları (N=16)

Not: Kurum temsilcileri bazı sorularla ilgili olarak birden fazla görüş bildirmişlerdir.

Mevcut durumda kurumların hiçbiri henüz koleksiyonlarını veri olarak koleksiyonlara dönüştürüp herhangi bir platformdan araştırmacıların hizmetine sunmamıştır. Sekiz kurum koleksiyonlarını veri olarak koleksiyonlar biçiminde paylaşmadıklarını, üç kurum araştırmacılar ile aralarında bir protokol imzalandığı takdirde paylaşabileceklerini belirtmişlerdir -ki bu üç kurumun dijital koleksiyonları herkese açıktır. Dört kurumun talep olursa üstverileri makinece okunabilir formatlarda toplu halde dışarı aktarabilecekleri imkânları bulunmaktadır. Bir kurum ise üstverilerini CSV, XML, JASON, RDF gibi makinece okunabilir formatlarda FAIR<sup>12</sup> ilkelerine uygun biçimde veri olarak koleksiyonlara dönüştürmekte ve GitHub üzerinden paylaşmaya hazırlanmaktadır (K4).

Çoğu kurum temsilcisine (N=9) göre veri olarak koleksiyonlara ilişkin en büyük engel telif haklarıdır. Nitelikli personelin olmaması (N=5), OCR bulunmaması (N=4), teknik yetersizlikler (N=4) ve veri güvenliği (N=2) önemli görülen diğer engeller arasındadır.

Üstverilerini veri olarak koleksiyonlara dönüştüren bir kurumun temsilcisi mevcut durumda herhangi bir engelleri olmadığını ama süreç boyunca nitelikli personel, zaman, bütçe ve teknik altyapı hususlarında zorluklar yaşadıklarını belirtmiştir (K4). Bir başka kurum temsilcisi ise bu konuda farkındalık eksiklikleri dışında bir engellerinin bulunmadığını bildirmiştir (K14):

Aslında dijitalleştirdiğimiz her materyalin txt dosyası elimizde mevcut, bunları sunmak için herhangi bir teknik engelimiz yok, ancak bizde bu konuda bir fikir kaybı oluşmamış. Bunları ziplemek, künye künye ayırmak ciddi anlamda bir iş ama zaten elimizde txt dosyaları var. Kütüphanecilik literatüründe böyle bir olgu çıkarsa yapabiliriz. Veriyi üretme konusunda sıkıntımız yok, sunmak için çalışmamız lazım.

İki kurum (K2, K5) dışında tüm kurumların koleksiyonlarını veri olarak koleksiyonlara dönüştürmek için nitelikli personele ihtiyacı vardır. Benzeri bir biçimde dört kurum (K3, K5, K10, K12) dışında bütün kurumlar rehber ihtiyacı duymaktadır. Bütçe bir diğer önemli gereksinim (N=12) olarak karşımıza çıkmaktadır. On bir kurumun koleksiyonları dönüştürme sürecinde teknik altyapıya, altı kurumun yönetim desteğine,

12 Bulunabilirlik (findability), erişilebilirlik (accessibility), birlikte çalışabilirlik (interoperability), yeniden kullanılabilirlik (reusability) ilkeleri.

dört kurumun zamana ihtiyacı bulunmaktadır.<sup>13</sup> Bir kurum temsilcisi veri olarak koleksiyonlara yönelik gereksinimleri şu şekilde özetlemiştir:

İlk önce yukarıdan başlamak lazım. Öncelikle bürokratik engelleri ortadan kaldıracak bir irade gerekiyor. Bir politika olmalı. Bununla ilgili bir personele ihtiyaç duyulur. Çok fazla problem yaşayacak bir kurum değiliz ama teknik altyapı oluşturulmalı. Bunu planlayacak bir birime ihtiyaç var. Çok büyük bir bütçe ihtiyacı olacağını sanmıyorum ama bütçe de gerekir (K14).

Akademi ve uygulamacılar arasında veya kurum içi birimler arasında işbirliğine gereksinim duyulmaktadır (N=5). Kütüphaneci, arşivci, veri yöneticisi, yazılımcı gibi çeşitli meslek mensuplarının işbirliği yapmasını ve hangi mesleğin hangi aşamada devreye gireceğinin akademisyen ve uygulamacılar tarafından önceden belirlenmesini önemli bulan (K1) ve derleme kütüphanelerinin bir araya gelerek tek bir yerden veri olarak koleksiyonlar hizmeti sunmasını öneren (K2) kurum temsilcileri bulunmaktadır.

## Tartışma

Bu araştırmada Türkiye'deki dijital insani bilimler araştırmacılarının veri olarak koleksiyonlara ilişkin ihtiyaç ve beklentilerini belirlemek ve kültürel bellek kurumlarının mevcut durumlarını saptamak amaçlanmıştır. İlk araştırma sorusu (AS1) çerçevesinde araştırmacıların veri olarak koleksiyonlara ihtiyaç duydukları, verilerin bulunması, toplanması ve analiz edilmesi aşamalarında verilere erişim, çözünürlük, indeksleme hataları, üstverilerin yetersizliği, standartların olmaması, verileri analize uygun biçime dönüştürme zorluğu, verilerin uygun formatlarda sunulmaması, OCR olmaması gibi birçok zorlukla karşılaştıkları görülmüştür. Bu zorluklar literatürde rapor edilen zorluklarla büyük ölçüde örtüşmektedir (örneğin, Wittmann ve diğerleri, 2019). İnsani bilimler çalışmalarının temel zorluklarından biri yüksek kaliteli verilere/üstverilere erişimdir. Bir metindeki sözcük sıklığı analizini gerçekleştirebilmek için bile dosya uzantısı .txt olan düz metin dosyalarına erişebilmeye, yüksek kaliteli bir OCR'a, telif hakkı gibi engellerle karşılaşmadan toplu halde dosyaları indirebilmeye ihtiyaç vardır (Wittmann ve diğerleri, 2019, s. 49-50). Bu zorlukların veri olarak koleksiyonlar yaklaşımının benimsenmesi ile büyük ölçüde kolaylaşacağı düşünülmektedir.

İkinci araştırma sorusu (AS2) doğrultusunda dijital insani bilimler araştırmacılarının kültürel bellek kurumlarından beklentileri tespit edilmiştir. Araştırmacılar en çok verilere/üstverilere makinece okunabilir formatlarda ve açık bir biçimde uzaktan ve toplu indirme seçenekleri ile erişebilmek istemektedirler. Araştırmacılar veriler dağınık olduğu için çeşitli zorluklarla karşılaşmakta ve bir toplu kataloğa ihtiyaç duymaktadırlar. Açık veri, bağlı veri, ilgili personel, OCR ve standartlara uygun üstveriler araştırmacıların diğer önemli beklentilerindedir. Araştırmacılar kültürel bellek kurumlarının koleksiyonlarını veri olarak koleksiyonlara dönüştürerek açık bir şekilde paylaşımlarını beklemekte ve bu konuyu önemsemektedirler. Bu bulgular kültürel bellek kurumlarında çalışanların teknolojik gelişmelere uygun biçimde dijital araştırmalarla ilgili yeni beceriler edinmeleri gerektiğini göstermektedir (Candela ve diğerleri, 2022, s. 1).

Üçüncü araştırma sorusu (AS3) ile kültürel bellek kurumlarının kültürel miras koleksiyonlarını veri olarak koleksiyonlara uygun biçimde sunup sunmadıkları sorusuna yanıt aranmıştır. Bazı kurumlar veri olarak koleksiyonlara ilişkin teknik engelleri olmadığından araştırmacılar istekte bulunduğu protokoller karşılığında verileri ya da üstverileri paylaşabilecek durumdadır. Hemen hemen tüm kurumlar veri olarak koleksiyonların önemli olduğunu düşünmekte, ancak hiçbir kurumun henüz bu yaklaşımı benimsemediği görülmektedir. Nitekim kurumlar var olan dijital koleksiyonlarının tamamını açık bir şekilde hizmete sunmamaktadır. Kurumlarda üyelikle, ücretli, yalnızca kurum içinden, sadece araştırmacı istekte bulunduğu, filigranlı ya da küçük görsel olarak veya sınırlı sayıda belgeye erişilebilen koleksiyonlar yer almaktadır. Halen sadece bir kurum üstverilerini veri olarak koleksiyonlara dönüştürmeye çalışmaktadır.

13 Bir kurum (K5) ise herhangi bir gereksinimden söz etmemiştir.

Dördüncü araştırma sorusu (AS4) kapsamında kültürel bellek kurumlarının koleksiyonlarını veri olarak koleksiyonlara dönüştürmek için gereksinimleri saptanmıştır. Bulgular kurumların önünde telif hakları, personel, OCR, teknik altyapı, veri güvenliği ve farkındalık, bilgi ve deneyim eksikliği gibi engeller olduğunu göstermektedir. Bu ve benzeri engeller (örneğin, üstveri tutarsızlıkları, OCR kalitesi, kurumlarda bilişimsel erişime uygun olmayan geleneksel sistemlerin kullanılması) literatürde de rapor edilmiş, dönüşüm sürecinde veri hazırlığı, veri seti oluşturma, lisanslama, üstveri oluşturma, uzman personel eksikliği, paylaşım platformunun belirlenmesi, depolama ve zaman gibi konularda zorluklarla karşılaşmıştır (Candela ve diğerleri, 2022, s. 1; 2023, s. 5-6). Çalışmamızın bulgularına göre bellek kurumlarının veri olarak koleksiyonlar yaklaşımını hayata geçirebilmeleri için nitelikli personele, bütçeye, teknik altyapıya, yönetim desteğine, işbirliğine, zamana ve konuya ilişkin bir rehber ihtiyaçları vardır. Benzeri gereksinimlere literatürde de yer verilmiş (Padilla ve diğerleri, 2019, s. 91), sürece başlamadan önce uygulama örnekleri, veri hazırlama, kurumsal strateji ve destek, makinece okunabilir üstveri, depolama, GDPR (Avrupa Birliği Genel Veri Koruma Tüzüğü) ve veri seçimi hakkında bilgi ve bilişimsel uzmanlık sahibi olmanın süreci kolaylaştıracağı vurgulanmıştır (Candela ve diğerleri, 2023, s. 7).

Elde ettiğimiz tüm bulgular araştırma hipotezini (H1) büyük ölçüde desteklemektedir: “Türkiye’deki kültürel bellek kurumlarının veri olarak kültürel miras koleksiyonlarına yönelik mevcut uygulama ve süreçleri dijital insani bilimler araştırmacılarının gereksinim ve beklentilerini karşılamamaktadır”.

Amaçlı örneklem tekniğiyle seçilen ve kısıtlı sayıda dijital insani bilimler araştırmacısı ve bellek kurumu temsilcisi ile yapılan görüşmelere dayanan bulgular ile genelleme yaparken dikkatli olmak gerektiğinin farkındayız. Ancak dijital insani bilimler araştırmacılarının veri yoğun araştırma yapan kişiler arasından, kurum temsilcilerinin ise bir dereceye kadar “iyi uygulama örnekleri”ne sahip kurumlardan seçildiği gözden uzak tutulmamalıdır. Başka bir deyişle, birçok kurumun henüz dijital koleksiyonlara sahip olmadığı için veri olarak koleksiyonlar yaklaşımını hayata geçirmeye hazırlıklı olmadığını söylemek mümkündür. Öte yandan, bu araştırmanın bulguları genel olarak önceki çalışmaların bulgularını destekler nitelikte olup dijital koleksiyonlarını veri olarak koleksiyonlara dönüştürmeyi planlayan kültürel bellek kurumları açısından yararlı olabilir.

## Sonuç ve Öneriler

Dijital insani bilimler araştırmalarına yönelik ilginin artmasıyla birlikte kültürel bellek kurumlarının koleksiyonlarını veri odaklı araştırmalara uygun biçimde sunmaları bir zorunluluk hâline gelmiştir. Araştırmamız kapsamında görüşülen kurumların henüz hiçbiri toplu indirme seçenekleri ve makinece okunabilir ve işlem yapılabilir formatlarda veriler sunmamaktadır. Öte yandan dijital koleksiyonların üstverileri genellikle yetersiz ve standartlara uygun değildir. Bu kapsamda öncelikle mevcut dijital koleksiyonların mümkün olduğunca tamamının çözünürlüğü yüksek formatlarda ve açık erişim olarak hizmete sunulması önerilmektedir. Sonraki aşama ise kurumların üstverilerini veya dijital koleksiyonlarını bilişimsel erişime ve makinece işlem yapmaya uygun biçimde hizmete sunmalarıdır. Bu aşamada hem mevcut dijital koleksiyonların veri olarak koleksiyonlara dönüştürülmesi hem de veri olarak koleksiyonların dijitalleştirme sürecinin bir parçası olarak ele alınması önemlidir.

Koleksiyonların veri olarak koleksiyonlara dönüştürülmesinin önündeki engellerin aşılması için kurumlarda veri olarak koleksiyonlardan sorumlu personel belirlenerek bu personele hizmet içi eğitim verilmesi yararlı olacaktır. Koleksiyonların makinece okunabilir formatlarda dijitalleştirilerek veri olarak koleksiyonlara dönüştürülmesi gerekmektedir. Bu bakımdan akademisyenler ve uygulamacılar ile fikir alışverişlerinde bulunulmasında ve diğer kurumlarla işbirliğine gidilmesinde yarar görülmektedir. İlk aşamada araştırmacılar

ile birlikte çalışarak onların ihtiyaçları çerçevesinde koleksiyonları veri olarak koleksiyonlara dönüştürmek hem süreci öğrenmek hem de veri olarak koleksiyonların bilimsel araştırmalara katkısını görebilmek açısından faydalı olacaktır. Dönüşüm sürecinde ilk aşamada projelendirme ya da dış kaynak kullanımı gibi seçenekler düşünülebilir. Ancak sonraki aşamalarda veri olarak koleksiyonlar dijitalleştirilmenin bir parçası olarak ele alınmalı ve dijitalleştirme sürecinde buna uygun bir yol izlenmelidir. Kurumlar veri olarak koleksiyonlara geçiş aşamasında daha önce koleksiyonlarını dönüştürmüş olan kurum örneklerinden, önceki bölümlerde değinilen kontrol listelerinden ve 50 adımdan oluşan yol haritasından yararlanabilirler (Padilla ve diğerleri, 2019; Candela ve diğerleri, 2023; Lee, 2023).

Konuya ilişkin nitelikli personel yetiştirilmesi ve farkındalık oluşturulması ile ilgili olarak Bilgi ve Belge Yönetimi (BBY) bölümlerinin de sorumluluk yüklenmesi gerektiği kanısındayız. Bu bölümlerin veri olarak koleksiyonlar için gerekli becerilere sahip mezunlar yetiştirmesi önemlidir. Bu kapsamda ders programlarında makinece okunabilir ve işlem yapılabilir koleksiyonlar ve veriler üretmek, bunları düzenlemek ve görselleştirmek için gereken teknik bilgilerin yanı sıra telif hakları, lisanslama, dijital insani bilimler araç ve yöntemleri gibi konulara da yer verilmesi gerektiği kanısındayız. Ayrıca, ulusal literatürde konuyla ilgili çalışmaların yayımlanması farkındalığın artmasına ve dijital insani bilimler araştırmacılarının ihtiyaçlarının daha kısa sürede karşılanmasına yardımcı olur. Dijital insani bilimler disiplinine kültürel bellek kurumlarının sağlayabileceği katkıların göz ardı edilmemesi, bu konuda hem akademik hem de uygulamalı çalışmalar yapılması gerektiği kanısındayız. Sonraki çalışmalarda koleksiyonlarını bilişimsel erişime uygun olarak erişime açmak isteyen kurumlara yol gösterici rehberler hazırlanması ve örnek koleksiyonlara dayalı veri odaklı çalışmalar gerçekleştirilerek ihtiyaçların daha somut bir şekilde saptanması önerilmektedir.

## Teşekkür

Bu araştırmaya yaptıkları değerli katkılar için görüştüğümüz dijital insani bilimler araştırmacılarına ve kurum temsilcilerine teşekkür ederiz.

## İzin ve Katkı Bildirimleri

**Etik Kurul İzni:** Hacettepe Üniversitesi Senatosu Etik Komisyonununun 08 Şubat 2022 tarihli kararı.

**Yazarlık Katkısı:**

M. Dişli: Kavramsallaştırma, yöntem, veri toplama, veri analizi, görselleştirme; özgün ve gözden geçirilmiş taslakların yazılması ve düzenlenmesi

Y. Tonta: Kavramsallaştırma, yöntem, taslakları gözden geçirme ve düzeltme, son taslak, denetleme

## Kaynakça

Adamlink Project. (2022). <https://adamlink.nl/>

Akça, S. (2018). Dijital insanî bilimler: Yeni bir yaklaşım. *Türk Kütüphaneciliği*, 32(3), 193-207. <https://doi.org/10.24146/tkd.2018.37>

Aladağ, F. (2020). *Cities and administrative divisions of the Ottoman Empire in the early 16th century: a case study for the application of digital history to Ottoman studies*. Yüksek Lisans Tezi. İstanbul Şehir Üniversitesi, İstanbul. <http://t.ly/NsOV1>

Ames, S. (2021). Transparency, provenance and collections as data: The National Library of Scotland's Data Foundry. *Liber Quarterly*, 31, 1-13. Doi: 10.18352/LQ.10371

Amsterdam Time Machine. (2022). <https://www.amsterdamtimemachine.nl/about-atm/>



- ANAMED Kütüphanesi. (2021). *Osmanlı Epir'ini haritalandırmak*. A. Yayıoğlu, A. Atçıl, F. Öncel. [https://www.youtube.com/watch?v=AUAtSB43p2s&ab\\_channel=ANAMEDLibrary](https://www.youtube.com/watch?v=AUAtSB43p2s&ab_channel=ANAMEDLibrary)
- Bekman, B. (2020). XIX. yüzyılın dönüşen İstanbul'unda mahalleler, kollar ve belediyeler: bir envanter ve haritalama çalışması. Yüksek Lisans Tezi. Marmara Üniversitesi, İstanbul. [http://t.ly/hV1V\\_](http://t.ly/hV1V_)
- Bell, G., Zafiroglu, A., Assaad, Z., Bradley, C., Cooper, N., O'Brien, E., Reid, K. ve Ruster, L. (2021). *Custodians & midwives: The library of the future*. The Australian National University. [https://cybernetics.anu.edu.au/assets/Custodians\\_and\\_Midwives\\_FINAL\\_web.pdf](https://cybernetics.anu.edu.au/assets/Custodians_and_Midwives_FINAL_web.pdf)
- Ben-Bassat, Y. ve Buessow, J. (2020). Applying digital methods to the study of a late Ottoman city: A social and spatial analysis of political partisanship in Gaza. *Journal of the Economic and Social History of the Orient*, 63(4), 505-554. Doi: 10.1163/15685209-12341517
- Berry, D. (2019). *What are the digital humanities?* The British Academy. <https://www.thebritishacademy.ac.uk/blog/what-are-digital-humanities/>
- Bozkurt, A., Hamutoğlu, N. B., Kaban, A. L., Taşçı, G. ve Aykul, M. (2021). Dijital bilgi çağı: Dijital toplum, dijital dönüşüm, dijital eğitim ve dijital yeterlilikler. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 7(2), 35-63. Doi: 10.51948/auad.911584
- British Library. (2022a). *British Library datasets*. [https://data.bl.uk/bl\\_labs\\_datasets/](https://data.bl.uk/bl_labs_datasets/)
- British Library. (2022b). *British Library Labs*. <https://www.bl.uk/projects/british-library-labs>
- British Library. (2023). *Free dataset downloads*. <https://www.bl.uk/collection-metadata/downloads>
- Candela, G., Gabriëls, N., Chambers, S., Pham, T. A., Ames, S., Fitzgerald, N., Hofmann, K., Harbo, V., Potter, A., Ferriter, M., Manchester, E., Irollo, A., Van Keer, E., Mahey, M., Holownia, O. ve Dobрева, M. (2023). A checklist to publish collections as data in GLAM institutions. (*arXiv:2304.02603*). arXiv. <http://arxiv.org/abs/2304.02603>
- Candela, G., Sáez, M. D., Escobar Esteban, M. ve Marco-Such, M. (2022). Reusing digital collections from GLAM institutions. *Journal of Information Science*, 48(2), 251-267. Doi: 10.1177/0165551520950246
- Chambers, S., Walsh, M., Caswell, M., Harder, G., Okumura, M., Corrin, J., Baeza Ventura, G., Antonijevic, S., Knazook, B., Narlock, M., Bailey, J., Neudecker, C., Downie, J. S., Layne-Worthey, G., van Strien, D., Irollo, A., Whitmire, A., Lee, J., Berry, D., ... Ridge, M. (2023). *Position statements: Collections as data: State of the field and future directions*. Doi: 10.5281/zenodo.7897735
- Colavizza, G., Blanke, T., Jeurgens, C. ve Noordegraaf, J. (2021). Archives and AI: An overview of current debates and future perspectives. *Journal on Computing and Cultural Heritage*, 15(1), 1-15. Doi: 10.1145/3479010
- Collections as Data Futures: A recap, a resource, next steps. (2023). *Collections as Data - Part to Whole*. <https://collectionsasdata.github.io/part2whole/recap/>
- Çayırözmez, N. A. (2023). *Türkiye'de dijital arkeolojik veri yönetimi*. Doktora Tezi. Ankara Üniversitesi, Ankara.
- Digital Humanities Lab. (2022). *DHLab*. <https://dhlab.nl/>
- Digital Ottoman Studies. (2022). <https://www.digitalottomanstudies.com/about-us?lang=tr>
- Digital Preservation Coalition. (2022). *Computational access: A beginner's guide for digital preservation practitioners*. <https://www.dpconline.org/digipres/implement-digipres/computational-access-guide>
- Ekim, B., Sertel, E. ve Kabadayı, M. E. (2021). Automatic road extraction from historical maps using deep learning techniques: A regional case study of Turkey in a German World War II Map. *ISPRS International Journal of Geo-Information*, 10(8), 492. Doi: 10.3390/ijgi10080492

- Europeana Pro. (2022). *Harvesting and downloads*. <https://pro.europeana.eu/page/harvesting-and-downloads>
- Gökçe, O. (2006). İçerik analizi: *Kuramsal ve pratik bilgiler*. Siyasal Kitabevi.
- Gökmen, B. (2015). Türkiye’de kadı isimli yerleşmelerin mekânsal analizi: Kadılık kurumunun yerleşme isimleri üzerindeki etkisi. Çankırı Karatekin Üniversitesi Karatekin Edebiyat Fakültesi Dergisi, 5(5), 39-64. <https://dergipark.org.tr/tr/pub/karefad/issue/19200/204140>
- Grallert, T. (2016). Mapping Ottoman Damascus through news reports: A practical approach. *The Digital Humanities and Islamic & Middle East Studies* içinde (s. 175-198). De Gruyter. Doi: 10.1515/9783110376517-009
- Gray, J. (2009). Jim Gray on e-science: A transformed scientific method. Hey, T., Tansley, S. ve Tolle, K. (Yay. haz.). *The Fourth Paradigm: Data-Intensive Scientific Discovery* içinde (s. xvii-xxxi). Microsoft Research. [https://www.microsoft.com/en-us/research/wp-content/uploads/2009/10/Fourth\\_Paradigm.pdf](https://www.microsoft.com/en-us/research/wp-content/uploads/2009/10/Fourth_Paradigm.pdf)
- Hadjikyriacou, A. (2018). *Introduction: Ottoman Recogito (OttRec)*. <https://medium.com/pelagios/introduction-ottoman-recogito-ottrec-3157478130fd>
- Hasan Ağa ve Kardeşi Anton Çelebi’nin Aşırı Tuhaf Hikayesi. (2022). <https://dhlabsabanciuniv.edu/tr/hasan-aga-ve-kardesi-anton-celebinin-asiri-tuhaf-hikayesi>
- HathiTrust. (2022). *HathiTrust Digital Library: Millions of books online*. <https://www.hathitrust.org/>
- HisGis. (2022). <https://hisgis.nl/>
- Johnston, L. (2011). *From records to data: It’s not just about collections any more*. <https://blogs.loc.gov/thesignal/2011/11/from-records-to-data-its-not-just-about-collections-any-more/>
- Karbeyaz, C., Can, E. F., Can, F. ve Kalpaklı, M. (2012). A content-based social network study of Evliyâ Çelebi’s Seyahatnâme-Bitlis section. *Computer and Information Sciences II*, 271-275. Doi: 10.1007/978-1-4471-2155-8\_34
- Karbeyaz, C., Can, E. F., Can, F. ve Kalpaklı, M. (2013). A social network study of Evliyâ Çelebi’s The Book of Travels-Seyahatnâme. *ENRICH 2013: The First Workshop on the Exploration, Navigation and Retrieval of Information in Cultural Heritage. August 1, 2013, Dublin*. <http://www.cs.bilkent.edu.tr/~canf/bookOfTravels.pdf>
- Koninklijke Bibliotheek. (2022). *KB Lab*. <https://lab.kb.nl/about-us>
- Lampert, C. ve Lapworth, E. (2020). *What do we mean by “Collections As Data” (CAD)?* <https://www.library.unlv.edu/whats-new-special-collections/2020/2020-03/what-do-we-mean-collections-data-cad-cory-lampert-emily>
- Lanzolla, G. ve Anderson, J. (2008). Digital transformation. *Business Strategy Review*, 19(2), 72-76. Doi: 10.1111/j.1467-8616.2008.00539.x
- Lee, B. C. G. (2023). The “Collections as ML Data” checklist for machine learning and cultural heritage. *Journal of the Association for Information Science and Technology*. Doi: 10.1002/asi.24765
- Library of Congress. (2016). *Collections as data: Stewardship and use models to enhance access*. <https://digitalpreservation.gov/meetings/dcs16.html>
- Library of Congress. (2022). *Search results from selected datasets, available online*. <https://www.loc.gov/collections/selected-datasets/>

- Mahey, M., Al-Abdulla, A., Ames, S., Bray, P., Candela, G., Chambers, S., Derven, C., Dobрева-McPherson, M., Gasser, K., Karner, S., Kokegei, K., Laursen, D., Potter, A., Straube, A., Wagner, S-C. ve Wilms, L. (2019). *Open a GLAM Lab*. Digital Cultural Heritage Innovation Labs, Book Sprint, Doha, Qatar, 23-27 September 2019. <https://glamlabs.pubpub.org/>
- Miksa, T., Neish, P., Walk, P. ve Rauber, A. (2018). Defining requirements for machine-actionable Data Management Plans. iPres. <http://ifs.tuwien.ac.at/~miksa/papers/2018-iPres-maDMPs.pdf>
- National Library of Australia. (2022). *Trove: Using the API*. <https://trove.nla.gov.au/about/create-something/using-api>
- National Library of Australia. (2023). *Trove: Case studies*. <https://trove.nla.gov.au/about/create-something/case-studies>
- Neuman, W. L. (2010). *Toplumsal araştırma yöntemleri: Nitel ve nicel yaklaşımlar*. Yayın Odası.
- Öner, O. (2019). *A collective biography study of musicians: patterns, networks and music as a "profession" in the late Ottoman era and the early republican years in İstanbul*. Doktora Tezi. İstanbul Şehir Üniversitesi, İstanbul.
- Padilla, T. (2016). Humanities data in the library: Integrity, form, access. *D-Lib Magazine*, 22(3/4). Doi: 10.1045/march2016-padilla
- Padilla, T., Allen, L., Frost, H., Potvin, S., Roke, E. R. ve Varner, S. (2019). *Final report: Always already computational: Collections as data*. Doi: 10.5281/zenodo.3152935
- Padilla, T. ve Higgins, D. (2014). Library collections as humanities data: The facet effect. *Public Services Quarterly*, 10(4), 324-335. Doi: 10.1080/15228959.2014.963780.
- Padilla, T. ve Ridge, M. (2019). *Collections as data*. HILT. <https://dhtraining.org/hilt/course/collections-as-data-2019/>
- Posner, M. (2015). *Humanities data: A necessary contradiction*. <http://miriamposner.com/blog/humanities-data-a-necessary-contradiction/>
- Royal Danish Library. (2022). *KB Labs*. <https://labs.kb.dk/>
- Rueda, D. M. (2018a). *Collections as data: Rethinking the digital archive in the digital age*. [https://www.library.unlv.edu/whats\\_new\\_in\\_special\\_collections/2018/07/collections-data-rethinking-digital-archive-digital-age](https://www.library.unlv.edu/whats_new_in_special_collections/2018/07/collections-data-rethinking-digital-archive-digital-age)
- Rueda, D. M. (2018b). *Collections as data: The process (and complications) of data visualization*. [https://www.library.unlv.edu/whats\\_new\\_in\\_special\\_collections/2018/07/collections-data-process-and-complications-data](https://www.library.unlv.edu/whats_new_in_special_collections/2018/07/collections-data-process-and-complications-data)
- Sabancı Üniversitesi Dijital Beşeri Bilimler Laboratuvarı. (2022). *DH Lab*. <https://dhlabs.sabanciuniv.edu/tr>
- Şavk, S. ve Doğu, B. (2021). Mapping Yeşilçam: A relational approach to the Turkish film industry. *CINEJ Cinema Journal*, 9(2), 193-230. Doi: 10.5195/cinej.2021.425
- Schmidt, B. (2012). *Women in the libraries*. Sapping Attention. <http://sappingattention.blogspot.com/2012/05/women-in-libraries.html>
- Sherratt, T. (2019). *From collection search to collections as data*. Heritage of the Air, Canberra. <https://doi.org/10.5281/zenodo.3551405>

- Sherratt, T. (2022). *Welcome to the GLAM Workbench*. <https://glam-workbench.net/>
- Stanford University. (2021). *Mapping Ottoman Epirus*. <https://mapoe.stanford.edu/>
- The Real Face of White Australia. (2022). *Living under the White Australia Policy*. <https://www.realfaceofwhiteaustralia.net/>
- The Santa Barbara Statement on Collections as Data (v2). (2018). Always Already Computational - Collections as Data. <https://collectionsasdata.github.io/statement/>
- Tonta, Y. ve Al, U. (2008). Cumhuriyet Dönemi Makaleler Bibliyografyası 1923-1999: Eleştirel bir değerlendirme. *Türk Kütüphaneciliği*, 22(3), 320-334. <https://dergipark.org.tr/pub/tk/624224>
- Tonta, Y. ve Madran, O. (2021). *Akademiden Notlar 42: Veri olarak dermeler*. <https://www.youtube.com/watch?v=IHBkpvVKT7M>
- Turco, M.L., Calvano, M. ve Giovanni, E.C. (2019). Data modeling for museum collections. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W9, 2019* (s. 433-440). 8th International Workshop 3D-ARCH "3D Virtual Reconstruction and Visualization of Complex Architectures", 6-8 February 2019, Bergamo, Italy. <https://isprs-archives.copernicus.org/articles/XLII-2-W9/433/2019/isprs-archives-XLII-2-W9-433-2019.pdf>
- Türkiye'yi Temsil Eden Kurucu Kuşaklar (2022). <https://dhlabsabanciuniv.edu/tr/turkiyeyi-temsil-eden-kurucu-kusaklar>
- Türnüklü, A. (2000). Eğitim bilim araştırmalarında etkin olarak kullanılacak nitel bir araştırma tekniği: Görüşme. *Kuram ve Uygulamada Eğitim Yönetimi*, 24, 543-559. <https://dergipark.org.tr/en/download/article-file/108517>
- UC San Diego Library. (2021). *Collections as data*. <http://suave.sdsc.edu/about/>
- Uğur, Y. (2014). *The historical interaction of the city with its mahalles: Ottoman Edirne in the late seventeenth and early eighteenth centuries*. Doktora Tezi. Boğaziçi Üniversitesi, İstanbul. <http://t.ly/V0pYB>
- Uğur, Y. (2018). *Osmanlı şehirlerini haritalamak: Sosyo-mekânsal benzerlikler ve özgünlükler (1520-1540)*. TÜBİTAK 1001 - Türkiye Bilimsel ve Teknolojik Araştırma Kurumu.
- University of St Andrews. (2021). *The Islamisation of Anatolia Project*. <https://www.islam-anatolia.ac.uk/>
- Wittmann, R., Neatrou, A., Cummings, R. ve Myntti, J. (2019). From digital library to open datasets. *Information Technology and Libraries*, 38(4), 49-61. Doi: 10.6017/ital.v38i4.11101
- Yıldırım, A. ve Şimşek, H. (2018). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin Yayıncılık.
- Yüksek, A. Y. (2021). Sufis and the Sufi Lodges in Istanbul in the late nineteenth century: A socio-spatial analysis. *Journal of Urban History*. Doi: 10.1177/00961442211025253
- Ziegler, S. L. (2020). Open data in cultural heritage institutions: Can we be better than data brokers? *Digital Humanities Quarterly*, 14(2). <http://www.digitalhumanities.org/dhq/vol/14/2/000462/000462.html>

### Ek 1: Veri olarak kültürel miras koleksiyonları: Kurumlar için görüşme soruları

1. Hangi kurumda görev yapıyorsunuz?
2. Hangi birimde çalışıyorsunuz?
3. Hangi pozisyonda çalışıyorsunuz?
4. Kurumunuzda bulunan dijital koleksiyonları kısaca tanıtabilir misiniz?
  - Osmanlı arşivleri
  - Gazete, fotoğraf arşivleri
  - Haritalar
  - Tıbbi koleksiyonlar
  - Coğrafi veriler
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
5. Dijital koleksiyonlarınızda çoğunlukla hangi tür veriler bulunmaktadır?
  - Metin, ses, video, görüntü vb.
6. Koleksiyonlarınız hangi formatlarda dijitalleştirildi?
7. Bu koleksiyonları uzun dönemli korumak amacıyla hangi formatlarda saklıyorsunuz?
8. Dijital koleksiyonlarınızı kullanıcılarla paylaşıyor musunuz? Cevabınız “Evet” ise süreç hakkında bilgi verebilir misiniz?
  - Kurum içi erişim
  - Üyelik ile erişim
  - Ücret karşılığında paylaşım
  - Araştırmacı istekte bulunduğu paylaşım
  - Açık erişim
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
9. Hassas ve kişisel verileri paylaşıyor musunuz? Paylaşıyorsanız süreç hakkında bilgi verebilir misiniz?
10. Veri güvenliğini ve gizliliğini sağlamak için neler yapıyorsunuz?
11. Kullanıcılarınız dijital koleksiyonlarınıza nasıl erişiyorlar?
  - Arama motorları, kurum web sayfası, kurumdaki bilgisayarlardan, kurumsal arşiv vb.
12. Dijital koleksiyonlarınız üzerinde araştırma yapmak isteyen araştırmacılara hangi konularda yardımcı oluyorsunuz?
13. Bu araştırmanın konusu olan “veri olarak koleksiyonlar”, dijital koleksiyonların veya üstverilerin bilgisayar destekli hesaplamalara uygun veri kümelerine dönüştürülmesi ve paylaşılması şeklinde tanımlanmaktadır. Buna göre veri olarak koleksiyonlar sizce önemli midir? Veri olarak koleksiyonların paylaşılmasının araştırmacılara ve kurumlara ne gibi katkıları olabilir?
14. Sizin kurumunuzda dijital koleksiyonlar veri olarak paylaşılıyor mu?
15. Evet \_\_\_\_\_ Hayır \_\_\_\_\_ (Lütfen 18. soruya gidiniz.)
16. Veri olarak dijital koleksiyonlarınızı nasıl ve hangi koşullarda paylaşıyorsunuz?
17. Veri olarak dijital koleksiyonlarınızı paylaşırken ne tür sorunlarla ya da engellerle karşılaşıyorsunuz?
18. Veri olarak dijital koleksiyonlarınızı FAIR ilkelerine -bulunabilir, erişilebilir, birlikte çalışabilir, yeniden kullanılabilir- uygun mudur? Lütfen açıklayınız. (Lütfen 22. soruya gidiniz.)
19. Koleksiyonlarınızı veri olarak koleksiyonlara dönüştürme ve paylaşmanın önündeki engeller nelerdir?
20. Koleksiyonlarınızı veri olarak koleksiyonlara dönüştürecek olsanız öncelikle hangi koleksiyonları dönüştürürdünüz?

21. Koleksiyonlarınızı veri olarak koleksiyonlara dnřtrme konusunda gereksinimleriniz nelerdir?
  - Personel
  - Ynetim desteęi
  - Teknik altyapı
  - Bte
  - Dięer (Ltfen belirtiniz): \_\_\_\_\_
22. Koleksiyonları veri olarak koleksiyonlara dnřtrmede bir rehber ihtiyaa duyuyor musunuz? Ltfen aıklayınız.
23. Arařtırma ile ilgili sizin eklemek istedięiniz grř ve nerileriniz nelerdir?
24. Bu arařtırmanın sonularından haberdar edilmek ister misiniz? Cevabınız "Evet" ise e-posta adresiniz:  
\_\_\_\_\_

## Ek 2: Veri olarak kültürel miras koleksiyonları: Araştırmacılar için görüşme soruları

1. Hangi kurumda görev yapıyorsunuz?
2. Çalışma alanınız nedir?
3. Araştırmalarınızda kullandığınız veri kaynakları nelerdir?
  - Osmanlı arşivleri
  - Gazete arşivleri
  - Özel koleksiyonlar
  - Sicil kayıtları
  - Tarihi haritalar
  - Genomik veriler
  - Coğrafi veri
  - Laboratuvar verileri
  - Diğer (Lütfen belirtiniz) : \_\_\_\_\_
4. Araştırmalarınızda hassas veya kişisel verilerden yararlanıyor musunuz? Bu verilere nereden ulaşıyor ve bu verileri nasıl kullanıyorsunuz?
5. En çok kullandığınız veri türleri nelerdir?
  - Metin
  - Ses
  - Görüntü
  - Video
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
6. Araştırmalarınızda çoğunlukla hangi formatlarda veriye ihtiyaç duyuyorsunuz?
  - csv
  - xls
  - txt
  - sav
  - xml
  - tiff
  - fastq
  - vcf
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
7. Araştırmalarınızda kullandığınız verilerin ortalama boyutu ne kadardır? (MB, GB, TB...)
8. Araştırmalarınızda kullandığınız analiz araçları nelerdir?
9. Araştırmalarda kullandığınız verilere nerelerden erişiyorsunuz?
  - Kütüphane
  - Arşiv
  - Müze
  - Kurumsal arşiv
  - Ticari arşivler
  - Veri tabanları
  - Biyobankalar
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
10. Araştırmalarınızda kullandığınız verileri nasıl buluyor ve elde ediyorsunuz? (arama motorları, kataloglar aracılığıyla; tek tek ya da topluca indirerek)

11. Araştırmalarınızda ihtiyaç duyulan verinin bulunması, toplanması ve analiz edilmesi aşamalarında ne tür zorluklarla karşılaşyorsunuz?
12. Kültürel bellek kurumlarında (kütüphaneler, arşivler, müzeler) verilerin okunabilirliğini ve analizini olumsuz etkileyen uygulamalar var mı? Varsa açıklayınız.
13. Araştırmalarınızda hiç kütüphane, arşiv veya müze çalışanlarından destek aldınız mı? Aldıysanız açıklayınız.
14. Araştırmalarınızda kültürel bellek kurumlarından beklentileriniz nelerdir?
15. Kültürel bellek kurumlarının koleksiyonlarını toplu halde indirme imkânlarıyla açık ve makinece okunabilir formatlarda paylaşması araştırmalarınızı kolaylaştırır mı? Evet ise açıklayınız.
16. Sizce kurumlar ne tür koleksiyonları makinece okunabilir formatlarda ve toplu indirme imkânları ile paylaşmalıdır? Bu koleksiyonların seçimi neye göre yapılmalıdır?
  - Tarihi değeri olan koleksiyonlar
  - Kültürel değeri olan koleksiyonlar
  - Ekonomik değeri olan koleksiyonlar
  - Sağlık hizmetlerinin gelişmesine katkıda bulunan koleksiyonlar
  - Kullanıcılar tarafından en sık kullanılan koleksiyonlar
  - Araştırmacının ihtiyacı olan koleksiyonlar
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
17. Kurumlar koleksiyonlarını veri olarak nerelerde paylaşmalıdır?
  - GitHub
  - OSF
  - Aperta
  - Zenodo
  - Kurum web sayfası
  - Kurumsal arşiv
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
18. Kurumlar koleksiyonlarını hangi koşullarda paylaşmalıdır?
  - İstekte bulunan araştırmacılara özel
  - Üyelik ile
  - Ücretli
  - Açık veri
  - Diğer (Lütfen belirtiniz): \_\_\_\_\_
19. Araştırma ile ilgili eklemek istediğiniz görüş ve önerileriniz nelerdir?
20. Bu araştırmanın sonuçlarından haberdar edilmek ister misiniz? Cevabınız “Evet” ise e-posta adresiniz: \_\_\_\_\_