



Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Bilgi ve Belge Yönetimi Anabilim Dalı

**ARAŞTIRMA VERİLERİNİN YÖNETİMİ: TÜBİTAK
PROJELERİNDE YER ALAN ARAŞTIRMACILAR ÜZERİNE BİR
DEĞERLENDİRME**

Canan TAVLUOĞLU

Doktora Tezi

Ankara, 2022

ARAŐTIRMA VERİLERİNİN YÖNETİMİ: TÜBİTAK PROJELERİNDE YER ALAN
ARAŐTIRMACILAR ÜZERİNE BİR DEĞERLENDİRME

Canan TAVLUOĐLU

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü
Bilgi ve Belge Yönetimi Anabilim Dalı

Doktora Tezi

Ankara, 2022

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinleri yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan **“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ay ertelenmiştir. ⁽²⁾
- Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

01/07/2022

Canan TAVLUOĞLU

1“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu** iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulunun** gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, **tezin yapıldığı kurum** tarafından verilir *. Kurum ve kuruluşlarla yapılan iş birliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, **ilgili kurum ve kuruluşun önerisi** ile **enstitü** veya **fakültenin** uygun görüşü üzerine **üniversite yönetim kurulu** tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

* Tez **danışmanın** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu tarafından karar verilir.**

ETİK BEYAN

Bu alıřmadaki bütn bilgi ve belgeleri akademik kurallar erevesinde elde ettiđimi, grsel, iřitsel ve yazılı tm bilgi ve sonuları bilimsel ahlak kurallarına uygun olarak sunduđumu, kullandıđım verilerde herhangi bir tahrifat yapmadıđımı, yararlandıđım kaynaklara bilimsel normlara uygun olarak atıfta bulunduđumu, tezimin kaynak gsterilen durumlar dıřında zgn olduđunu, **Prof. Dr. S. Serap KURBANOĐLU** danıřmanlıđında tarafımdan retildiđini ve Hacettepe niversitesi Sosyal Bilimler Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.

Canan TAVLUOĐLU

TEŞEKKÜR

Değerli zamanını, engin bilgisini ve emeğini hiçbir zaman esirgemeyen; çalışmamın her aşamasında bana rehberlik eden, yolumu aydınlatan çok kıymetli, biricik tez danışmanım Prof. Dr. Serap Kurbanoglu'na sonsuz teşekkürlerimi sunarım.

Tezimin ilk aşamasından son aşamasına kadar tez izleme komitemde yer alarak yapıcı eleştirileriyle yol gösteren saygı değer hocam Prof. Dr. Umut Al'a; beni her zaman cesaretlendiren, yaptığı katkılarla tezimi zenginleştiren çok kıymetli hocam Doç. Dr. Gülten Alır Derbent'e çok teşekkür ederim. Tez savunma sınavımda yer almayı kabul ederek tezimi titizlikle okuyan, görüşlerini paylaşan ve önemli katkılar sağlayan değerli hocalarım Prof. Dr. Tülay Oğuz'a ve Doç. Dr. Yurdağül Ünal'a teşekkürü bir borç bilirim.

Doktora eğitimim boyunca derslerine katılma fırsatı bulduğum kıymetli bölüm hocalarım Sayın Prof. Dr. Yaşar Tonta'ya, Prof. Dr. Nazan Özenç Uçak'a, Prof. Dr. İnci Önal'a, Prof. Dr. Bülent Yılmaz'a, Prof. Dr. Özgür Külcü'ye ve Doç. Dr. İrem Soydal'a çok teşekkür ederim.

Ayrıca çok zorlu bir süreçte, başta tıp doktorları olmak üzere değerli zamanlarını ayırarak anketimi dolduran ve araştırmanın verilerinin oluşmasını sağlayan tüm araştırmacılara minnettarım.

Sorun çözücü ve iş bitirici arkadaşlarım Mehmet Taşkol'a, Kerim Sert'e ve Merve Avcı'ya sağladıkları katkılardan dolayı ayrıca teşekkür ederim.

Ve son olarak bu uzun süreçte desteklerini benden esirgemeyen sevgili aileme, üzerimdeki hakkımı ve emeğini hiçbir zaman ödeyemeyeceğim canım anneme ve babama ve kızlarım Ada ile Çağla'ya da sonsuz teşekkürler.

ÖZET

TAVLUOĞLU, Canan. *Araştırma Verilerinin Yönetimi: TÜBİTAK Projelerinde Yer Alan Araştırmacılar Üzerine Bir Değerlendirme*, Doktora Tezi, Ankara, 2022.

Doğrulanabilirlik, hesap verebilirlik, görünürlük, geçerlilik, verimlilik, tekrarlanabilirlik ve kaynak tasarrufu sağlama özellikleri nedeniyle araştırma verilerinin paylaşımı giderek önem kazanmaktadır. Sağladığı ekonomik yarar ve kamu yararı nedeniyle, birçok kurum ve kuruluş politika belgelerini araştırma verilerini ve veri paylaşımını kapsayacak şekilde düzenlemektedir.

Veri paylaşımı ve araştırma verilerinin yönetimi birbirlerini tamamlayan iki önemli konudur. Araştırmacıların bu alandaki mevcut durumunu saptamak konuyla ilgili planlama ve hizmet geliştirilmede katkı sağlamaktadır.

Bu çalışma, TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu) desteğiyle yürütülen araştırma projelerinde yer alan araştırmacıların, araştırma verilerinin yönetimi ve paylaşımı konusundaki uygulamalarını, davranışlarını ve tutumlarını belirlemek amacıyla yapılmıştır.

Araştırmaya 73 farklı üniversiteden 408 araştırmacı katılmıştır. Konuya ilişkin mevcut durumu saptamak için araştırmacılara 34 sorudan oluşan bir anket uygulanmıştır. Bulgulara göre, araştırmacılar büyük verilerle çalışmamakta, genelde veri yönetim planı hazırlamamakta, verilerini daha çok kişisel bilgisayarlarda saklamakta ve uzun süreli korumaya hazırlamamaktadır. Araştırmacıların verilerini paylaştığı ancak herkese açık bir şekilde paylaşım yapma düzeyinin düşük olduğu görülmektedir. Ayrıca araştırmacılar, araştırma verilerinin yönetimi konusunda eğitime gereksinim duymaktadır ve en çok veri yönetimi planı hazırlama konusunda eğitim gereksinimi vardır.

Araştırma verilerinin yönetimi ve veri paylaşımının yaygınlaşması için ulusal ve kurumsal düzeyde hazırlanacak politikalar önemlidir. Bu konuda, Türkiye’de ilerleme kaydedilmesi için ilgili kişi, kurum ve organizasyonların üstlenmesi gereken roller çalışma kapsamında ortaya konulmuştur.

Anahtar Sözcükler

Araştırma verilerinin yönetimi, veri paylaşımı, veri uygulamaları, araştırma verisi, açık veri

ABSTRACT

TAVLUOĞLU, Canan. *Research Data Management: An Evaluation of Researchers Involved in TÜBİTAK Projects*, PhD Thesis, Ankara, 2022.

Research data sharing is becoming increasingly important because of its verifiability, accountability, visibility, validity, efficiency, reproducibility, and resource-saving features. Many institutions and organizations revise their policy documents to cover research data management and data sharing due to the economic and public benefit it provides.

Research data sharing and research data management are two important complementary issues. Determining the current situation of researchers in this regard contributes to the planning and development of services on the subject.

This study aimed to determine the research data management and sharing practices, behaviours and attitudes of researchers who are involved in research projects carried out with the support of TÜBİTAK (The Scientific and Technological Research Council of Turkey).

408 researchers from 73 different universities participated in the study. Within this study, a questionnaire consisting of 34 questions was conducted in order to determine the current situation of the researchers. According to the findings, researchers do not work with big data, generally do not prepare a data management plan, store their data mostly on personal computers and do not prepare their data for long-term protection. It is seen that researchers share their data, however the level of sharing their data publicly is low. Additionally, researchers need training in the management of research data, and training in preparing a data management plan is most needed.

Policies to be prepared at the national and institutional level are important for the research data management and sharing. In this regard, the roles that the relevant persons, institutions, and organizations should undertake in order to make progress in Turkey were presented within the scope of this study.

Keywords

Research data management, data sharing, data practices, research data, open data

İÇİNDEKİLER

| | |
|--|-----|
| KABUL VE ONAY | i |
| YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI | ii |
| ETİK BEYAN | iii |
| TEŞEKKÜR | iv |
| ÖZET | v |
| ABSTRACT | vi |
| İÇİNDEKİLER | vii |
| KISALTMALAR DİZİNİ | x |
| TABLOLAR DİZİNİ | xi |
| ŞEKİLLER DİZİNİ | xii |
| 1. BÖLÜM: GİRİŞ | 1 |
| 1.1 KONUNUN ÖNEMİ | 1 |
| 1.2 ARAŞTIRMANIN AMACI | 4 |
| 1.3 ARAŞTIRMA PROBLEMİ | 4 |
| 1.4 HİPOTEZ | 4 |
| 1.5 ARAŞTIRMANIN KAPSAMI | 5 |
| 1.6 ARAŞTIRMANIN YÖNTEMİ | 7 |
| 1.7 ARAŞTIRMANIN DÜZENİ | 10 |
| 1.8 KAYNAKLAR | 10 |
| 2. BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ VE PAYLAŞIMI | 11 |
| 2.1 TARİHSEL GELİŞİM | 11 |
| 2.2 ARAŞTIRMA VERİSİ | 15 |
| 2.3 ARAŞTIRMA VERİLERİNİN YÖNETİMİ | 17 |
| 2.3.1 Veri Yönetim Politikaları..... | 18 |
| 2.3.2 Veri Yaşam Döngüsü..... | 20 |
| 2.4 VERİ PAYLAŞIMI | 21 |

| | | |
|------------|--|-----------|
| 2.4.1 | FAIR İlkeleri | 23 |
| 2.5 | İLGİLİ ÇALIŞMALAR | 24 |
| 2.5.1 | Türkiye'deki Durum | 27 |
| 3. | BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ SÜRECİNİN AŞAMALARI | 45 |
| 3.1 | PLANLAMA | 45 |
| 3.2 | VERİ TOPLAMA VE DÜZENLEME | 47 |
| 3.2.1 | Düzenleme | 47 |
| 3.2.2 | Veri Dokümantasyonu | 50 |
| 3.2.3 | Üst Veri ve Üst Veri Standartları | 51 |
| 3.3 | DEPOLAMA VE KORUMA | 53 |
| 3.3.1 | Depolama ve Yedekleme | 54 |
| 3.3.2 | Veri Güvenliği | 55 |
| 3.3.3 | Uzun Süreli Koruma | 55 |
| 3.4 | YAYIMLAMA VE PAYLAŞMA | 57 |
| 3.5 | KEŞFETME VE YENİDEN KULLANMA | 61 |
| 4. | BÖLÜM: BULGULAR VE DEĞERLENDİRME | 63 |
| 4.1 | DEMOGRAFİK BİLGİLER | 64 |
| 4.2 | ARAŞTIRMA VERİLERİ YÖNETİMİ KONUSUNDA FARKINDALIK DURUMU VE UYGULAMA EĞİLİMLERİ | 66 |
| 4.2.1 | Dosya Türleri | 66 |
| 4.2.2 | Veri Elde Etme Biçimleri | 69 |
| 4.2.3 | Veri Miktarı | 74 |
| 4.2.4 | Depolama ve Saklama | 76 |
| 4.2.5 | Yedekleme | 79 |
| 4.2.6 | Uzun Süreli Koruma | 80 |
| 4.2.7 | Veriyi Tanımlama: Üst Veri Kullanımı ve Dokümantasyon İşlemleri | 83 |
| 4.2.8 | Veri Yönetim Planı | 87 |
| 4.3 | VERİ PAYLAŞIMI | 90 |

| | | |
|---------------------------------|--|------------|
| 4.3.1 | Veri Paylaşım Durumları | 90 |
| 4.3.2 | Veri Paylaşım Yolları..... | 94 |
| 4.3.3 | Veri Paylaşmama Nedenleri..... | 96 |
| 4.3.4 | Araştırma Verilerinin Yönetimindeki Temel Problemler | 98 |
| 4.4 | AVY KONUSUNDA EĞİTİM GEREKSİNİMİ | 99 |
| 4.4.1 | Eğitim Gereksinimi Duyulan Konular | 101 |
| 4.5 | KATILIMCI GÖRÜŞLERİ | 104 |
| 5. | BÖLÜM: SONUÇ VE ÖNERİLER | 107 |
| 5.1 | SONUÇLAR | 108 |
| 5.2 | ÖNERİLER | 115 |
| KAYNAKÇA | | 124 |
| EK 1. Anket Soruları | | 134 |
| EK 2. Etik Kurul İzni | | 143 |
| EK 3. Orijinallik Raporu | | 144 |

KISALTMALAR DİZİNİ

| | | |
|-------------|---|---|
| AB | : | Avrupa Birliği |
| ALLEA | : | All European Academics |
| AÜ | : | Ankara Üniversitesi |
| AVY | : | Araştırma Verilerinin Yönetimi |
| BBY | : | Bilgi ve Belge Yönetimi Bölümü |
| DDI | : | Data Documentation Initiative |
| Dublin Core | : | DC |
| G8 | : | The Group of Eight |
| HÜ | : | Hacettepe Üniversitesi |
| ISO | : | International Organization for Standardization |
| IWGDD | : | Interagency Working Group On Digital Data |
| İTÜ | : | İstanbul Teknik Üniversitesi |
| İÜ | : | İstanbul Üniversitesi |
| JISC | : | Joint Information Systems Committee |
| MAG | : | Mühendislik Araştırma Destek Grubu |
| MFAG | : | Matematik, Fizik Araştırma Destek Grubu |
| NASA | : | National Aeronautics and Space Administration |
| NIH | : | National Institutes of Health |
| NISO | : | National Information Standards Organization |
| NSB | : | National Science Board |
| NSF | : | National Science Foundation |
| OAI-PMH | : | Open Archives Initiative Protocol for Metadata Harvesting |
| ODTÜ | : | Orta Doğu Teknik Üniversitesi |
| PLoS | : | The Public Library of Science |
| RCUK | : | Research Councils United Kingdom |
| RDA | : | Research Data Alliance |
| SOBAG | : | Sosyal ve Beşerî Bilimler Araştırma Destek Grubu |
| SPSS | : | Statistical Package for the Social Sciences |
| t.y. | : | Tarih yok |
| TÜBİTAK | : | Türkiye Bilimsel ve Teknolojik Araştırma Kurumu |
| UAVS | : | Ulusal Araştırma Verileri Sempozyumu |
| UKRI | : | United Kingdom Research and Innovation |
| ULAKBİM | : | Ulusal Akademik Ağ ve Bilgi Merkezi |
| URAP | : | University Ranking by Academic Performance |
| VYP | : | Veri Yönetim Planı |
| YÖK | : | Yükseköğretim Kurulu |

TABLolar DİZİNİ

| | |
|---|----|
| Tablo 1. TÜBİTAK destekli proje sayıları (2014-2018)..... | 6 |
| Tablo 2. Anketin yapısı | 8 |
| Tablo 3. Formatına göre veriler..... | 17 |
| Tablo 4. Üst veri türleri | 52 |
| Tablo 5. Saklama için tercih edilen formatlar. | 56 |
| Tablo 6. Lisanslar ve koşulları | 60 |
| Tablo 7. Katılımcıların yaş aralıkları..... | 65 |
| Tablo 8. Katılımcıların deneyim bilgileri..... | 65 |
| Tablo 9. Disipline göre dış kaynaklı veri kullanma durumu..... | 72 |
| Tablo 10. Veri saklama ortamları..... | 77 |
| Tablo 11. Demografik özelliklere göre veri saklama ortamları | 78 |
| Tablo 12. Kullanılan üst veri standartları..... | 84 |
| Tablo 13. TÜBİTAK Açık Bilim Politikası farkındalık durumu | 90 |

ŞEKİLLER DİZİNİ

| | |
|--|-----|
| Şekil 1. Araştırmacıların mensup oldukları kurumlar | 7 |
| Şekil 2. Veri yaşam döngüsü | 20 |
| Şekil 3. Araştırma verilerinin yönetimi sürecinde Türkiye'deki mevcut durum..... | 35 |
| Şekil 4. Araştırma verilerinin yönetimde paydaşlara düşen görev ve sorumluluklar | 37 |
| Şekil 5. Katılımcıların kurumlara göre dağılımı | 64 |
| Şekil 6. Katılımcıların unvanlara göre dağılımı..... | 65 |
| Şekil 7. Disipline göre kullanılan dosya türleri (%). | 67 |
| Şekil 8. Araştırmacıların veri elde etme biçimleri (%)..... | 69 |
| Şekil 9. Disipline göre katılımcıların veri elde etme biçimleri (%)..... | 71 |
| Şekil 10. Disipline göre dış kaynaklı veri kullanım biçimleri (%) | 74 |
| Şekil 11. Disipline göre kullanılan veri miktarları (%) | 75 |
| Şekil 12. Yaşa göre verilerin saklanma ortamları | 79 |
| Şekil 13. Uzun süreli korumada sorumluluk (%)..... | 83 |
| Şekil 14. Yapılan dokümantasyon işlemleri (%) | 86 |
| Şekil 15. VYP hazırlayan araştırmacıların kullandıkları alanlar (%) | 88 |
| Şekil 16. Katılımcıların disipline göre veri paylaşım durumları (%)..... | 91 |
| Şekil 17. Katılımcıların veri paylaşma koşulları (%)..... | 92 |
| Şekil 18. Veri paylaşmama nedenleri (%) | 96 |
| Şekil 19. AVY alanında temel problemler..... | 99 |
| Şekil 20. AVY konusunda eğitim gereksinimi (%) | 100 |
| Şekil 21. Disipline göre eğitim gereksinimleri (%) | 102 |
| Şekil 22. Yaşa göre eğitim gereksinimleri (%)..... | 103 |
| Şekil 23. Deneyime göre TÜBİTAK AVY Eğitim Portalı farkındalık durumu (%).... | 104 |

1. BÖLÜM: GİRİŞ

1.1 KONUNUN ÖNEMİ

Bilgi ve iletişim teknolojilerindeki gelişmelere paralel olarak bilimsel çalışmalarda üretilen veri miktarında (biyoinformatik alanında genetik dizilimler ve protein yapısı verileri, sinirbilimde çeşitli türdeki beyin görüntüleri, astronomide gökyüzü araştırmaları ve sanal gözlem verileri gibi) büyük bir artış yaşanmıştır (Arzberger ve diğerleri, 2004, s. 138). Bununla birlikte, verinin depolanması, paylaşımı ve yeniden kullanımı daha önemli hale gelmiştir. Coğrafi sınırlamaların ortadan kalktığı, kıtalar arası ortak çalışmaların yürütüldüğü, iş birliğinin yaygınlaştığı bu dönemde araştırmanın doğası değişmiş, araştırmanın dayanağı olan verilerin değeri artmıştır.

Uluslararası iş birliği ve veri paylaşımının en iyi örneği olan İnsan Genomu Projesi, 1990 yılında başlayıp 2003 yılında tamamlanmış ve proje insanlığa büyük katkı sağlamıştır. Projede elde edilen genom dizilimleri DNA veri bankası olan GenBank'ta korunmaktadır (Ray, 2014). İnsan genomu projesinin ekonomik getirisi yapılan yatırımın yaklaşık 7,7 kat fazlasıdır (The data harvest, 2014, s.5).

Bilimin birikimli olarak ilerlediği göz önüne alındığında, veri paylaşımının önemi yadsınamaz. Patent, telif hakkı gibi belgeler, entelektüel mülkiyet haklarının belirlenmesi ve korunmasında gerekli olmakla birlikte fikir sahibinin buluşunu/fikrini açıkça paylaştığı durumlar da vardır. Örneğin, ilk bilgisayar veya World Wide Web için patent alınmamıştır (Tonta, 2016). Açıklığın ve paylaşımın teknolojik ilerlemeyi hızlandırdığı görülmektedir.

Araştırma verilerinin yönetimi, planlamayla başlayan ve verilerin toplanması, düzenlenmesi, saklanması, korunması, yayımlanması ve paylaşılması aşamalarını kapsayan bir süreçtir. Verilerin geçen zaman içinde bağlamının ve içeriğinin anlaşılabilmesi için bu süreç çok önemlidir. Araştırma verilerinin yönetimi ve veri paylaşımının sağladığı yararlar şu şekildedir (Patel, 2016; Mantra Research Data Management Training, 2020e):

- Araştırmanın verimliliğini artırmaktadır.
- Uzun vadede zaman ve kaynak tasarrufu sağlamaktadır.

- Araştırmanın görünürlüğünü ve etkisini artırmakta, dolayısıyla atıf sayılarında artış sağlamaktadır.
- Araştırma sonuçlarının doğrulanmasını sağlayarak güvenilirliği artırmaktadır.
- Verinin güvenliğini artırarak, veri kaybı riskini en aza indirmektedir.
- Araştırmanın bütünlüğünü ve tekrarlanabilirliği sağlamaktadır.
- Diğer araştırmacılar ve farklı bağlamdaki araştırmalar için verinin tekrar kullanılmasını sağlamakta ve veri toplama için harcanacak çabanın tekrarlanmasını önlemektedir.
- Fon sağlayıcı kuruluşun hibe gereksinimlerini karşılamaktadır.

Veri paylaşımının sağladığı avantajların yanı sıra veri paylaşmamanın da dezavantajları bulunmaktadır. Araştırmacılar ve ekonomistler veri paylaşmamanın bir maliyeti olduğuna inanmaktadır (The data harvest, 2014, s. 27). Özellikle sağlık alanında paylaşım yapmamanın bedeli çok ağır olabilmektedir. Dünya çapındaki kolera salgınının önlenmesinde ve kontrol edilmesinde uluslararası verilere erişim yardımcı olmuştur (Arzberger ve diğerleri, 2004, s. 139). Salgının çevresel faktörlerle ilişkisi disiplinler arası (epidemioloji, sosyal bilimler, deniz biyolojisi, mikrobiyoloji alanına ait veriler, genomik veriler, NASA uzaktan algılama verileri gibi) verilere dayanarak araştırılmıştır (Arzberger ve diğerleri, 2004, s. 139).

Kamu kaynakları kullanılarak gerçekleştirilen araştırmalara ait verilerin yayımlanması için hükümetler, fon sağlayıcı kuruluşlar, dergi yayıncıları, akademik kurumlar ve araştırma kurumları önemli itici güç olmuştur (Curty, 2015, s. 3). Vatandaşın verdiği vergiler kullanılarak kamu desteğiyle yürütülen araştırmalarda kullanılan verilerin mümkün olduğunca paylaşılması görüşü benimsenmiştir.

Birleşik Krallık, Amerika Birleşik Devletleri, Avustralya gibi birçok ülkede, araştırma verilerinin uzun vadede değeri ve etkisinin farkında olan fon sağlayıcılar, kamuya getiri sağlamak için, araştırma verisi yönetim planı oluşturulmasını ve araştırma verilerinin veri depolarında paylaşılmasını zorunlu kılmıştır (Kahn, Higgs, Davidson ve Jones, 2014, s. 296; Si ve diğerleri, 2015). Böylelikle araştırma verilerinin uzun ömürlü ve sürekli kullanılabilir olması sağlanmaktadır (Pryor, 2012, s. 2).

Dünyadaki diğer uygulamalar incelendiğinde; araştırma verileri konusunda bir yol haritası çizildiği, bu konuya ilişkin politikalar oluşturulduğu, akademik kütüphanelerin veri yönetimi konusunda araştırmacılara destek verdiği, kurumların/fon sağlayıcıların alt

yapı, bakım, uzun süreli koruma gibi konularda hizmet sunduğu görülmektedir. Ulusal politikaların eksikliği, araştırma verisine erişimdeki en önemli engellerden bir tanesidir (European Commission, 2012). Veri yönetiminin başarılı bir şekilde gerçekleşmesi için temel gereklilik ulusal ve kurumsal düzeyde politikaların geliştirilmesi ve gerekli düzenlemelerin yapılmasıdır (Whyte ve Tedds, 2011; Ünal ve Kurbanoğlu, 2018, s. 307).

Ülkemizde ise akademik çevrenin çabalarıyla, açık bilim, açık veri, veri okuryazarlığı gibi konulara yönelik çeşitli çalışmalar, projeler yapıldığı (bkz. Tonta, 2013b; Aydınoglu, Doğan ve Taşkın, 2017; Ünal ve Kurbanoğlu, 2018; Zencir, 2019; Doğan, Taşkın ve Aydınoglu, 2021) ancak konunun henüz gelişme aşamasında olduğu görülmektedir. TÜBİTAK Açık Bilim Politikası yayımlanmadan önce 2018 yılında yapılan bir çalışmada, Türkiye’de araştırma verilerinin büyük bir çoğunluğunun erişime açık olmamasının dolayısıyla yeniden kullanılabilir olmamasının konuya ilişkin politika, strateji ve direktif eksikliğinden kaynaklandığı belirtilmiştir (Ünal ve Kurbanoğlu, 2018, s. 291). Geçen sürede TÜBİTAK ve bazı üniversiteler Açık Bilim Politikalarını yayımlamışlardır ancak araştırma verilerinin yönetimi konusu eksikliğini hissettirmektedir. Birçok kurumun herhangi bir veri yönetim ve paylaşım politikasının olmadığı ve kütüphanelerin henüz bu alanda araştırmacılara destek verecek düzeyde olmadığı görülmektedir.

Türkiye’de yapılan çalışmalar, araştırmacıların araştırma verileri yönetimine ilişkin bilgi/beceri eksikliği olduğunu ve konuya ilişkin (özellikle veri yönetim planı konusunda) eğitim gereksinimi duyduklarını ortaya koymaktadır (Aydınoglu ve diğerleri, 2017, s. 282; Ünal ve Kurbanoğlu, 2018, s. 305).

Veri paylaşımı ve araştırma verilerinin yönetimi (AVY) birbirlerini tamamlayan iki önemli konudur. Araştırmacıların bu alandaki mevcut durumunu saptamanın konuyla ilgili planlama ve hizmet geliştirmede katkı sağlayacağı düşünülmektedir. Bu çalışmada, TÜBİTAK destekli araştırmalarda yer alan araştırmacıların, araştırma verilerinin yönetimine (planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) yönelik uygulamaları ve tutumları saptanmaya çalışılmıştır. Eldeki bulgulara dayalı yapılan önerilerin Türkiye’de araştırma verileri yönetimi konusundaki eksikliklerin giderilmesinde yardımcı olacağı düşünülmektedir. Bu çalışma, araştırma evreninin Türkiye’nin önde gelen üniversitelerinde TÜBİTAK desteğiyle yürütülmüş projelerde yer alan araştırmacılardan oluşması ve disiplin haricindeki diğer demografik

özelliklere (yaş, deneyim, kurum, unvan) göre de AVY uygulamalarında farklılık olup olmadığının incelenmesi nedeniyle Türkiye üzerinde yapılan/yapılmakta olan diğer çalışmalardan farklılık göstermektedir.

1.2 ARAŞTIRMANIN AMACI

Bu çalışmanın temel amacı, Türkiye'nin URAP sıralamasına göre ilk beş üniversitesinde 2014-2018 yılları arasında TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu) desteğiyle tamamlanmış araştırma projelerinde yer alan araştırmacıların araştırma verilerinin yönetimi konusundaki uygulamalarını ve tutumlarını belirlemektir.

1.3 ARAŞTIRMA PROBLEMİ

Çalışmanın amacı doğrultusunda, araştırma problemi "TÜBİTAK tarafından desteklenen projelerde görev alan araştırmacılar araştırma verilerinin yönetimi sürecini içerdiği tüm aşamalarla (planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) gerçekleştirmekte midir?" şeklinde oluşturulmuştur.

Temel araştırma problemi doğrultusunda, aşağıda listelenen alt sorulara yanıt aranmıştır:

- Araştırmacılar AVY sürecinin planlama aşamasını gerçekleştirmekte midir?
- Araştırmacılar AVY sürecinin düzenleme aşamasını gerçekleştirmekte midir?
- Araştırmacılar AVY sürecinin saklama ve koruma aşamasını gerçekleştirmekte midir?
- Araştırmacılar AVY sürecinin yayımlama ve paylaşma aşamasını gerçekleştirmekte midir?
- Araştırmacıların AVY sürecindeki mevcut uygulamaları demografik özelliklere göre farklılık göstermekte midir?

1.4 HİPOTEZ

Araştırmanın hipotezi "TÜBİTAK tarafından desteklenen projelerde görev alan araştırmacılar, araştırma verilerinin yönetimi sürecini içerdiği tüm aşamalarla (planlama,

veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) bir bütün olarak gerçekleştirilmemektedir” şeklinde oluşturulmuştur.

Araştırma hipotezine bağlı alt hipotezler aşağıda sunulmuştur:

- Araştırmacılar AVY sürecinin planlama aşamasını gerçekleştirilmemektedir.
- Araştırmacılar AVY sürecinin düzenleme aşamasını gerçekleştirilmemektedir.
- Araştırmacılar AVY sürecinin saklama ve koruma aşamasını gerçekleştirilmemektedir.
- Araştırmacılar AVY sürecinin yayımlama ve paylaşma aşamasını gerçekleştirilmemektedir.
- Araştırmacıların AVY sürecindeki mevcut uygulamaları demografik özelliklere göre farklılık göstermektedir.

1.5 ARAŞTIRMANIN KAPSAMI

Araştırma, URAP'a (University Ranking by Academic Performance–Akademik Performansa Göre Üniversite Sıralaması) göre ilk beşte yer alan üniversitelerde 2014-2018 yılları arasında TÜBİTAK desteğiyle gerçekleştirilen projelerde yer alan araştırmacıları kapsamaktadır.

Orta Doğu Teknik Üniversitesi Enformatik Enstitüsü tarafından kurulan URAP Araştırma Laboratuvarı, üniversiteler için sıralama sistemi geliştirmekte ve üniversitelerin akademik performanslarını değerlendirebilmelerine katkı sağlamaktadır. 2018-2019 URAP sıralamasına göre ilk beşte yer alan üniversiteler Hacettepe, Orta Doğu Teknik, İstanbul, İstanbul Teknik ve Ankara üniversiteleridir (URAP, 2018).

Araştırma kapsamında öncelikle URAP sıralamasına göre Türkiye'nin ilk beş üniversitesinde 2014-2018 yılları arasında TÜBİTAK desteğiyle yürütülen projeler belirlenmiştir. Kasım 2018'de TÜBİTAK Destekli Projeler veri tabanında¹ yazarın çalıştığı kuruma göre yapılan sorguda 564 projenin desteklendiği ancak indeksleme hatası veya mükerrer projelerin olması nedeniyle tekil proje sayısının 558 olduğu tespit

¹ <http://uvt.ulakbim.gov.tr/uvt/index.php?cwid=3&vtadi=TPRJ>

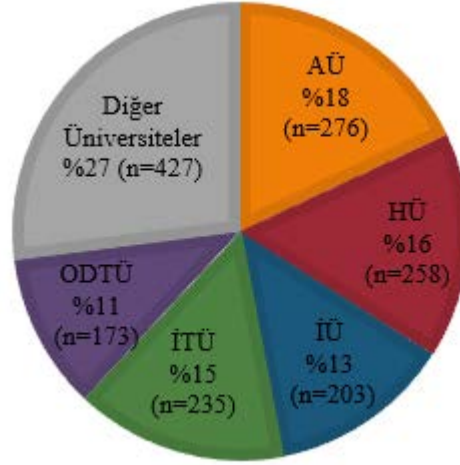
edilmiştir. 15 Kasım 2018 tarihi itibarıyla, TÜBİTAK Destekli Projeler veri tabanına web üzerinden erişim kaldırılmış ve TÜBİTAK destekli projeler TR-Dizin² üzerinden sorgulanmaya başlanmıştır. TR-Dizin’de benzer bir sorgulama yapıldığında, eski veri tabanı ile yeni veri tabanındaki projelerin aynı olmadığı belirlenmiştir. Araştırma kapsamındaki beş üniversitenin destek aldığı proje sayısı TR-Dizin veri tabanında Haziran 2019’da yapılan sorguda 553 olup yanlış indekslenenler ve mükerrer projeler düşüldüğünde sayı 439’a düşmüştür. Her iki veri tabanındaki projeler karşılaştırılmış ve eski veri tabanında olmayıp TR-Dizin’den elde edilen 131 proje de araştırma kapsamına dâhil edilmiştir. 2014-2018 yılları arasında TÜBİTAK desteği alan 689 projenin destek alan kurumlara göre dağılımı Tablo 1’de sunulmaktadır.

Tablo 1. TÜBİTAK destekli proje sayıları (2014-2018)

| Kurum | TÜBİTAK Proje veri tabanı | TR-Dizin | Birleşim |
|--------------------------------------|--|-----------------|-----------------|
| Ankara Üniversitesi (AÜ) | 134 | 127 | 153 |
| Hacettepe Üniversitesi (HÜ) | 120 | 72 | 138 |
| İstanbul Teknik Üniversitesi (İTÜ) | 110 | 98 | 149 |
| İstanbul Üniversitesi (İÜ) | 73 | 70 | 98 |
| Orta Doğu Teknik Üniversitesi (ODTÜ) | 121 | 72 | 151 |
| Toplam | 558 | 439 | 689 |

Söz konusu 689 projede görev alan tekil araştırmacı sayısı 2648 olarak belirlenmiştir. YÖK Akademik Arama veri tabanı ve üniversitelerin akademik bilgi sistemlerinden yararlanılarak 2648 araştırmacının iletişim bilgileri saptanmaya çalışılmıştır. Araştırmacıların iletişim bilgileri saptanırken birden fazla ad veya soyada sahip olunması, soyadı değişikliği, isim benzerlikleri gibi nedenlerle doğru araştırmacıyı belirlemenin sorun yarattığı durumlar olmuştur. Böyle durumlarda, proje adı ve projede yer alan diğer araştırmacı bilgilerinden faydalanılarak doğru kişiler saptanmıştır. Ayrıca, YÖK Akademik Arama veri tabanında bilgi mevcutsa projeler, birlikte çalıştığı kişiler ve yayınlar sayfalarından da yararlanılmıştır. Söz konusu çalışmaların sonucunda yurtiçi üniversitelerde halen görev yapan 1572 araştırmacı olduğu belirlenmiştir. Araştırma evrenini oluşturan bu araştırmacıların görev aldıkları kurumlar Şekil 1’de sunulmuştur.

² <https://trdizin.gov.tr/>



Şekil 1. Araştırmacıların mensup oldukları kurumlar

1.6 ARAŞTIRMANIN YÖNTEMİ

Araştırmada betimleme yöntemi kullanılmış, veriler anket tekniği ile toplanmıştır. Araştırma evrenini oluşturan araştırmacıların konuya ilişkin uygulamalarını ve tutumlarını belirlemek amacıyla literatürdeki çalışmalardan yararlanılarak bir anket hazırlanmıştır.

01.09.2019-31.12.2019 tarihleri arasında çeşitli disiplinlerden 14 araştırmacıya yüz yüze veya e-posta yoluyla pilot uygulama yapılmıştır. Pilot uygulama sonucunda soruların daha anlaşılır hale getirilmesi için anket üzerinde bazı eklemeler ve düzenlemeler yapılmıştır. Yapılan değişiklikler aşağıda özetlenmiştir:

- Araştırmacıların dijital (sayısal) veri kullanıp kullanmadığı anketin ilk sorusu olarak eklenmiştir, “hayır” seçeneğinin işaretlenmesi durumunda anket sonlandırılmak üzere yapılandırılmıştır.
- Anketin ilk kısmına araştırma verisinin tanımı eklenmiş ve araştırma verisinin ne olduğu örneklerle açıklanmıştır.
- Araştırma verileri yönetimi ve veri yönetim planı tanımları ilgili ilk sorunun başına eklenmiştir.
- Bazı sorularda seçeneklerin sıralamasında (olumsuz ifadelerin ilk seçeneğe taşınması gibi) değişiklik yapılmıştır.
- Üst veri bilgilerinin kaydedilmesine ilişkin bir soruda ayrıntılı bilgi edinmek açısından seçenekler artırılmıştır.
- Üst veri standartlarına ilişkin iki soru birleştirilmiştir.

- Bir soruda veri tabanları ifadesinin kütüphane veri tabanları ile karıştırılması sonucu, veri tabanları yerine veri arşivleri (Zenodo, DRYAD, Figshare gibi) ifadesi kullanılmıştır.
- Demografik bilgileri içeren sorular son bölüme taşınmıştır.

Ankette, bir ön sözle çalışma hakkında bilgilendirme yapılmış, araştırma verisinin tanımı ve örnekler verilmiştir. Yanlış anlamalara meydan vermemek için kullanılan temel kavramlara (uzun süreli koruma, üst veri, veri yönetim planı gibi) ilişkin tanımlar yapılmıştır.

Anket üç ana bölümden oluşmaktadır. İlk bölümde araştırmacıların araştırma verileri yönetimi konusunda (veri toplama, depolama ve koruma, planlama) uygulama eğilimlerini ve eğitim gereksinimlerini belirlemeyi amaçlayan 23 soru, ikinci bölümde veri paylaşımına yönelik yaklaşım ve tutumlarını belirlemek amacıyla altı soru, üçüncü ve son bölümde ise demografik bilgileri kapsayan beş soru yer almaktadır (Ek. 1). Anketin yapısına ilişkin ayrıntılı bilgi Tablo 2’de sunulmuştur.

Tablo 2. Anketin yapısı

| No | İlgili olduğu alan | İçerik | Soru no. |
|----|---------------------------|---|----------------|
| 1 | Veri toplama ve düzenleme | Araştırma verisi türleri, elde etme biçimi, veri miktarı | S1-S5, S11-S13 |
| 2 | Depolama ve koruma | Depolama, yedekleme, saklama, uzun süreli koruma | S6-S10 |
| 3 | Planlama | Veri yönetim planı | S14-S18 |
| 4 | Eğitim gereksinimi | Eğitim gereksinimi duyulan konular, genel öneriler, diğer | S20-S23, S29 |
| 5 | Paylaşma | Paylaşım yaklaşımları ve koşulları, paylaşmama nedenleri | S19, S24-S28 |
| 6 | Demografik bilgiler | Kurum, disiplin, unvan, deneyim, yaş | S30-S34 |

26.02.2020-15.01.2021 tarihleri arasında araştırma evrenini oluşturan tüm araştırmacılara (1572 kişi) anket gönderilmiş ve 442 geçerli yanıt alınmıştır. Anketin yanıtlanma oranı %28,1’dir. Anket uygulanmaya başlandıktan kısa bir süre sonra COVID-19 salgını başlamış ve ülke çapında tam kapanmaya gidilmiştir bu nedenle anketin uygulanma aralığı uzamıştır.

Anketin demografik bilgiler bölümünde, araştırma faaliyetine geç başlama ve/veya farklı kurumlarda farklı görevlerde bulunma gibi nedenlerle yaş, unvan ve deneyim bilgilerinin paralellik göstermeyeceği düşünülmüştür. Bu nedenle, katılımcılara hem yaş, hem deneyim, hem de unvan bilgileri sorulmuştur. Deneyim bilgisinde, herhangi bir kurumda çalışılıp çalışılmadığına bakılmaksızın, araştırma faaliyetinde bulunulan süre (yıl) esas alınmıştır.

Ankette doğrudan sorulan “bölüm”, “yaş” ve “deneyim” soruları sonradan kategorilere ayrılmıştır. Disiplin kategorizasyonu yapılırken TÜBİTAK proje kodlarından (MAG-Mühendislik Araştırma Destek Grubu, SOBAG-Sosyal ve Beşeri Bilimler Araştırma Destek Grubu, MFAG-Matematik, Fizik Araştırma Destek Grubu gibi) yararlanılmış ve tıp (tıp ve sağlık bilimleri, eczacılık, diş hekimliği gibi), fen (matematik, fizik, kimya, biyoloji gibi) mühendislik, sosyal bilimler (sosyoloji, arkeoloji, dil bilim, siyaset bilimi, iktisat gibi) ve ziraat (toprak bilimleri, su bilimleri, ormancılık, veterinerlik gibi) olmak üzere beş kategori belirlenmiştir. Yaşa ve deneyime göre sınıflandırma yapılırken ise, saçılım grafiklerinden yararlanılmış ve her iki alan için de dört grup olması gerektiğine karar verilmiştir. Her bir gruptaki aralığın eşit olması için, değişim aralığı (en büyük değer ile en küçük değer farkı) grup sayısına bölünerek sınıf aralığı hesaplanmış ve buna göre gruplandırma yapılmıştır.

SPSS üzerinde tanımlayıcı istatistik ve Ki-kare analizleri yapılmıştır. Uygun olan unsurlar arasında yaşa, deneyime, disipline, kuruma, unvana göre farklılık olup olmadığı Ki-kare analizi ile test edilmiştir.

SPSS programına veri girişi yapılırken, birden fazla seçeneğin işaretlenebildiği sorularda her bir seçenek evet/hayır şeklinde cevaplanabilen ayrı bir soru gibi değerlendirilmiştir. İşaretlenen seçeneklerde cevap “evet”, işaretlenmeyen seçeneklerde cevap “hayır” kabul edilmiştir. Örneğin, araştırmacılara, kullanılan dosya formatları ile ilgili standart ofis dokümanları, görseller, videolar gibi birden fazla seçeneğin olduğu ve işaretlenebildiği bir soru yöneltilmiştir. Ancak her bir seçenek “Standart ofis dokümanı kullanıyor musunuz?” gibi bir soru olarak değerlendirilmiştir. Araştırmacı bu seçeneği işaretlemişse “evet, kullanıyor”, işaretlememişse “hayır, kullanmıyor” şeklinde değerlendirilmiştir. Böylelikle birden fazla seçeneğin işaretlendiği sorularda da Ki-kare analizleri yapmak mümkün olmuş, disiplin, deneyim, unvan, yaş gibi unsurlara göre farklılıklar olup olmadığı ortaya konmaya çalışılmıştır. Sadece bazı seçeneklerde istatistiksel açıdan

anlamli sonu ıkmasının temel nedeni budur. Bu durum, arařtırmanın sınırlılıkları kapsamında deęerlendirilebilir.

1.7 ARAŐTIRMANIN DÜZENİ

Arařtırma beř ana bölümden oluřmaktadır.

Giriř bölümü olan birinci bölümde; arařtırma konusunun önemi, arařtırmanın amacı, hipotezler, arařtırmanın kapsamı, arařtırma yöntemi ve veri toplama teknikleri, arařtırmanın düzeni ile yararlanılan kaynaklar hakkında bilgilere yer verilmektedir.

İkinci bölüm; arařtırma verilerinin yönetimi ve paylaşımına iliřkin ne, neden ve nasıl sorularının yanıtlarını içeren bilgileri kapsamaktadır.

Üüncü bölümde arařtırma verileri yönetimine yönelik uygulamalar veri yařam döngüsü kapsamında açıklanmıřtır.

Dördüncü bölüm, uygulanan anket sonucunda elde edilen verilere yönelik bulgular, literatürdeki bulgularla karřılařtırma ve deęerlendirmeden oluřmaktadır.

Son bölümde ise; arařtırma bulgularına dayanan sonuçlar, hipotezlerin doęrulanıp doęrulanmadığı bilgisi ve öneriler yer almaktadır.

1.8 KAYNAKLAR

Arařtırma konusuna yönelik olarak literatür taraması yapılmıřtır. Library and Information Science Abstract (LISA), Emerald, ScienceDirect, Wiley Online Journal, EBSCO, SpringerLink, Web of Science, Scopus veri tabanları, Proquest Dissertations and Theses Global, YÖK Tez Merkezi, TR Dizin, Toplu Katalog, Bilgi Dünyası, Türk Kütüphanecilięi taranan kaynaklardandır.

2. BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ VE PAYLAŞIMI

Yaşanan teknolojik gelişmelerle birlikte; verinin formatı çeşitlenmiş, üretilen veri miktarı artmış ve araştırmaların doğası değişmiştir. E-bilim dönemi olarak adlandırılan bu dönemde, küresel çapta iş birlikleri yaygınlaşmış, katlanarak artan veri miktarını tanımlamak için “veri tufanı” betimlemesi kullanılmaya başlanmıştır (Hey ve Trefethen, 2003, s. 809) Araştırma verilerinin elektronik olması; paylaşım kolaylığı, yeni ve farklı analizler yapabilme, tekrarlanabilirlik olanağı gibi birçok fırsat sunmaktadır. Ayrıca az gelişmiş ülkelerde fon bulma sorunu yaşayan araştırmacıların veri paylaşımı sayesinde araştırmalara katılabilmesi sağlanmaktadır (Briney, 2015, s. 13-14).

Veri yönetim ve paylaşım politikalarının gelişiminde ekonomik etmenler ve e-araştırmaların etkisi büyüktür. Araştırma verilerinin paylaşılmaya elverişli olması ve paylaşım sonucu ortaya çıkacak ekonomik fayda politikaların uygulanmasını sağlamıştır. Araştırma verilerinin yönetimi ve paylaşımına yönelik önemli olayları kapsayan “Tarihsel Gelişim” bölümü aşağıda sunulmuştur.

2.1 TARİHSEL GELİŞİM

2000’lerin sonundaki küresel ekonomik durgunluk fon sağlayıcı kuruluşların bütçelerine de yansımış, sonuç olarak araştırma fonlarında kısıtlamalara gidilmiştir. Söz konusu kısıtlamalar veri kaybının önlenmesi, hesap verilebilirliğin sağlanması, veri paylaşımı ve verinin tekrar kullanımı yoluyla araştırma çıktılarından maksimum fayda sağlanması üzerinde odaklanmaya neden olmuştur (Briney, 2015, s. 13-14). 2000 yılından önce konuya dikkat çeken kişiler görülmekle birlikte, bazı disiplinlerde (astronomi, biyoloji, fizik gibi) veri paylaşım uygulamaları da gerçekleştirilmiş ancak açık veriyle ilgili düzenlemeler 2000’lerin ortalarında başlamıştır.

Bu tarihten sonra ivme kazanan araştırma verilerinin yönetimi ve paylaşımına yönelik gelişmeler şu şekildedir:

- Açık Toplum Enstitüsü tarafından 1-2 Aralık 2001 tarihinde düzenlenen toplantı, 2002 yılında Budapeşte Açık Erişim Girişiminin oluşmasına yol açmıştır. Açık erişimin ve yararlarının tartışıldığı toplantı sonrasında Açık Erişim Bildirgesi yayımlanmıştır

(Polat, 2006, s. 58; Hey ve Trefethen, 2008, s. 20). 2003 yılında, Bethesda’da yapılan bir toplantı sonrasında açık erişimin ilkeleri ve uygulamalarına ilişkin kararları içeren Açık Erişim Yayıncılığı Üzerine Bethesda Bildirimi yayımlanmıştır (Polat, 2006, s. 59). Aynı yıl, Berlin’de bir toplantı daha gerçekleştirilmiş olup toplantıda alınan kararlar Fen ve İnsan Bilimlerinde Bilgiye Açık Erişim Üzerine Berlin Bildirgesi adı altında yayımlanmıştır. Bu bildirgenin diğerlerinden farkı açık erişim tanımında, ham veri ve üst veri kavramlarına da yer vermesidir (Hey ve Trefethen, 2008, s. 20).

- 2003 yılında JISC’nin (Joint Information Systems Committee-Ortak Bilgi Sistemleri Komitesi) isteği üzerine Lord ve MacDonald e-bilime ilişkin bir rapor hazırlamışlardır. Bu kapsamlı raporda, araştırma verilerinin yönetimindeki sorunlardan bahsedilmekte ve çözüm önerileri sunulmaktadır. Yayınlarla yönelik bir açık erişim hareketinin olduğu ancak verileri kapsamadığı; uzun süreli depolamaya ilişkin düzenlemenin olmadığı; veri yönetimi ve veri altyapısı için hükümet düzeyinde genel bir strateji eksikliği var eksikliği olduğu konularına değinilmiştir. Bu raporda ayrıca, veri yönetim stratejileri, verilerin tanımlanması ve arşivlenmesi için oluşturulabilecek mekanizma önerileri ortaya konmuştur (Lord ve MacDonald, 2003).
- 2003 yılında Amerikan Ulusal Sağlık Enstitüsü Veri Paylaşım Politikası ve Uygulama Rehberi (NIH Data Sharing Policy and Implementation Guidance) oluşturulmuş; belirli bir miktar hibe alan araştırmacılardan veri paylaşım planı, eğer veri paylaşılmıyacaksa bunun nedenlerinin belirtilmesi istenmiştir (National Institutes of Health (NIH), 2003).
- 2004 Ocak ayında -o zamanki- OECD üyesi 30 ülkeyle, Çin, İsrail, Rusya ve Güney Afrika’nın da dâhil olduğu Paris’teki bir toplantıda, araştırma verilerine erişim üzerine politikaların önemi masaya yatırılmıştır. Bu toplantı sonucunda “OECD Kamu Destekli Araştırma Verilerine Erişim Bildirgesi” yayımlanmış ve kamu fonlarıyla üretilen araştırmaların verilerine erişim için OECD’den bir dizi ilke ve rehber geliştirmesi istenmiştir. Oluşturulan taslak üzerine fikir birliğine varmak için 2006’da Paris’te bir çalıştay düzenlenmiş ve 2007’de kamu fonlarıyla üretilen araştırma verilerine uygun maliyetli erişimi kolaylaştırmak için ilkeler yayımlanmıştır (OECD, 2007, s. 7).
- 2005’te Birleşik Devletler Ulusal Bilim Kurulu (National Science Board-NSB), Ulusal Bilim Vakfı’nın (National Science Foundation-NSF) finanse ettiği araştırmalarda

Veri Yönetim Planı (VYP) gerekliliğini rapor etmiş ancak politika belgesinde bir değişiklik olmamıştır (Smale, Unsworth, Denyer ve Barr, 2018, s. 6).

- Birleşik Krallık'ta altı fon sağlayıcı kuruluşun konuyla ilgili aldığı danışmanlık hizmeti (Digital Archiving Consultancy) sonrası sunulan raporda (2005), veri paylaşımı desteklenmekte, VYP gereksinimlerine dikkat çekilmekte ve öneriler yer almaktadır (Digital Archiving Consultancy, Bioinformatics Research Centre, & National e-Science Centre, 2005). 2006 yılında Medical Research Center (Tıbbi Araştırma Merkezi) ve 2007 yılında ise Wellcome Trust bu önerileri dikkate almıştır (Smale ve diğerleri, 2018, s. 6-7). Wellcome Trust veri yönetim ve paylaşım politikasında, destek alan araştırmacıların mümkün olduğunca az kısıtlamayla en üst düzeyde araştırma verilerine erişimine yönelik beklentisini belirtmiştir (Wellcome Trust, 2008).
- Amerika Birleşik Devletleri hükümetinin dijital veriye erişim stratejik plan uygulamalarını geliştirmek ve desteklemek amacıyla kurulan IWGDD (Interagency Working Group On Digital Data-Dijital Veriler Üzerine Kurumlar Arası Çalışma Grubu) 2009 yılında bir rapor yayımlamıştır (IWGDD, 2009, s. 5). Bu raporun içeriğinde dijital veri erişimini ve kullanımını en üst düzeye çıkarma ve VYP sürecini teşvik etme önerileri bulunmaktadır (IWGDD, 2009)
- Açık Bilgi Vakfı, 2009'da taslağı oluşturulan ve 2010 da kabul edilen Panton ilkelerini benimsemiştir. Bu ilkeler, açık veri kavramında neyin açık olduğu konusundaki karışıklığı azaltmak için özellikle lisanslama konusuna yer vererek verilerin yayımlanmasında dikkat edilecek hususları içermektedir (Molloy, 2011, s. 2).
- Birleşik Devletler Ulusal Bilim Kurulu'nun (National Science Board-NSB) Mayıs 2010'daki toplantısında, Ulusal Bilim Vakfı (National Science Foundation-NSF) yetkilileri araştırma verilerinin paylaşılmasına ilişkin mevcut politikanın uygulanmasında bir değişiklik olduğunu ve araştırma önerilerinin bir veri yönetim planı içermesi zorunluluğunun uygulanacağını duyurmuştur (National Science Foundation, 2010) ve 2011'de araştırma başvurularında veri paylaşım planı ekleme zorunluluğu getirilmiştir (National Science Foundation, 2011).
- Birleşik Krallık Araştırma Konseyi (Research Councils United Kingdom-RCUK) -şimdiki adıyla UK Research and Innovation (UKRI)- 2011 yılında araştırma verilerinin yönetimi ve paylaşımı üzerine beklentileri içeren yedi ortak ilke yayımlamıştır. Bu

ilkeler, veri politikası konusunda araştırma kurumlarına kapsamlı bir çerçeve sunmaktadır (UKRI, 2011).

- 2012 yılında, Türkiye Bilimler Akademisi'nin de üye olduğu ALLEA (All European Academics-Tüm Avrupa Akademileri), "21. Yüzyıl İçin Açık Bilim" (Open Science for the 21st century) bildirgesinde, açık bilim vizyonunun yerine getirilmesi için kamu fonuyla üretilen yayınlara ait veriler için açık paylaşım ilkelerinin uygulanması gerektiğine yer vermiştir (Kroes, 2012).
- 2013 yılında G8 zirvesinde, bilim bakanları bilimsel araştırma verilerine herkes tarafından erişim sağlanması, verilerin yayımlanması ve yeniden kullanımı içeren temel ilkeleri benimsemiştir (Willets ve diğerleri, 2013).
- Avrupa Komisyonu'nun 2011'de araştırmacılara uyguladığı anket sonuçlarına göre veri erişimine ve korumaya yönelik problemler olduğu ortaya çıkmıştır (European Commission, 2012). Bunun üzerine, 2 Temmuz 2013'te Brüksel'de paydaşlarla açık araştırma verileri üzerine bir toplantı düzenlenmiştir (Spichtinger ve Siren, 2017, s. 14). Bu gelişmeler, kamu tarafından finanse edilen araştırma verilerine açık erişim için yasal düzenleme yapılmasını sağlamıştır (Regulation (EU) No 1291/2013/291/2013, madde 28). Avrupa Komisyonu tarafından fonlanan Ufuk 2020 kapsamındaki kamu tarafından finanse edilen araştırmalardan kaynaklanan araştırma verilerine açık erişimin mahremiyet, ulusal güvenlik ve entelektüel mülkiyet hakları dikkate alınarak desteklenmesi kararı alınmıştır.
- Önce pilot uygulamayla (2014-2016) sonrasında ise 2017 yılından başlayarak Avrupa Komisyonu desteğiyle gerçekleştirilen Ufuk 2020 projelerine ait araştırma verileri için açık olma zorunluluğu getirilmiştir (EU Commission, 2016)
- 2015 yılında Birleşik Krallık araştırma kuruluşları araştırma verilerinin erişilebilir olması için on ilke içeren Açık Araştırma Verileri Antlaşması'nı yayımlamıştır (Concordat on Open Research Data, 2015).
- Science, Nature ve PLoS gibi birçok dergi yayıncısı da veri yönetim ve paylaşım politikalarını yayımlamışlardır (Briney, 2015, s. 13).

Sonuç olarak birçok fon sağlayıcı kuruluş, üniversite ve bilimsel dergi, veri paylaşım ve yönetim politikalarını geliştirmişlerdir. Araştırmacılardan bu süreci iyi yönetmeleri ve verilerini mümkün olduğunca paylaşmaları beklenmektedir. Araştırma sırasında iyi

yönetilmiş veriler paylaşım aşamasında daha az hazırlık gerektirmektedir (Briney, 2015, s. 13). Araştırma verilerinin yönetimi sürecinin aşamaları ve bu aşamalara yönelik uygulamalar [Bölüm 3](#)'te verilmiştir.

2.2 ARAŞTIRMA VERİSİ

Araştırma verisinin tanımı araştırmacıya, disipline ve fon sağlayıcı kuruluşa göre değişiklik gösterebilmektedir. Dolayısıyla tanım üzerinde bir fikir birliği yoktur (University of Leicester, 2019). Ancak kapsamlı bir tanım yapmak gerekirse; **araştırma verileri**, araştırmada temel kaynak olarak kullanılan ve araştırma sonuçlarını üretmek için toplanan, gözlemlenen veya oluşturulan; sayısal-metinsel, görüntü, ses gibi dosyaları kapsayan herhangi bir formattaki ham, temizlenmiş veya işlenmiş verilerdir (OpenAIRE Research Data Management Briefing Paper, 2016, s. 3; Mantra Research Data Management Training, 2020d; Organisation for Economic Co-operation and Development, 2007).

Veri kelimesi farklı disiplinlerdeki kişilere farklı şeyler ifade edebilmektedir. Araştırmanın temelini oluşturan “araştırma verisi” kavramı yerine disipline özgü farklı bir ifade de kullanılabilir (Cox ve Verbaan, 2018, s. 20). Örneğin, tarih alanında çalışan bir araştırmacı için araştırmanın dayanağı, birincil kaynaklar, metinler veya arşiv malzemeleri olup veri olarak ifade edilmemektedir.

Araştırma verileri; laboratuvar notları, günlükler, fosil, kayaç, tahta, yaprak, doku örneği gibi fiziksel formlarda olabileceği gibi, sonradan dijital ortama aktarılmış veya tamamen elektronik ortamda üretilmiş/toplanmış da olabilir. Bazı araştırmacılar için elektronik olan her şey veriyken bazı araştırmacılar için hem analog hem de elektronik malzemeler veridir (Mantra Research Data Management Training, 2020d).

Beşerî bilimler için birincil kaynaklar veya metinler; sosyal bilimler için anket sonuçları, görüşme notları ve istatistikler; fen bilimleri için deneylerin, gözlemlerin çıktıları araştırma verisi olabilir (Burnham, 2012, s. 3). Literatürde, tamamen teoriğe dayalı olan felsefe gibi bazı alanlarda ise araştırma verisi bulunmadığı ifade edilse de (Cox ve Verbaan, 2018, s. 20), bu alanlarda çalışma sonucunu destekleyen açıklamalar, notlar, kuramlar veridir.

Kaynağına Göre Veriler

Kaynağına göre veriler dört grupta sınıflandırılmıştır (Briney, 2015, s. 16; University of California Curation Center, 2020):

- *Gözlemsel Veriler*: Olayların ve davranışların belirli bir zaman ve yerde gerçek zamanlı olarak gözlenmesi sonucu ortaya çıkan verilerdir. Bu tür verilerin yeri başka bir veriyle doldurulamadığı/tekrarlanamadığı için korunması çok önemlidir. Örneğin, hava durumu, sensör okumaları, anket sonuçları gibi.
- *Deneysel Veriler*: Laboratuvarlarda veya kontrollü ortamlarda deney sonucu üretilen ve benzer koşullarda aynı sonuca ulaşılabilen verilerdir. Tekrarlanması mümkün olmakla birlikte, maliyetli ve zaman alıcı olabilmektedir. Örneğin, Büyük Hadron çarpıştırıcısından gelen ölçümler, gen dizilimleri, manyetik alan okumaları gibi.
- *Simulasyon Verileri*: Gerçekte var olan sistemin veya sürecin makine desteğiyle modellenmesi sonucu test modellerinden üretilen verilerdir.
- *Derlenmiş Veriler*: Var olan veri setlerinden yani ikincil kaynaklardan toplanan, tekrar üretilen ancak üretilmesi emek, zaman isteyen, maliyetli olan verilerdir. Örneğin, metin ve veri madenciliği, derlenmiş veri tabanları, üç boyutlu modellemeler gibi.

Formatına Göre Veriler

Araştırmalarda çeşitli dosya formatlarında ve birçok farklı türde veri kullanılmaktadır. Veri türüne göre araştırma verilerinin içeriği ve içerebileceği dosya formatları Tablo 3'te sunulmuştur (Mantra Research Data Management Training, 2020a).

Araştırmalarda kullanılan farklı türdeki verilerin yapılandırılması veya saklanmasında çeşitli dosya formatları kullanılmaktadır. Gelecekte verilerin erişilebilir olması için seçilen dosya formatı önemli olmakla birlikte önerilen dosya formatları [Bölüm 3.3.3](#)' de aktarılmıştır.

Tablo 3. Formatına göre veriler

| Veri türleri | İçerik | Dosya formatları |
|-------------------|---|--|
| Metin | Anketler, röportaj transkriptleri, kod çizelgeleri, metodolojisi, iş akışları, standart çalışma prosedürleri, protokoller, saha not defterleri, günlükler | txt (düz metin), pdf, docx (Word) html, xml |
| Dijital | Anket yanıtları, tıbbi kayıtlar, test yanıtları, elektronik çizelgeler (spreadsheets), ölçüm aletleri, mekânsal bilgi | dta (Stata) sav, sdq, spv (SPSS) xlsx, csv (Excel) gpx, kml (GIS) |
| Multimedya | Fotoğraflar, ses kayıtları, videolar | jpeg, png, tiff, mp3, wav, mpeg |
| Geliştirilmiş kod | Kaynak kodu, algoritmalar, betik (script) | Python, Java, mat (MATLAB) |
| Sözdizim (syntax) | Veri hazırlama, bağlantı, istatistiksel analiz gibi veri işleme adımlarını gerçekleştirmede kullanılan yazılma özgü kod dosyaları | dta (Stata) sav, sdq, spv (SPSS), R, mat (MATLAB) |
| Disipline özgü | | FITS - [Astronomi] CIF - [Kimya] GRIB - [Meteoroloji] |

2.3 ARAŞTIRMA VERİLERİNİN YÖNETİMİ

Bu bölümde, araştırma verileri yönetimiyle ilgili olarak **ne** (tanım, amaç, kapsam), **neden** (önemi ve politika belgeleri) ve **nasıl** (veri yaşam döngüsü) sorularına yanıt verilmeye çalışılmıştır.

Araştırma verisi yönetimi, verinin araştırma döngüsüne girmesinden sonuçların dağıtım ve arşivlenmesine kadar olan süreçteki düzenlemeler ile ilgilidir. Araştırma verisi yönetimi, sonuçların güvenli bir şekilde doğrulanması ve mevcut bilgiler üzerine yeni ve yenilikçi araştırmaların gerçekleşmesini sağlamayı amaçlamaktadır (Whyte ve Tedds, 2011).

Araştırma verilerinin yönetiminin kapsamı şu şekildedir (Tonta, 2013b):

- Veri yönetimi planlaması
- Veri yaratma/tanımlama
- Veri belgeleme (veri toplama ve yönetim için standartlar ve yöntemler)
- Veriye erişim, kullanım ve paylaşım
- Etik ve entelektüel mülkiyet
- Depolama ve yedekleme
- Uzun süreli veri koruma stratejisi

2.3.1 Veri Yönetim Politikaları

Araştırma verileri bireysel olarak araştırmacılar tarafından uzun vadede kusursuz olarak korunamamaktadır. Uzun vadede verinin korunması için politikalara ihtiyaç vardır (Vines ve diğerleri, 2014).

Araştırma verileri yönetimine ilişkin politikalar; kütüphane, bilgi teknolojileri hizmetleri, araştırma ofisi, hukuk ofisi, akademik personel gibi birçok paydaşın katkı sağlayarak geliştirdikleri bir süreçtir (Cox, Kennan, Lyon ve Pinfield, 2017, s. 2186). Bu politikalar ulusal mevzuat, ulusal politikalar, fon sağlayıcıların politikaları, bölgesel ve yerel politikalar, kurumsal politikalar, dergi politikaları gibi çok çeşitli kaynakta yer alabilmektedir (Briney, 2015, s. 34).

Üniversitelerin AVY ilişkin politika, hizmet ve altyapı geliştirmeye yönelik hedeflerinde kütüphaneler de rol almaya başlamıştır. Araştırma çıktılarının korunması, veri yönetimi ve derleme, açık erişim gereksinimleri, üst veri kullanımı gibi konulara yönelik bilgileri ile kütüphaneciler AVY sürecine katkı sağlayabilmektedir. Birçok araştırma kurumu kütüphanesi veri hizmetlerini, veri kütüphanecileri ve araştırmacıların iş birliğiyle geliştirmektedir. Ancak kütüphanelerin bu kapsamdaki rolü halen belirsizliğini korumaktadır (Cox ve diğerleri, 2017, s. 2182; Stewart ve Crossley, 2013, s. 14; Briney, Goben ve Zilinski, 2015, s. 2).

İngiltere’de birçok fon sağlayıcı kuruluş araştırmacılardan araştırma çıktılarına erişim, yönetim ve uzun süreli koruma konularının yer aldığı bir bildirim göndermelerini beklemektedir (Pryor, 2012, s. 2). Benzer şekilde, Amerika’da birçok araştırma kurumu,

araştırmacılar veri yönetim ve paylaşım planlarını talep etmektedir. Ulusal Sağlık Enstitüsü (National Institutes for Health-NIH) ve Ulusal Bilim Kurumu (National Science Foundation-NSF) bu kurumlardan en bilinenleridir. Üniversiteler proje başvurularında istenen şartları yerine getirebilmeleri için araştırmacılara destek olmaktadır (Jones, 2012, s. 61). Birçok kütüphane araştırmacılara veri yönetimine ilişkin yardım etmekte; Araştırma Kütüphaneleri Derneği (Association of Research Libraries-ARL) ise, kütüphanecilere destek sağlamaktadır (Hswe ve Holt, 2010).

Almanya'daki bir araştırma kuruluşu olan DFG (Deutsche Forschungsgemeinschaft-German Research Foundation), fon sağladığı araştırmalara açık erişim ve koruma konularında destek sunmakta; araştırma verilerinin nasıl yönetileceği ve paylaşılacağını dikkate almaktadır. Hollanda'da iki ana araştırma kurumu (The Royal Netherlands Academy of Arts and Sciences-KNAW ve Netherlands Organisation for Scientific Research-NWO), araştırma verilerine ve yayınlarına açık erişimi teşvik etmektedir (Jones, 2012, s. 60).

DCC (Digital Curation Center-Dijital Kürasyon Merkezi) tarafından bazı fon sağlayıcıların politika belgelerinin kapsamı incelenmiştir ve karşılaştırmalı bir tablo oluşturulmuştur. Bu tabloda yer alan öğeler dikkate alınarak politika belgesinin kapsamında olması gereken unsurlar şu şekilde belirlenmiştir:

- Veri: Veri seti politikası veya elektronik kaynağa erişim ve bakım konusunun bildirilmesi
- Zaman sınırı: Araştırma çıktılarının korunması veya içeriğin erişilebilir olması için gerekli sürenin ayarlanması.
- Veri planı: Projeye destek alabilmek için veri derleme, yönetim ve paylaşım konularını kapsayan bir plan gerekliliği.
- Erişim/paylaşma: Veri paylaşımı ve tekrar kullanım için verinin açık erişim dergilerde, e-arşivlerde depolanması.
- Uzun süreli derleme: Araştırma çıktılarının uzun süreli koruma ve bakımı için gerekli şartlar.
- Gözetim: Şartların sağlanıp sağlanmadığının gözlenmesi.
- Rehber: Sıkça sorulan sorular, uygulama rehberleri, personel desteği gibi konuları içeren bir rehber hazırlama.

- E-arşiv: Yayımlanmış araştırma çıktılarının erişilebilir olduğu bir e-arşiv sunma.
- Veri merkezi: Yayımlanmamış elektronik kaynakları veya verileri yönetmek için bir veri merkezinin sağlanması.
- Maliyet: Yayın ücretleri ve veri yönetimi/paylaşım maliyetlerini karşılamak için istekli olunması (DCC, 2017).

Fon sağlayıcı kuruluşların hibe gereksinimini karşılamak için araştırmacılar araştırma verilerini iyi yönetebilmelidir. Araştırma sürecini veri yaşam döngüsü aşamalarına göre şekillendirmek veri yönetimi açısından kolaylık sağlamaktadır. Bunun için bir sonraki bölümde veri yaşam döngüsü konusuna değeri verilmiştir.

2.3.2 Veri Yaşam Döngüsü

Araştırma verilerine yönelik uygulamalar veri yaşam döngüsündeki aşamaları takip etmektedir. Birçok veri merkezinin, üniversitelerin veya organizasyonların benimsediği farklı veri yaşam döngüsü modelleri bulunmakla birlikte, bu çalışma için benimsenen veri yaşam döngüsü modeli Şekil 2’de sunulmuştur. Veri yaşam döngüleri veri yönetimindeki süreci araştırmacılara aşama aşama göstermesi açısından faydalıdır.



Şekil 2. Veri yaşam döngüsü

Araştırma henüz gerçekleştirilmeden planlama aşamasıyla başlayan veri yaşam döngüsü, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma, keşfetme ve

yeniden kullanma aşamalarıyla sürekli devam etmektedir. Oluşturulan veri yaşam döngüsü modeline dayalı AVY uygulamaları [Bölüm 3](#)'te sunulmuştur.

2.4 VERİ PAYLAŞIMI

Bilimsel yayınların araştırma verileriyle birlikte yayımlanması kamu yararı sağlamasının yanı sıra araştırmacı ve fon sağlayıcı kuruluş açısından da birçok avantaj sağlamaktadır. Bu bölümde veri paylaşımının yararları açıklanmıştır.

Veri duplikasyonunu azaltır: Benzer verileri tekrar toplamak yerine mevcut verileri kullanarak yapılan araştırmaların geliştirilmesine olanak tanımaktadır (USGS, t.y.). Veri toplamak için harcanan zaman ve maliyetten tasarruf sağlar.

Verinin yeniden kullanımını sağlar: Belirli bir konuda yapılmış olan birden fazla çalışmanın sonuçlarının birleştirilmesiyle araştırmacıların meta-analizler yapmalarına da olanak sağlar (USGS, t.y.). Diğer araştırmacılar çeşitli yaklaşımlar ve teknikler kullanarak farklı bir bakış açısı geliştirebilir (Chawinga ve Zinn, 2019, s. 113). Örneğin, tıp tarihine Gage vakası olarak geçen ve olayın ilginçliği nedeniyle Harvard Üniversitesi Warren Tıp Müzesi'nde korunan kafatasıyla ilgili veriler yeni teknoloji yardımıyla tekrar kullanılarak yeni keşiflere neden olmuştur. Kısaca Gage vakasından bahsetmek gerekirse, 1848 yılında Phineas Gage bir iş kazası geçirmiş, elmacık kemiğinden giren demir bir çubuğun kafatasını delerek çıkması sonucu beyninin bir kısmı parçalanmasına rağmen hayatta kalmıştır. Gage, kaza sonucu beyindeki hasara rağmen motor işlevlerinin hasar görmemesi ancak kişiliğindeki ciddi değişiklikler olması nedeniyle tıp literatürüne girmiştir. Gage'in ölümünden 150 yıl sonra, Dr. Damasio ve çalışma arkadaşları kafatası fotoğrafları ve koordinatlarını kullanarak kafatasının üç boyutlu görüntüsünü elde etmişler, beyin hangi kısımlarının parçalandığını belirleyebilmişlerdir. Gage'in beyni korunamamasına rağmen elde edilen veriler ve teknoloji yardımıyla kaza bilgisayar ortamında tekrarlanmış, kişilikteki değişiklikten beyin prefrontal korteks denilen bölgesinin sorumlu olduğu bulunmuştur (Karaçay, 2018). Yeniden kullanıma başka bir örnek de astronomi alanından verilebilir. 1998 yılında Hubble Uzay Teleskobu ile çekilmiş bir fotoğraf, 11 yıl sonra yeni bir teknik kullanarak analiz edilmiş ve güneş sistemi dışında yeni bir gezegen keşfedilmiştir (Space.com, 2009).

Görünürlüğü, etkiyi ve atıf oranlarını artırır: Verilerin erişilebilir ortamlarda sunulması hem çalışmanın hem de verilerin görünürlüğünü ve etkisini artırmakta, bu da atıf oranlarına yansımaktadır (Piwowar, Day ve Fridsma, 2007).

İş birliği olanağı yaratır: Aynı konuda çalışan araştırmacıların birbirleriyle bağlantı kurarak ortak çalışmalar yapmasına vesile olabilmektedir (USGS, t.y.). Küresel çapta iş birliği olarak İnsan Genomu Projesi örnek gösterilebilir.

Şeffaflık, doğrulanabilirlik ve güvenilirlik ortamı hazırlar: Araştırma çıktılarının verilerle birlikte yayımlanması şeffaf bir yapı sunmakta, doğrulanabilirliği sağlamakta ve güvenilirliği artırmaktadır. 2000 ve 2010 yılları arasında PubMed veri tabanından geri çekilen 742 araştırma makalesinin geri çekilme nedenlerinin araştırıldığı bir çalışmada, çalışmaların yaklaşık dörtte birinin veri fabrikasyonu (data fabrication) veya tahrifatını (data falsification) içeren iyi niyetli olmayan davranışlar içermesi nedeniyle geri çekildiği belirlenmiştir (Steen, 2011, s. 250). Yine bu çalışmada, geri çekilen makale sayısının giderek arttığı gözlemlenmiştir. Tespit edilen çalışmalar, buzdağının görünen kısmı olabilir. Araştırma sayısının giderek arttığı günümüzde mevcut kontrol mekanizmasıyla etik dışı davranışların tespiti oldukça zordur. Bilimsel çıktı yayımlanmadan önce, araştırma hatalarını engellemek amacıyla doğrulama ve onaylama için hakemler bazen verileri isteyebilmektedir (Chawinga ve Zinn, 2019, s. 115). Ancak sadece hakem değerlendirmesi hataların veya uydurma verilerin tespit edilmesinde yeterli değildir. Çünkü duplikasyon ve fabrikasyon nedeniyle geri çekilen bazı makaleler ancak yayımlanma işleminden sonra ortaya çıkmaktadır (Al ve Soydal, 2015, s. 29). Şeffaf bir yapı sunan veri paylaşımı; hakem değerlendirmelerinin de ötesinde verilerin diğer araştırmacılar tarafından incelemesine olanak sağladığı için bilimsel yayınların güvenilirliğini artırma potansiyeline sahiptir (Nicol, Caruso ve Archambault, 2013, s. ii).

Ekonomik getiri sağlar: 2014 yılında İnsan Genomu Projesinin ekonomik getirisinin yapılan yatırımdan 7,7 kat fazla olduğu belirlenmiştir (The data harvest, 2014, s. 5). Günümüzde söz konusu ekonomik getiri daha da artmış olabilir. Veri paylaşımının ekonomik getirisinin farkında olan kurum ve kuruluşlar hem araştırmacılara teşvik sunmakta hem de veri arşivleri için yatırımlar yapmaktadır. Almanya'da bir araştırma enstitüsü (Berlin Institute of Health), araştırmacıların başkaları tarafından yayımlanmış verileri tekrar kullanarak bir klinik öncesi çalışma veya bir makale yayımlamaları halinde araştırmacılara 1000 € sunarak akademik teşvikleri yeniden yapılandırmaktadır (Grove, 2020).

Benzer şekilde, veri arşivleme altyapısına yapılacak finansal yatırım da önemli ölçüde bir bilimsel getiri sağlamaktadır (Piwowar, Vision ve Whitlock, 2011, s. 285). Uzun süreli korumaya yardımcı olan veri arşivleri, kalıcı tekil tanımlayıcılar sağlayarak verilere atıf verilebilmesini ve verilerin keşfedilebilmesini kolaylaştırmaktadır (Princeton University Library, 2020). Veri paylaşım ortamı olan veri arşivlerine yapılan yatırımın, araştırmaya yapılan yatırıma göre harcanan dolar başına etkileyici biçimde daha fazla yayına yol açtığı hesaplanmıştır. Hibe olarak ödenen 400.000 dolarlık bir yatırım ortalama 16 araştırmayı desteklemekte, veri arşivlerine yapılacak aynı tutardaki yatırım ise paylaşılan verilerin tekrar kullanımıyla yaklaşık 1000 tane araştırma çıktısının oluşmasını sağlamaktadır (Piwowar ve diğerleri, 2011, s. 285).

Bilimsel verilerin yeniden kullanımını destekleyen altyapının iyileştirilmesi gereği, 2016 yılında akademik camia, yayıncılar, fon sağlayıcı kuruluşlar gibi paydaşlar tarafından FAIR ilkeleri geliştirilmiştir (Wilkinson ve diğerleri, 2016).

2.4.1 FAIR İlkeleri

Araştırma verilerinin yeniden kullanımını sağlamak için verilerin FAIR ilkelerine uygun hale getirilmesi gerekmektedir. FAIR, İngilizce Findability, Accessibility, Interoperability ve Reusability kavramlarının baş harflerinden oluşan bir kısaltmadır. FAIR ilkelerinin kapsamı aşağıda açıklanmıştır.

Bulunabilirlik (Findability): Araştırmacının gereksinim duyduğu veriyi kolayca bulabilmesini ifade eden bu ilke, verinin doğrudan veri arşivlerinde veya akademik arama motorlarında (Google Dataset, Scholar, MS Academics) taranması sonucu bulunabilmesi ile ilgilidir (Bote ve Termens, 2019). Verinin standart üst verilerle tanımlanması, verilere kalıcı bir tanımlayıcı atanması, verinin aranabilir bir ortama kaydedilmesi ve indekslenmesi eylemlerini içerir (FORCE11, 2011-2021).

Erişilebilirlik (Accessibility): Kullanıcıların veriye ulaşması ve indirebilmesini ifade etmektedir. Bu süreçte, standartlaştırılmış iletişim protokolü kullanılarak (üst)verilere erişilebilir, gerektiğinde bir kimlik doğrulama ve yetkilendirme süreci uygulanabilir (FORCE11, 2011-2021).

Birlikte Çalışabilirlik (Interoperability): Üst verilerin standartlara uygun hazırlanması ve OAI-PMH gibi birlikte çalışabilirlik protokolleri ile uyumlu olması sistemlerin veri

alışverişi yapmasını sağlamaktadır. Verilerin görünürlüğünün ve etkisinin artması birlikte çalışabilir sistemlerle mümkün olup ülkemizdeki bazı arşivlerin ve YÖK Ulusal Tez Merkezi'nin bu şartları sağlayamadığı için diğer arşivler veya arama motorları tarafından harmanlanamadığı bilinmektedir (Tonta ve Akbulut, 2019). Verilerin keşfedilebilirliği harmanlama ile mümkündür.

Yeniden Kullanım (Reusability): Verilerin lisanslanması ve ilgili kaynaklarla ilişkilendirilerek yayımlanması sonucu yeniden kullanıma sunulması ilkesidir. Şimdiye kadar tanımlanan ilkelerin tümü bu ilkenin gerçekleşmesine hizmet etmektedir (FORCE11, 2011-2021). FAIR ilkeleri birbirini tamamlayan bir bütünün parçasıdır. Yeniden kullanım için verilerin keşfedilebilir olması, verilerin keşfedilebilir olması içinse üst veri standartlarına ve protokollere uygun olarak hazırlanan tanımlayıcı bilgilerin veri arşivinde kaydedilmesi gerekmektedir.

2.5 İLGİLİ ÇALIŞMALAR

Literatürde; araştırma verilerine yönelik e-bilim, açık veri, araştırma verilerinin yönetimi, planlama, politika geliştirme, veri okuryazarlığı gibi çeşitli açılardan konunun ele alındığı görülmektedir. Araştırma konusu kapsamında; kullanıcı araştırmaları, kütüphanelerdeki uygulamalar, politika belgelerinin incelenmesi ve teorik bilgiler içeren birçok yayın bulunmaktadır.

Araştırma Verilerinin Yönetimi (AVY) konusuna ilişkin Avustralya, Kanada, Almanya, İrlanda, Hollanda, Yeni Zelanda ve Birleşik Krallık'taki üniversite kütüphanelerinin araştırma verileri yönetimini nasıl destekledikleri ve hizmetlerini gelecekte nasıl geliştirmeyi planladıklarına ilişkin uluslararası bir çalışma yapılmıştır. Araştırma verilerine yönelik olarak sunulan hizmetleri hem ülkelere hem de önceki çalışmalara göre karşılaştıran bu çalışmada, hizmet gelişiminin sınırlı olduğu, kaynak yetersizliği, iş birliği eksikliği ve beceri eksikliği gibi sorunlar belirlenmiştir (Cox ve diğerleri, 2017).

Herhangi bir coğrafi alan veya disiplin kısıtlaması olmaksızın, 1995-2016 yılları arasında AVY ile ilgili yayınlar incelenmiş, 2010 yılından sonra konuyla ilgili yapılan çalışmaların arttığı gözlemlenmiştir. Coğrafi bölgelere göre dağılımı incelendiğinde çalışmaların yaklaşık yarısının Kuzey Amerika'da yapıldığı saptanmıştır. Disipline göre dağılıma bakıldığında, çalışmaların yaklaşık %10'u bilgi bilim ve kütüphanecilik alanı ile ilgili

olmakla birlikte çoğunlukla veri paylaşımı, erişim kontrolü gibi alt konuları kapsayan “veriye erişim” konusu ile üst veri oluşturma/dokümantasyon, veri arşivlemeyi kapsayan “veri koruma” konularına odaklanıldığı belirlenmiştir (Perrier ve diğerleri, 2017).

Ekim 2012’de, *USA News* tarafından hazırlanan dünyanın en iyi üniversitesi sıralamasında ilk 100’de yer alan, 87 üniversitenin akademik kütüphanelerinin araştırma verilerine yönelik sundukları hizmetlere ilişkin mevcut durumu saptamak ve bu konuda kütüphanelere yol gösterici bazı stratejilere değinmek amacıyla bir çalışma yapılmıştır. Araştırma sonuçlarına göre, üniversite kütüphanelerinin %57,5’inin, veri yönetimi ve paylaşım hizmetleri sunduğu ortaya çıkmıştır. Kütüphanelerin araştırmacılarla ve idari birimlerle iş birliklerini güçlendirmelerinin önerildiği bu çalışmada, kütüphanelerin veri paylaşımı anlayışının benimsenmesinde ve veri yönetim planı hazırlamada araştırmacılara yardımcı olmaları gerektiği vurgulanmaktadır (Si, Xing, Zhuang, Hua ve Zhou, 2015).

Karayıpler’de bir üniversite kütüphanesi tarafından, araştırma verilerinin yönetimindeki güncel uygulamaları ve bu konuda kütüphanenin rolünü belirlemek amacıyla bir çalışma yapılmıştır. Katılımcılar, çoğunlukla flash diskler, kurum dışı sabit diskler ve kendilerine e-posta gönderme yöntemlerini kullanarak araştırma verilerini depolamakta ve yedeklemektedir. Bulut çözümler ve ikinci bir bilgisayar kullanarak depolama ve yedekleme yapma çok tercih edilmemektedir. Katılımcıların veri yönetimine ilişkin desteğe ihtiyaç duydukları temel konu, uzun süreli koruma için arşivleme olarak belirlenmiştir. Katılımcıların veri yönetimine yönelik bilgi ve deneyim düzeylerinin, farkındalıklarının düşük olduğu, araştırmacıları destekleme konusunda kütüphaneye rol düştüğü belirtilmektedir (Renwick, Winter ve Gill, 2017).

Kaliforniya Politeknik Devlet Üniversitesi’nde yapılan bir çalışmada, araştırmacıların veri toplama konularına yönelik olarak farkındalıklarını, davranışlarını ve tutumlarını belirlemek amacıyla bir anket uygulanmıştır. Katılımcılar, araştırma verilerinin sorumluluğunun kendilerine ait olduğunu düşünmektedir. Ayrıca öğretim üyelerinin çoğunluğu veri yönetimine yönelik olarak rehberlik ihtiyacı duymaktadır (Scaramozzino, Ramirez ve McGaughey, 2012, s. 360).

Buyss ve Shaw (2015), Şikago’da bir üniversitedeki araştırmacıların veri yönetimine ilişkin mevcut durumun ve gelecekteki ihtiyaçların belirlenmesi amacıyla bir çalışma yapmışlardır. Araştırmacıların yarısından fazlasının yönetim planına sahip olmadığı;

bunun temel nedeninin veri yönetiminin planlanması hakkında bilgi eksikliği olduğu belirlenmiştir. Özellikle araştırma verilerine uzun süreli erişim sağlama ve koruma, veri depolama ve yedekleme konularına ilişkin, araştırmacıların eğitim ihtiyacı duydukları gözlemlenmiştir.

Oregon Devlet Üniversitesi'ndeki araştırmacıların veri uygulamalarına yönelik yapılan bir araştırmada, birçok öğretim üyesinin verileri tanımlarken üst veri kullandığı belirlenmiştir. Ancak araştırmacılar üst veri standartları yerine kendi araştırma gruplarınca standartlaştırılan üst verileri kullanmaktadır. Araştırmacıların, yaklaşık %20'sinin listelenen üst veri standartlarını bilmediği ortaya çıkmıştır (Whitmire, Boock ve Sutton, 2015, s. 390).

Avrupa'da yapılan bir araştırmada, araştırmacıların %87'si araştırma verilerine erişimde sorun olduğunu düşünmektedir. Araştırma verisine erişimdeki en önemli engel ulusal/bölgesel strateji/politikaların araştırma verilerine erişim için uygun olmayışdır (European Commission, 2012).

Çin'de bulunan on bir üniversitenin araştırmacılarına uygulanan anket sonuçlarına göre, araştırmacıların veri hizmetlerine ihtiyaç duydukları ortaya çıkmıştır. Araştırmacıların %65'ten fazlasının verilerini bir kez kaybettiği; araştırma verilerinin %73'ünün erişilebilir durumda olmadığı; katılımcıların %50'den fazlasının verilerini sadece 2 yıl süresince koruduğu ve kalıcı korumanın dikkate alınmadığı sonucuna ulaşılmıştır (Liu ve Ding, 2016). Başka bir çalışma kapsamında; araştırmacılar verilerin depolanması, çalışma belgelerinin ve veri setlerinin yönetimi konusunda sorun yaşadıklarını bildirmiştir. Verilerin düzensiz olması yararlı bulguların kaybolmasına veya erişilemez olmasına neden olmaktadır (Kroll ve Forsman, 2010, s. 5)

Berman (2017), Vermont Üniversitesi'nde (ABD), bilim insanlarının veri yönetim davranışlarını araştırmış ve araştırmacıların desteğe ihtiyaçları olduğu alanları belirlenmiştir. Vermont Üniversitesi'nde araştırma verileri yönetimi için altyapı, üst veri, veri analizi ve istatistiksel destek olmak üzere dört ana konuda hizmetlerini geliştirilmesi önerisinde bulunmuştur.

2014 yılında yapılan bir çalışmada, yayımlanmasının üzerinden belirli bir yıl (en az iki, en çok 22) geçmiş 516 makalenin araştırma verilerinin kullanılabilirliği araştırılmıştır. Yayımlanan verilerin kullanılabilirliği, yılda %17 oranında düşmektedir. Verilerin hala var olma olasılığı ile makalenin yaşı arasında negatif bir ilişki bulunmuştur. Yayımlanma

tarihi eski makalelerin verilerine erişimin azalmasının ana nedeni, depolama ortamına erişilememesi ve böylece verinin kaybolma oranının artması olarak bildirilmiştir (Vines ve diğerleri, 2014). Birçok derginin yayın politikalarında ham veri setlerinin paylaşılması şartı bulunmaktadır. Açık erişimli dergileri bünyesinde bulunduran PLoS (The Public Library of Science) dergileri de istek olması halinde yazarın verilerini paylaşmasını istemektedir. Savage ve Vickers (2009), PLoS dergilerinin veri paylaşım politikasına dayanarak yaptıkları çalışmada, iki PLoS dergisinde (PLoS Medicine ve PLoS Clinical Trials) yayın yapan on yazardan veri isteğinde bulunmuş ve sadece bir yazar verisini paylaşmıştır. Tenopir, Birch ve Allard (2012) veri yoğun bilimin araştırmacılar için zorluklar yarattığını, eğitim almak için zamanları olmayan bilim insanları için veri yönetimi beceri setine ihtiyaç duyulduğunu bildirmiştir. Bilim insanlarının veri paylaşımına ilişkin yaklaşımlarını belirlemek amacıyla yapılan bir çalışmada, veri paylaşımının önündeki en önemli engellerden bir tanesinin kurumsal rehberlik ve destek eksikliği olduğu belirlenmiştir (Tenopir ve diğerleri, 2011)

E-bilim hakkında öğrenmeye ilgi duyan 63 kütüphaneciyi kapsayan bir araştırmaya göre, kütüphaneciler veri yönetimi hizmetlerine yönelik becerilerini geliştirmek istemekte, özellikle büyük veri setlerinin tanımlanması ve derlenmesi hakkında eğitim ihtiyacı duymaktadır (Creamer, Morales, Crespo, Kafel ve Martin, 2012).

Araştırma verileri hizmetlerine yönelik kütüphane yöneticileri ve kütüphanecilere iki ayrı anket uygulanarak bu hizmetlerin uygulanmasına yönelik görüşler belirlenmek istenmiştir. Kütüphanelerin çoğunda araştırma verileri hizmetlerinin uygulanmadığı ancak bu konunun planlandığı görülmektedir. Kütüphaneciler, veri yönetim hizmetlerine dair becerilerinin gelişmesi için kütüphanenin sürekli eğitim sunmasına ihtiyaç duymaktadırlar (Tenopir, Sandusky, Allard ve Birch, 2014, s. 89).

2.5.1 Türkiye'deki Durum

Türkiye'de araştırma verilerinin yönetimi konusunda yapılan çalışmalar az olmakla birlikte, konuyla ilgili bazı kişi ve kurumların farkındalığı artırmak için çabaladığı görülmektedir. Bu bölümde araştırma verileri konusunda ülkemizde yapılan çalışmalara, düzenlenen etkinliklere ve Türkiye'deki araştırmacılar üzerinde yapılan bilimsel çalışmaların çıktıklarına yer verilmiştir.

Türkiye'nin dâhil olduğu bazı AB projeleri kapsamında araştırma verisi ve açık veriye yönelik çalışmalar yapılmaktadır. Bilimsel bilgi altyapısını genişletme amaçlı OpenAIRE2020; açık erişim politikaları geliştirme amaçlı MedOAnet ve PASTEUR4OA projeleri AB projelerine örnek olarak verilebilir (Gürdal ve Bitri, 2015). MedOAnet projesi kapsamında yapılan bir ankette açık erişim politika belgelerinde araştırma verilerinin yönetiminden bahsedilmediği saptanmıştır (Tonta, 2012; Tonta, 2013a). Benzer şekilde, PASTEUR4OA projesi kapsamında Türkiye'deki kurumsal arşivlerin açık erişim politikaları incelenmiş ve politika belgelerinin %60'ında araştırma verilerinden söz edilmediği belirtilmiştir (Tonta, 2015).

Türkiye'deki çevre bilimciler ile bilgi ve belge yönetimi alanındaki akademisyenlerin veri uygulamalarını belirlemek için görüşme tekniği kullanılarak bir çalışma yapılmıştır. Katılımcılar araştırma verilerini kişisel olarak yedeklemekte ancak veri paylaşımı konusunda çekingen davranmaktadırlar. Verilerin yönetilmesinde kurumsal desteğin olmadığı görülmektedir. Bilgi bilimcilere göre, veri yönetimi ve veri paylaşımı sadece akademik çevrede konuşulmaya başlanmıştır (Allard ve Aydınoglu, 2012).

Tonta (2013b), 2. Ulusal Açık Erişim Çalıştayı'nda araştırma verilerinin yönetimi ve açık veri konulu sunumda TÜBİTAK'ın, YÖK'ün ve üniversitelerin veri yönetim politikaları olmadığını vurgulamıştır.

Kasım 2014'te Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü (HÜ BBY) ile Goethe-Institut Ankara tarafından Uluslararası Araştırma Verilerinin Yönetimi Çalıştayı gerçekleştirilmiştir. Tonta ve Al (2015) çalıştaydaki izlenimlerini, yapılan sunumları ve paylaşılan bilgileri görüş olarak yayımlamışlardır. Çalıştayda; araştırma verisi kavramı, araştırma verisi yönetimi ile ilgili sorunlar, entelektüel mülkiyet hakları, yasal düzenlemeler, veri paylaşımı gibi konular tartışılmıştır. Yine HÜ BBY Bölümü tarafından 2014 yılında düzenlenen uluslararası bir sempozyumun ana teması araştırma verileri yönetimi ve bilgi keşfi olarak belirlenmiştir.

2015 yılı Akademik Bilişim Konferansı'nda sunulan bir bildiri; AB Ufuk 2020 programı kapsamında araştırma verisi, araştırma verisi yönetimi, işlevi, kapsamı ve yararları değerlendirilmiştir. Konunun önemine dair Türkiye'de farkındalık yaratılması amaçlanmıştır (Gürdal ve Bitri, 2015).

Açık erişim konusu çerçevesinde yaşanan gelişmeleri takip etmek ve ulusal politikalar oluşturmak için bilimsel çalışmalar yapmak amacıyla TÜBİTAK ULAKBİM bünyesinde

Açık Bilim Komitesi oluşturulmuştur. Komitenin görev ve sorumlulukları arasında açık bilim, açık erişim, açık veri konularını takip etmek ve bu konulara ilişkin mevzuat çalışması yapmak yer almaktadır (Cahit Arf Bilgi Merkezi, 2017).

Türk araştırmacıların AVY alanındaki uygulamalarını belirlemek, beceri ve farkındalık düzeylerini saptamak ve Türkiye’deki güncel politikaları değerlendirmek amacıyla yapılan bir çalışmada, 25 üniversitedeki 532 araştırmacıya anket uygulanmıştır. Araştırmacıların, üst veri standartlarına göre veri toplama/derleme veya veri depolama gibi araştırma verileri yönetimi konusunda bilgi ve beceri eksikliği olduğu ortaya çıkmıştır. Araştırma verileri yönetimi konusunda farkındalık düzeyi en yüksek katılımcıların araştırma görevlileri olduğu belirlenmiştir. Katılımcıların çoğunlukla veri paylaşımı konusunda olumlu davranışa sahip olduğu, araştırma verisini paylaşmayanların en çok “verilerinin diğerlerinin erişimini istememesi” nedeniyle paylaşmadığı belirlenmiştir (Aydınoglu ve diğerleri, 2017).

Türkiye’de ilk defa TÜBİTAK Ulusal Akademik Ağ ve Bilgi Merkezi (ULAKBİM) öncülüğünde 26-27 Eylül 2017 tarihinde *Araştırma Verileri Yönetimi* konusunda bir eğitim programı düzenlenmiştir. Eğitime TÜBİTAK’ın farklı birimlerinde ve çeşitli üniversitelerde çalışan 100’den fazla kişi katılım sağlamıştır.

11 Eylül 2018 tarihinde Boğaziçi Üniversitesi’nde “*Araştırma Verilerinin Yönetimi ve Açık Bilim Çalıştayı*”, 12-13 Ekim 2017 tarihlerinde ise Ankara’da “*Açık Veri Portalı Çalıştayı*” düzenlenmiştir (Araştırma Verilerini Yönetimi ve Açık Bilim Çalıştayı Gerçekleştirildi, 2018; Açık Veri Portalı Çalıştayı Düzenlendi, 2018)

Beş kıtada gerçekleştirilen çok uluslu bir araştırma projesinin parçası olarak Türkiye’deki araştırmacılar üzerinde yapılan bir çalışmayla, araştırmacıların veri yönetimi konusuna ilişkin mevcut durumu saptanmıştır. Araştırmacıların veri yönetim planı uygulamaları hakkındaki farkındalık ve görüşlerinin de yer aldığı bu çalışmada, araştırmacıların çok azının veri yönetim planı kullandığı belirlenmiştir. Ayrıca araştırmacıların büyük çoğunluğu (%73) kurumlarının veri yönetim planının olup olmadığını bilmemektedir (Ünal ve Kurbanoglu, 2018, s. 299).

Araştırma verilerinin yönetimi konusunda Türkiye’deki gelişmeler ışığında, YÖK Açık Bilim Politikasının kapsamı, TÜBİTAK ULAKBİM tarafından hazırlanan Araştırma

Verileri Yönetimi Eğitim Portalı ve YÖK tarafından üniversitelere gönderilen örnek politika belgesi incelenmiştir.

TÜBİTAK destekli proje çıktıları ile projede üretilen araştırma verilerinin yönetimi, saklanması, arşivlenmesi, derlenmesi ve korunması amacıyla oluşturulan TÜBİTAK Açık Bilim Politikası 14.03.2019 tarihinde yürürlüğe girmiştir. Politika başlığı bölümünde, yayınların ve araştırma verilerinin yönetilmesi, saklanması, arşivlenmesi ve korunması konularının bu politikanın çerçevesini oluşturduğu ifade edilse de, araştırma verilerine yönelik bilgiler politikada yer almamaktadır. İlkeler bölümünde yer alan ve eksik olduğu düşünülen hususlar şu şekildedir:

- TÜBİTAK, kendi desteğiyle üretilmiş yayınların TÜBİTAK Açık Arşivinde depolanmasını zorunlu tutmaktadır. Bu yayınlara ait verilerin TÜBİTAK Açık Arşivinde depolanmasına yönelik bir bilgi bulunmamaktadır.
- Yayınlarla ait üst verilerin açık, taranabilir ve makine tarafından okunabilir olması gerektiği ifade edilmekte ancak üst veri standartlarından bahsedilmemektedir.
- Veri yönetimi planı oluşturmak için, TÜBİTAK tarafından şablon ve kılavuz hazırlanacağı belirtilmektedir ancak henüz bir şablon bulunmamaktadır.
- TÜBİTAK desteğiyle üretilmiş yayınların ve yayınlara ilişkin araştırma verilerinin açık erişim olması önerilmekle birlikte, bu verilerin nerede depolanacağına ilişkin bir bilgi sunulmamaktadır.
- Araştırma verileri yönetimine ilişkin araştırmacılara destek sunmayla ilgili bir madde bulunmamaktadır.

Genel olarak; verilerin sahibi kim, veriler nerede-nasıl saklanacak, verilerin güvenli bir şekilde saklanması sorumluluğu kime ait, uzun süreli koruma nasıl sağlanacak gibi birçok sorunun yanıtı bulunmamaktadır. Bu politika, açık erişim adına güzel bir gelişme olmakla birlikte, araştırma verileri yönetim politikası değildir.

Araştırma verileri yönetimi konusunda araştırmacılara, proje yöneticilerine, kurum ve kuruluşlara destek olmak amacıyla TÜBİTAK ULAKBİM tarafından Araştırma Verileri Yönetimi Eğitim Portalı erişime açılmıştır. Bu portalda, araştırma verilerinin yaratılmasından, tekrar kullanılmasına kadar olan süreci kapsayan araştırma verilerinin yönetilmesi konusu aşama aşama ayrıntılı olarak anlatılmaktadır. Eğitim portalından

APERTA'ya geçiş bulunmaktadır. APERTA, TÜBİTAK tarafında yürütülen veya desteklenen projelerin araştırma çıktıları ve verilerinin erişimine olanak sağlayan TÜBİTAK açık arşividir (Araştırma Verileri Yönetimi Eğitim Portalı, t.y.). TÜBİTAK-ULAKBİM Açık Ders Platformunda “Araştırma Verilerinin Yönetimi” Dersi içeriği oluşturulmuştur (<https://acikders.ulakbim.gov.tr/course/view.php?id=2>)

Açık bilim ve açık erişim çalışmaları çerçevesinde YÖK, üniversitelere bir Açık Bilim Politikası hazırlanması gerektiğini bildirmiştir (15.02.2019 tarihli ve 12867 sayılı resmi yazı). Ayrıca Mayıs 2019'da tüm üniversitelere, yardımcı olmak amacıyla örnek bir açık bilim politika belgesi göndermiştir. Üniversitelere gönderilen, *Üniversiteler İçin Model Açık Bilim Politikası* adlı belge, TÜBİTAK Açık Bilim Politikasına göre daha kapsamlıdır.

26.03.2019 tarihinde yürürlüğe giren İzmir Yüksek Teknoloji Enstitüsü Açık Bilim Politikası, bu konuda ilk üniversite politikası olma özelliği taşımaktadır. Bu politika belgesi, YÖK tarafından üniversitelere gönderilen belgeyle benzerlikler göstermektedir. *Üniversiteler İçin Model Açık Bilim Politikası* adlı belgede, araştırma verilerinin uygun bir arşive konulması zorunlu tutulmakta, veri yönetim planları için bir veri yöneticisi atanması gerektiğinden, verilerin saklanması, depolanması ve uzun süreli korunması sorumluluğunun üniversiteye ait olduğundan bahsedilmektedir. En üst kuruluş olarak YÖK'ün araştırma verileri yönetimi konusunda model politika geliştirmenin ötesinde, üniversitede çalışan personelin gerekli eğitimi almasını sağlaması, hazırlanmış olan eğitim portalının tanıtılması ve kullanılmasını teşvik etmesi, gerekirse uygulama için TÜBİTAK ile iş birliği yaparak üniversitelere yol göstermesi beklenmektedir.

Araştırma verileri konusunda düzenlenen etkinlikler incelendiğinde, etkinlik sayısının ve bu konuda farkındalığın giderek arttığı görülmektedir.

2019 Mart ve Nisan aylarında, “*FAIR Veri Prensipleri ve Araştırma Verilerinin Yönetimi*” ve “*Türkiye’de Açık Veri Politikalarının Geliştirilmesi ve Açık Veri Beyannamesinin Önemi*” başlıklı iki webinar düzenlenmiştir. Düzenlenen ilk webinarında, Hollanda TU Delft Üniversitesi’nden Yasemin Türkyılmaz-van der Velden, verinin bulunabilir, erişilebilir, birlikte çalışılabilir ve yeniden kullanılabilir olmasının önemini anlatmış, üst veriler, veri arşivleri ve veri yönetim planları hakkında bilgi vermiştir (Creative Commons Türkiye, 2019a). Diğer webinarında ise Kadir Has Üniversitesi’nden Pınar Dağ,

açık veriden, açık verinin öneminden, açık veri konusunda dünyadaki gelişmelerden bahsetmiş ve Türkiye’de yapılması gerekenler hakkında önerilerde bulunmuştur (Creative Commons Türkiye, 2019b).

11 Nisan 2019 tarihinde İYTE ev sahipliğinde *Açık Bilim Yolunda: Bilimi Özgür Bırak* Semineri düzenlenmiştir. Araştırmacılar ve kütüphanecilere hitap eden bu seminerde; açık bilim, açık veri, FAIR ilkeleri, açık lisanslar ve araştırma verisi yönetimi konularında sunumlar yapılmıştır (İYTE, 2019).

2019 yılında Ankara Üniversitesi akademisyenleri üzerine araştırma verilerinin yönetimine ilişkin tutumlarını belirlemek amacıyla bir tez çalışması yürütülmüştür. Bu tez çalışmasında, araştırmacıların verilerini yönetirken kapsamlı bir planlama yapmadıkları ve bu konuya yönelik eğitim ve düzenlemelere gereksinimleri olduğu sonucuna ulaşılmıştır (Zencir, 2019).

YÖK bünyesinde *Açık Bilim ve Açık Erişim Çalışma Grubu* altında alanında uzman kişilerin danışmanlığında, bazı üniversite kütüphanecilerinin yer aldığı *Araştırma Verisi ve Açık Veri Alt Çalışma Grubu* oluşturulmuştur. Bu çalışma grubunun amacı, açık veri ve araştırma verisi konusunda dünyadaki gelişmelerin takip edilmesi, veri yönetim planı hazırlanması, ilgili kurumlarla iş birliği yapılması, eğitimler verilmesi gibi bir dizi uygulamayla Türk üniversitelerinde araştırma verileri yönetiminin geliştirilmesinde katkı sağlamaktır (Gürdal, 2020, 16 Ocak).

Açık veri ve araştırma verileri konularındaki gelişmeleri kurumlarla/araştırmacılarla paylaşmak ve uygulama örnekleriyle bir yol haritası belirlenmesinde kurumlara/araştırmacılara yardımcı olmak amacıyla 25-26 Mayıs 2021 tarihlerinde çevrimiçi *Ulusal Araştırma Verileri Sempozyumu (UAVS)* gerçekleştirilmiştir (Açık Veri, 2022). Aynı yıl Kasım ayında ise, *Açık Veri Yönetimi ve Paylaşımı Çalıştayı* düzenlenmiştir (Hacettepe Üniversitesi Biyoinformatik Anabilim Dalı, t.y.).

Araştırma Verileri Yönetiminde Paydaşların Sorumlulukları

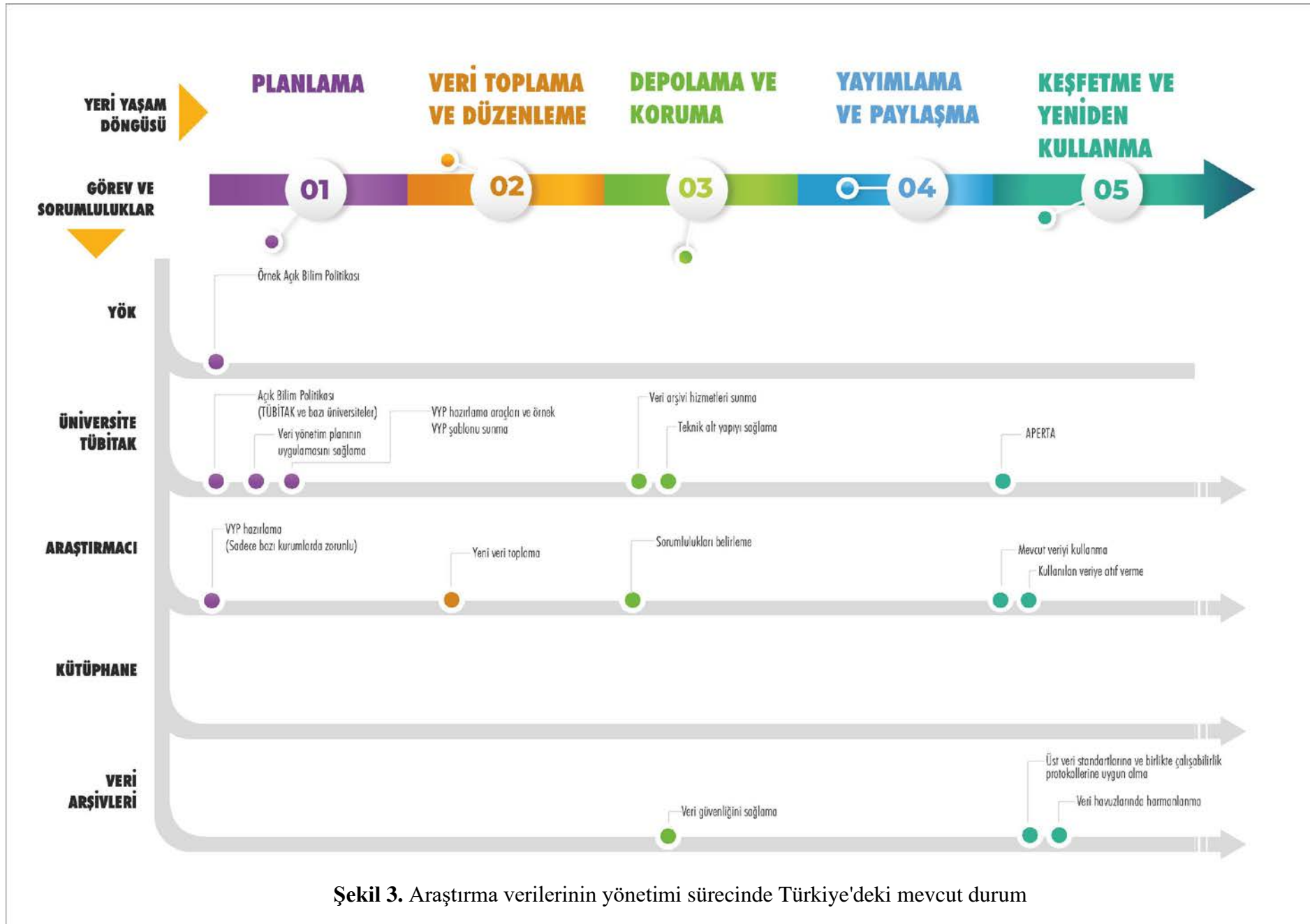
Veri paylaşımının bilimsel açıdan doğrulanabilirlik, güvenilirlik, tekrarlanabilirlik imkânı sunması ve verilerin tekrar kullanımının getirdiği yararlar nedeniyle kurumlar, fon sağlayıcı kuruluşlar ve yayıncılar araştırma verilerinin açık olmasını istemekte ve politika belgelerini buna göre hazırlamaktadır. Verilerin paylaşılabilir, paylaşılan verinin

başkaları tarafından anlaşılabilir ve kullanılabilir olması için araştırma sırasındaki sürecin iyi yönetilmesi gerekir. Bu süreç kısa vadede zaman kaybı gibi görülse de uzun vadede katma değeri yüksektir. Veri paylaşımı bu alanda bir bilgi birikimi ve beceriler bütünü gerektirir. AVY çok kolay olmayan ve araştırmacıların mevcut bilgileriyle yönetmekte zorlanacağı bir süreç olduğundan sorumlu kurum ve kuruluşlar mümkün olduğunca araştırmacıya yardımcı ve destek olmalıdır. Bu bağlamda, fon sağlayıcı kuruluş, üniversite, üniversitenin farklı birimleri (kütüphane gibi), veri arşivleri gibi çeşitli paydaşlara bazı sorumluluklar düşmektedir.

Türkiye özelinde de araştırma verilerinin yönetimi ve paylaşımı konularında ilerleme kaydedilmesi için ilgili kişi ve kurumların üstlenmesi gereken roller ve sorumluluklar bulunmaktadır. Araştırma verilerinin yönetimi konusunda paydaşların üstlendiği rol ve sorumluluklara ilişkin Türkiye’deki mevcut durum Şekil 3 ile özetlenmiştir. Aslında olması gereken ve bu çalışma kapsamında tasarlanan rol ve sorumluluklar ise Şekil 4’te sunulmuştur. Şekiller, Cape Town Üniversitesi’nin “Araştırma verilerinin yönetimi roller ve sorumluluklar” şeması temel alınarak oluşturulmuştur (University of Cape Town, 2022).

Araştırma verilerine yönelik olarak Türkiye’de TÜBİTAK-YÖK iş birliğiyle bir aşama kaydedilmiş durumdadır. Bazı üniversitelerden kütüphaneciler de araştırma grubunda yer alarak bu sürece dâhil olmuştur. TÜBİTAK Açık Bilim Politikası oluşturmuş, YÖK üniversitelere örnek bir açık bilim politikası göndermiştir. Ancak kütüphanelerin bu alandaki rolü konusunda henüz bir açıklık getirilmemiştir. Araştırma verilerinin yönetimi açısından TÜBİTAK Açık Bilim Politikası yeterli değildir. Sayıları çok az olmakla birlikte, İYTE ve Koç Üniversitesi kütüphanesi gibi bazı öncü kütüphaneler araştırma verileri yönetimine ilişkin mensuplarına hizmet geliştirmeye başlamıştır. Ancak Türkiye’deki üniversite kütüphanelerinin büyük çoğunluğu henüz araştırma verileri yönetimine ilişkin bir hizmet sunmamaktadır. Araştırmalarının hemen hemen her aşamasında araştırmacıya destek ve yardımcı olan, araştırmacının bilgi gereksinimini karşılayan, bu amaçla hizmet geliştiren kütüphaneler, bilginin ham maddesi olan ve araştırmaların temelini oluşturan verilerin yönetiminde de araştırmacıya katkıda bulunabilirler. Bilginin toplanması, düzenlenmesi, sınıflanması, depolanması ve erişime sunulması gibi konulardaki mesleki becerileriyle kütüphaneciler, araştırma verileri

yönetiminde arařtırmacıya destek olacak en uygun meslek grubudur. Nitekim çeřitli ülkelerde kütüphanelerin bu alanda önemli roller üstlendiđi görölmektedir.



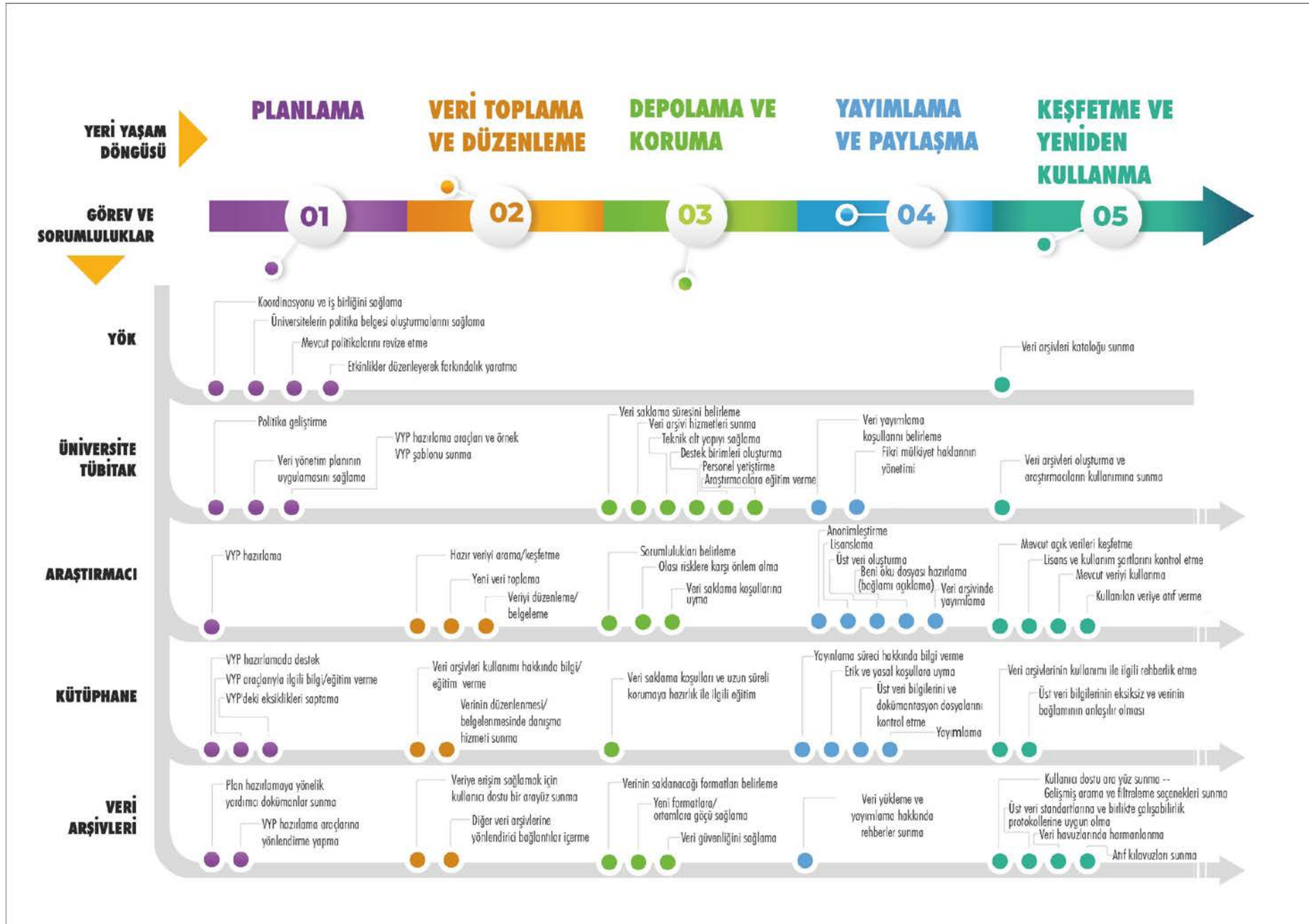
Üst düzey kuruluş olarak YÖK, farkındalığı artırarak üniversite kütüphanelerini ön plana çıkarabilir. TÜBİTAK Ulusal Akademik Ağ ve Bilgi Merkezi (ULAKBİM) ve üniversite kütüphaneleri bu süreçte aktif rol alabilir. Kütüphaneciler yeterli donanıma ulaştıklarında araştırmacılara veri yönetiminin her aşamasında destek olmak için hizmet geliştirebilirler.

Türkiye'deki araştırmacılar üzerinde yapılan çalışmalara göre, araştırmacıların veri yönetim planı hazırlamamaktadır (Doğan ve diğerleri, 2021, s. 57; Ünal ve Kurbanoglu, 2018, s. 299; Zencir, 2019, s. 146). Araştırmacılar verilerini farklı kaynaklardan elde etmekte ve kendi verisini üretmektedir (Ünal ve Kurbanoglu, 2018, s. 295). Verilerin depolanması ve yedeklenmesi aşamalarında sorumlu kişi ve/veya kişiler projeden önce belirlenmektedir (Zencir, 2019, s. 148). Türk araştırmacılar tarafından gerçekleştirilen araştırmalara ait verilerin büyük çoğunluğunun erişilebilir ve yeniden kullanılabilir olmadığı (Ünal ve Kurbanoglu, 2018, s. 306); açık erişim arşivlerin ve kurumsal arşivlerin kullanım oranının oldukça düşük olduğu görülmektedir (Doğan ve diğerleri, 2021, s. 57).

TÜBİTAK Açık Arşivi Aperta hizmete sunulmuştur. Veri güvenliğini sağlayan bu arşiv, birlikte çalışabilirlik protokollerine uygundur ancak araştırmacılara yardımcı dokümanlar ve kullanıcı dostu bir arayüz sunma özelliklerine sahip değildir.

Şekil 4 ile, ülkemizde araştırma verilerinin oluşumundan yayımlanmasına kadar olan süreçte kaliteli, anlaşılır ve tekrar kullanılabilir araştırma verilerini sunmak için paydaşlar tarafından yapılması gerekenlerin bütüncül bir yaklaşımla ele alınması amaçlanmıştır.

Söz konusu süreçte bazı sorumlulukların paylaşılması gerekmekte ve birden fazla paydaş birlikte rol alabilmektedir. Örneğin depolama araştırmacı tarafından yapılabileceği gibi kütüphane, üniversite ve/veya veri sağlayıcı tarafından da yapılabilir. Benzer şekilde birbiriyle iç içe olan bazı aşamalarda (veri toplama ve keşfetme gibi) bazı sorumluluklar her bir alan için tekrar edilmiştir.



Şekil 4. Araştırma verilerinin yönetimde paydaşlara düşen görev ve sorumluluklar

Veri yaşam döngüsü boyunca paydaşların üstlenmesi gereken görevlere ve sorumluluklara YÖK, TÜBİTAK/Üniversite, araştırmacılar, kütüphane ve veri arşivleri çerçevesinden tek tek değinilmiştir. Araştırmacılara yönelik öneriler bu çalışmadan elde edilen bulgular ışığında sonuç kısmında ayrıntısıyla ayrıca ele alınacaktır.

YÖK

Araştırma verilerinin yönetimi ve paylaşımı konusunda üniversiteler arasında iş birliğini ve koordinasyonu sağlamak açısından YÖK'e büyük bir sorumluluk düşmektedir. Yükseköğretim Kanununun Yükseköğretimin amacıyla ilgili 4. Maddesi (c) bendinde;

“Yükseköğretim kurumları olarak yüksek düzeyde bilimsel çalışma ve araştırma yapmak, bilgi ve teknoloji üretmek, bilim verilerini yaymak, ulusal alanda gelişme ve kalkınmaya destek olmak, yurt içi ve yurt dışı kurumlarla iş birliği yapmak suretiyle bilim dünyasının seçkin bir üyesi haline gelmek, evrensel ve çağdaş gelişmeye katkıda bulunmaktır”

ifadesi yer almaktadır. Bilim verilerinin yayılması veri paylaşımıyla mümkündür. YÖK araştırma verilerinin paylaşılmasını politika belgeleriyle destekleyerek tetikleyici bir unsur yaratabilir. Örneğin, Yükseköğretim Kurumları Bilimsel Araştırma Projeleri Hakkında Yönetmeliği'ne proje başvurularında VYP sunulması; doğrulanabilirlik, hesap verebilirlik, şeffaflık gereği araştırma verilerinin yönetim sürecinin nitelikli yapılması ve verilerin mümkün olduğunca paylaşılması maddeleri eklenebilir.

YÖK, üniversitelerde AVY konusunun yaygınlaşmasına öncülük ederek farkındalığın oluşmasını ve/veya artırılmasını sağlayabilir. Bunu Üniversite yönetimiyle iş birliği halinde bilgilendirici toplantılar düzenleme, konunun uzmanı konuşmacılar davet etme, çeşitli etkinlikler gerçekleştirme yollarıyla yapabilir. YÖK, üniversiteleri açık veriye özendirilmeli ve üniversitelerin politika belgelerini oluşturmasını sağlamalıdır.

YÖK'ün talebi üzerine birçok üniversite nasıl kurumsal arşivlerini kurduysa, araştırma verileri için de aynı şey yapılabilir ancak burada sürdürülebilirlik önemlidir. Başlangıç yapıldıktan sonra denetlemelerle, eğitimlerle süreklilik sağlanmalıdır. Çünkü teori ve pratik birbirlerinden farklı olabilmektedir. Başka bir deyişle, yeni bir uygulamaya adapte olma sürecinde karşılaşılan sorunlar farklılaşabilmekte, edinilen bilgi çözüm sunmada yeterli olamayabilmektedir. Ara sıra düzenlenecek toplantılarla, soru-cevap

etkinlikleriyle kütüphanecilere destek olunabilir. Aynı zamanda YÖK, TÜBİTAK iş birliğiyle Türkiye'nin veri arşivlerini içeren bir katalog sunabilir. Veri arşivleri kataloğu, araştırmacılara hem verilerini yayımlama hem de yayımlanmış verileri tekrar kullanma aşamasında veri arşivlerinin belirlenmesinde yardımcı olacaktır.

TÜBİTAK/Üniversite

Açık araştırma verileri ile ilgili olarak genel anlamda üniversitelerin kat edeceği çok yol vardır.

TÜBİTAK halihazırda Açık Bilim Politikası'nı oluşturmuştur ancak bu politika belgesi araştırma verilerinin yönetimini ve paylaşımını kapsayacak şekilde revize edilebilir. Politika belgesi olmayan üniversitelerde ise, herhangi bir politikanın bir parçası olmayan doğrudan araştırma verileri ile ilgili (Araştırma Verilerinin Yönetimi ve Paylaşım Politikası gibi) politika belgesi oluşturulabilir Politika belgesinde üniversitenin ve mensuplarının sorumluluklarının yanı sıra, veri sahipliği, etik ve yasal konular, veri depolama, yayımlama, paylaşım, FAIR ilkeleri gibi konular yer almalıdır. Araştırma verilerinin mümkün olduğunca açık ve erişilebilir olması için araştırmacı teşvik edilmelidir.

Üniversite ve/veya TÜBİTAK'ın üstlenmesi gereken görevler veri yaşam döngüsü aşamalarına göre şu şekilde olmalıdır:

- Planlama: Araştırmacının, araştırma öncesinde bir veri yönetim planına sahip olma zorunluluğu politika belgesinde belirtilmeli ve bu ilkenin uygulanıp uygulanmadığı kontrol edilmelidir. TÜBİTAK bünyesinde böyle bir zorunluluk olup kontrol sağlama üniversitelere göre görece daha kolaydır. Proje önerisi aşamasında VYP kontrol sağlanabilir. Ancak üniversiteler için proje haricindeki araştırma çıktılarında VYP mevcudiyet kontrolünün yapılması zordur. Bu aşamada dergi politikaları da önem kazanmaktadır. TRDizinde indekslenen ve/veya DergiPark üzerinden yayımlanan dergilerin politikalarında, yayın yapma ön şartı olarak VYP hazırlama ve araştırma verilerinin paylaşılması zorunluluğu getirilebilir. Dergilerin bunu yapmaları içinse, TRDizin'de indekslenme koşulları ve/veya DergiPark'ta dergi yayımlama koşulları arasına dergi politikalarının bu

şekilde geliştirmesi şartı eklenebilir. Aynı zamanda dergi yayımlayan üniversiteler, dergi politikalarını bu doğrultuda güncelleyebilir. Fon sağlayıcı kuruluş, yayıncılar ve üniversiteler gibi farklı yerlerden araştırmacının VYP hazırlama zorunluluğunun olması zaman içinde araştırmacıların bütün çalışmalarına uyguladıkları bir alışkanlığa dönüşecektir. Kütüphanelerin sundukları akademik yazı desteği hizmetlerinde başvuru esasları arasına çalışma yayına kabul edildiğinde verilerle birlikte kurumsal arşivde paylaşımına sunulması ilkesi eklenebilir.

Politika belgesinde gerekli düzenlemeler yapıldıktan sonra, ilgili kurum plan hazırlama aşamasında araştırmacıya her türlü desteği vermek için donanımlı personel istihdam etmelidir. VYP hazırlama araçları, örnek bir veri yönetim planı şablonu ve kontrol listesi sunulmalıdır.

- Depolama ve Koruma: Saklama süresi ve imha gibi konular politika belgesinde yer almalıdır. TÜBİTAK, APERTA aracılığıyla Türkiye Açık Arşivi'nin alt yapısını oluşturmuş ve kullanıma sunmuştur. TÜBİTAK'tan farklı olarak Üniversite aynı zamanda araştırma sırasında verilerin saklanması koşullarıyla ilgili de yatırımlar uygulamalıdır. Üniversiteler uygun teknik alt yapıyı sağlayarak depolama ortamları, yedekleme hizmeti ve uzun süreli korumanın yapılacağı ücretsiz veri arşivleri sunmalı veya yer göstermelidir. Bunun için bilgi işlem/bilişim teknolojileri, kütüphane gibi birimlerden destek alınmalı, mevcut personel donanımlı hale getirilmeli ve/veya yeni personel istihdamı sağlanmalıdır. Araştırmacılara bu konuda eğitim vermek kütüphane ve bilgi işlem/bilişim teknolojileri iş birliğiyle üniversitenin sorumluluğundadır.
- Yayımlama ve Paylaşma: Verilerin açık erişilebilir olması, kısıtlı erişim olduğu durumlarda gerekçe bildirilmesi, verilerin yeniden kullanımını sağlamak için iyi dokümante edilmesi, lisanslama yapılması, DOI alınması ve fikri mülkiyet haklarının yönetimi konuları politika belgesinde yer almalıdır. Araştırma verilerinin yayımlanması durumunda, araştırmacıya teşvik sunulabilir. Üniversitenin performansa dayalı değerlendirme sisteminde araştırma verilerinin paylaşımı da bir ölçüt olarak eklenebilir.

- Keşfetme ve Yeniden Kullanma: TÜBİTAK Aperta aracılığıyla üniversiteler ise kütüphaneleri aracılığıyla veri arşivleri sunarak ve/veya yer göstererek bu sürece katkı sağlayabilirler.

Politika belgesi tek başına kesinlikle yeterli değildir. Buna paralel olarak Üniversitenin gerekli insan gücünü sağlama, mevcut personeli eğitime, farkındalık kazandırmaya yönelik düzenlenecek etkinlikler, teknik alt yapı ve bakım hizmetleri için bütçe ayırması gerekmektedir. İlgili kurumlar ve birimlerle iş birliği çerçevesinde bir yapılanmaya gitmelidir. Üniversitenin proje ofisi, kütüphane, bilgi işlem/bilişim teknolojileri gibi birimlerinde çalışanlardan oluşan araştırma verileri yönetiminden sorumlu ayrı bir birim kurulabilir, yeni destek personeli (veri kütüphanecisi, veri uzmanı gibi) istihdam edilebilir. Bilgi ve Belge Yönetimi bölümü olan üniversitelerde araştırma verileri yönetimi, veri kütüphaneciliği, veri okuryazarlığı gibi dersler müfredata eklenebilir. Üniversite kütüphanelerinde halihazırda görevli kütüphanecilerin de gerekli mesleki eğitimlerle yenilikleri takip etmeleri sağlanabilir. Konunun uzmanları tarafından verilecek bu eğitimler mesleki dernekler veya YÖK vasıtasıyla bölgesel olarak düzenlenebilir. Personel gerekli donanıma ve beceriye sahip olduktan sonra diğer araştırmacılara da destek verebilir.

Üniversiteler kütüphaneleri aracılığıyla, TÜBİTAK ise ULAKBİM ve/veya diğer kaynaklarını kullanarak araştırmacıların veri yönetimi ile ilgili eğitim ihtiyacını giderebilir. Web sayfalarında sunulan hizmetlerin tanıtımı, destekleyici rehberler, veri yönetiminde kolaylık sağlayıcı araçlara yönlendirmeler, politika belgesine erişim, dış bağlantılara erişim (TÜBİTAK gibi) sağlanabilir. Araştırma verileri yönetiminin tüm aşamalarını kapsayan bir eğitim portalı TÜBİTAK tarafından oluşturulmuş olup geliştirilmeye de devam etmektedir. Bu portalın interaktif olması, disipline özgü seçim imkânı sunarak ilerlenmesi, destekleyici videolar içermesi, bölüm sonlarında konuya özel küçük testler içermesi sağlanabilir.

Araştırmacı

Araştırma fikriyle birlikte araştırmacıyı bağlayıcı bir veya birden fazla politika belgesi bulunmaktadır. Bunlardan ilki araştırmacının mensup olduğu kuruma, bir diğeri ise araştırma için bir destek alınacaksa fon sağlayıcı kuruluşa ait politika belgesidir.

Araştırmacı politika belgesindeki zorunlu olan ve önerilen hususları mümkün olduğunca gerçekleştirmekle yükümlüdür. Araştırmacıların yükümlülüklerini yerine getirebilmesi için araştırma verileri yönetimi sürecinin tüm aşamalarını (planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) gerçekleştirilebilmelidir. Bu sürecin iyi yönetebilmesi için ilgili kurumlar araştırmacının eğitim ihtiyacını gidermelidir.

Kütüphane

Kütüphaneler, araştırma verileri yönetiminin her aşamasında araştırmacılara yardımcı olmalıdır. Anket, görüşme gibi yollarla mevcut durumu saptayarak gereksinim analizi yapabilir, ihtiyaca göre genel eğitimler düzenleyebilir. Disipline özgü farklılıklar gözetilerek fakülte kütüphanecileri konularında uzmanlaşabilir. Bu eğitimin yanında araştırmacının isteği doğrultusunda eğitim gereksiniminin karşılanacağı veri yaşam döngüsü aşamaları şu şekilde olmalıdır:

- Planlama: Veri kütüphanecisi, araştırmacıya planlama araçları hakkında bilgi vermeli, VYP hazırlamasında yardımcı olmalı ve kontrol listeleri yardımıyla eksiklikleri tespit ederek gerekli düzenlemelerin yapılmasını sağlamalıdır.
- Veri Toplama ve Düzenleme: Mevcut verilerin keşfi için veri arşivleri hakkında bilgi sunan eğitimler verebilir. Yeni veri toplanacaksa veri sözlüklerinin kullanılması, tutarlı dosya adlandırılması yapılması, beni oku dosyası kullanılması, verinin standart üst verilerle tanımlanması gibi işlemler anlatılarak kaliteli dokümantasyon desteği sunulmalıdır.
- Depolama ve Koruma: Araştırma sırasında yedeklemenin ne sıklıkta yapılacağı, verilerin hangi ortamlarda depolanacağı, veri güvenliğinin nasıl sağlanacağı, uzun süreli korumada nasıl bir yol izleneceğine dair eğitim verilmelidir.
- Yayımlama ve Paylaşma: Verinin tekrar kullanılabilirliğini ve görünürlüğünü sağlamak için yayımlanması aşamasında lisanslamanın nasıl yapılacağı, hangi ortamlarda erişime sunulabileceği konuları hakkında bilgi verilmelidir. Kurumsal arşivde sunulacak verilerin yayımlanması aşamasında, veri sahipliği, veri mahremiyeti gibi etik ve yasal koşullar dikkate alınmalı, üst veri bilgileri ve dokümantasyon dosyaları kontrol edilmeli, kalıcı tanımlayıcı (DOI (Digital

Object Identifiers) gibi) desteği sunulmalıdır. Veriler, ilgili araştırma çıktılarıyla bağlantılı olarak erişime sunulmalıdır.

- Keşfetme ve Yeniden Kullanma: Veri paylaşılması durumunda, birlikte çalışabilirlik protokolleri ile uyumlu sistemler sunarak, standart üst veri kullanımı ve bağlamın anlaşılır olması sağlanarak keşfedilebilirlik artırılmalıdır. Veri kullanımı durumunda ise veri toplama aşamasında olduğu gibi veri arşivlerinin kullanılması sağlanmalıdır.

Araştırmacılara eğitim verilmesinin yanı sıra, özellikle kariyerlerinin başındaki lisansüstü öğrencilerin araştırma verileri yönetimi ile ilgili ders almaları sağlanabilir. Nasıl ki araştırma yönetimi dersi lisansüstü öğrencileri için zorunlu bir ders ise, araştırma verileri yönetimi de başlangıçta bu dersin bir parçası sonrasında ise zorunlu bir ders olabilir. En azından bir araştırma projesi veya tez yazacak, yeterli araştırma deneyimi olmayan öğrenciler için bu ders faydalı olacaktır. Kütüphaneler web sitelerinde, araştırma verilerinin tanımı ve araştırma verilerinin yönetiminin önemine yönelik kaynaklar sunabilir ve/veya diğer sitelere yönlendirme yapabilirler.

Veri Arşivleri

Araştırmacı üniversitenin veya fon sağlayıcı kuruluşun sağladığı imkanlar doğrultusunda uygun arşive verilerini yüklemelidir. Bu arşivler, kurum/kuruluşların veri arşivleri olabileceği gibi konu arşivleri veya genel kullanımlı arşivler de olabilir. Veri arşivleri, hazırlayacağı rehberler, Sıkça Sorulan Sorular veya Yardım sayfalarıyla araştırmacılara bilgi ve hizmet sunabilirler. Veri arşivlerinin sorumlulukları şu şekilde olmalıdır:

- Planlama: Veri arşivlerinde, veri yönetim planı hazırlamaya yönelik yardımcı dokümanlara yer vermeli, şablon bir plan sunmalı ve/veya plan hazırlama araçlarına (DMPTool gibi) yönlendirme yapmalıdır.
- Veri Toplama ve Düzenleme: Veriye erişim sağlamak için kullanıcı dostu bir ara yüz sunmalı, diğer veri arşivlerine ve/veya kataloglarına yönlendirici bağlantılar içermelidir.

- Depolama ve Koruma: Verinin saklanacağı uygun formatları araştırmacıya bildirmeli, teknoloji eskimesi halinde veriyi yeni formatlara ve ortamlara taşımali, veri güvenliğini sağlamalıdır.
- Yayımlama ve Paylaşma: Veri yükleme ve yayımlama süreci hakkında rehberler sunmalı, sistemin nasıl kullanılacağını açıklamalıdır.
- Keşfetme ve Yeniden Kullanma: İhtiyaç duyulan veri/veri setinin bulunabilmesi için kullanıcı dostu bir ara yüz, gelişmiş arama seçenekleri ve çeşitli uçlardan (konu, kurum, coğrafi alan gibi) filtreleme imkânı sunmalıdır. Birlikte çalışabilirlik protokolleriyle uyumlu olmalı, veri havuzlarında harmanlanmalı ve diğer veri arşivlerine yönlendirici bağlantılar içermelidir. Ayrıca kullanıcılara atıf kılavuzları sunmalıdır.

Genel olarak bu süreçte yapılması gerekenler şöyle özetlenebilir. Planlama aşamasında verinin paylaşılıp paylaşılmayacağına karar verilmelidir. Paylaşım için herhangi bir engel yoksa araştırma öncesinde gerekli planlamalar yapılmalıdır. Bu planlama araştırma sırasında ve sonrasında yol gösterici olmalıdır. Veriler uzun süreli korumaya adım adım hazırlanmalıdır. Standart üst veriler kullanılarak tanımlanmış, iyi yapılandırılmış ve belgelenmiş veriler paylaşılmadan denetlenmelidir. Veri paylaşım ortamına göre, bu akran değerlendirilmesi olabileceği gibi yayıncı-hakem değerlendirmesi veya veri kütüphanecisi tarafından araştırma verilerinin kalitesinin kontrol edilmesi olabilir. Gerekli kontroller yapıldıktan sonra araştırma verileri lisanslanarak veri arşivlerinde erişime sunulmalıdır. Veri güvenliği için daha sağlıklı bir ortam sunan veri arşivleri, aynı zamanda verinin erişilebilir, keşfedilebilir ve tekrar kullanılabilir olmasını sağlar.

Araştırma verileri yönetimi konusunda ülkemizde bazı gelişmeler olmaktadır ancak üniversiteler düzeyinde henüz çok aşama kaydedilememiştir. Bu nedenle üniversite ayağında yapabilecekler konusundaki önerilere ağırlık verilmiştir. Kamu kaynaklarıyla üretilen verilerin kamu yararı için tekrar kullanıma sunulması ancak iyi yönetilmiş araştırma verileri ile mümkündür. Bu süreçte araştırmacıya destek olunmalıdır.

3. BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ SÜRECİNİN AŞAMALARI

Bu bölümde, araştırma verileri yönetimi planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma, keşfetme ve yeniden kullanma olmak üzere beş başlık altında incelenmiş, her başlık altında konuya ilişkin araştırmacılara yol gösterecek bilgiler sunulmuştur.

3.1 PLANLAMA

Planlama, araştırma verileri yönetiminin ilk aşaması olup proje başlamadan veri yaşam döngüsünde yer alan konulara ilişkin (veri toplama, saklama, uzun süreli koruma, etik konular gibi) gelecekte ne yapılacağına planlanması sürecidir. Fon sağlayıcıların zorunlu tuttuğu veri yönetim planı ise araştırma projesi yürütülürken veya tamamlandığında, elde edilen araştırma verilerinin nasıl yönetileceğine dair bilgiler içeren bir belgedir (OpenAIRE Ufuk 2020, t.y.). İyi bir veri yönetiminin ve araştırma verisinin bulunabilir, erişilebilir, karşılıklı işler ve yeniden kullanılabilir olmasının parçasıdır (European Commission, 2016, s. 7). Fon sağlayıcıların planı gerekli kılması, uzun vadede para tasarrufu sağlaması, araştırmanın bütünlüğünü desteklemesi, araştırmaya değer katması ve daha kaliteli araştırma için öncülük etmesi nedeniyle gereklidir (Donnelly, 2012, s. 85-86; Zhang, 2016).

Araştırmacı için yazılı bir veri yönetim planı; verinin nasıl yönetileceğine ilişkin sorunların çözümlenmesinde ve iş birliği yapılan araştırmacılarla aktif veri paylaşımında kullanışlı bir araçtır. Ortak çalışma yapan araştırmacıların; isimlendirme, düzenleme, depolama gibi hususlarda aynı düzeni kullanması birbirlerinin verilerinin kullanımında kolaylık sağlamaktadır. Veri yönetim planı, araştırmadan ayrılan bir kişinin verilerinin bulunup kullanılması veya araştırmaya sonradan dâhil olan birinin mevcut düzeni anlaması için yol gösterici ve zaman kazandırıcı bir kaynaktır (Briney, 2015).

Veri yaşam döngüsündeki aşamaların altında, araştırmada kullanılacak verilere dair neyin nasıl yapılacağına ilişkin sorulara yanıt veren veri yönetim planında yanıtlanması gereken sorular şu şekildedir (Zhang, 2016):

- Toplanacak/işlenecek veri nedir?
 - Verinin büyüklüğü nedir?
 - Kullanılacak dosya formatları nelerdir?
 - Veri kaynağı nedir?
 - Verinin sahibi kimdir?
- Veriler nasıl belgelenecek ve düzenlenecek?
 - Veri nasıl belgelenecek?
 - Hangi üst veri standartları kullanılacak?
 - Dizin yapısı nasıl olacak?
 - Dosya adlandırma kuralları kullanılacak mı?
- Depolama ve veri güvenliği nasıl sağlanacak?
 - Veri nerede depolanacak?
 - Hangi dosyalar, nerede, ne sıklıkla yedeklenecek?
 - Depolamada ihtiyaç duyulan bütçe ve kaynaklar nelerdir?
 - Hassas veriler için uygulanacak politikalar nelerdir?
 - Hangi güvenlik önlemleri alınacak?
- Araştırma bittikten sonra veri nasıl yönetilecek?
 - Proje bittikten sonra veriler ne kadar süreyle saklanacak?
 - Uzun süreli koruma için veri nasıl hazırlanacak?
 - Koruma için ihtiyaç duyulan bütçe kaynaklar nelerdir?
- Veriler tekrar kullanılabilir hale nasıl getirilecek?
 - Hangi veriler paylaşılacak?
 - Veriler ne zaman paylaşılacak?
 - Veriler nerede paylaşılacak?
 - Paylaşım için veri nasıl hazırlanacak?
 - Paylaşım için ihtiyaç duyulan bütçe kaynaklar nelerdir?

Araştırmacılara veri yönetim planı şablonu sağlayan veri yönetim araçları bulunmaktadır. Kaliforniya Üniversitesi Kürasyon Merkezi tarafından DMPTool, Dijital Kürasyon Merkezi tarafından DMPOnline araçları geliştirilerek araştırmacıların hizmetine sunulmuştur. Her iki araç da fon sağlayıcıların politika belgelerinde istenen gereksinimlere dayanan birçok şablon sunmakta, araştırmacıların veri yönetimi planları oluşturmasına, gözden geçirmesine ve paylaşmasına yardımcı olmaktadır (DMPOnline, 2020; DMPTool, 2020).

3.2 VERİ TOPLAMA VE DÜZENLEME

Araştırma öncesinde, mevcut olan bir veri mi kullanılacağı yoksa verinin yeni mi oluşturulacağı tasarlandıktan sonra veri toplama ve düzenleme süreci başlamaktadır.

Veri toplanmadan önce, verilerin daha önceki çalışmalarda üretilme ihtimaline karşın veri arşivleri taranmalıdır (Doğan, 2019). Veri arşivleri ile ilgiliye [Bölüm 3.5](#)'ten erişilebilir.

Halihazırda mevcut veri yoksa, veri toplamak için kullanılacak yöntem (deney, gözlem, görüşme, anket, ölçüm vb.) belirlenmeli ve belirli bir düzen içerisinde çalışılmalıdır (Doğan, 2019).

3.2.1 Düzenleme

Araştırma verilerinin yönetiminde bir diğer önemli husus da dosyaların/klasörlerin düzeni ve isimlendirilmesidir. Çalışmanın veya projenin başında mantıklı ve tutarlı bir isimlendirme ile dosya düzeni oluşturmak, dosyayı açmadan içeriğin anlaşılmasını sağlayarak hem zaman kazandırmakta hem de aranan dosyanın bulunmasını kolaylaştırmaktadır (University of Cambridge, 2020b). İyi bir adlandırma stratejisi, özellikle iş birliği yapılan çalışmalarda karışıklığı önlemektedir.

Veri dosyalarının adlandırılmasında; gelecekte kolay erişim için **organizasyona**, içeriğe özgü bilgi sunması nedeniyle **bağlama** ve belirlenen adlandırma kurallarının her dosya için uygulanmasını içeren **tutarlılığa** dikkat etmek önemlidir (Mantra Research Data Management Training, 2020b). Çünkü iyi bir dosya adlandırma; dosyanın içeriğine, konumuna ve sürümüne ilişkin yararlı ipuçları sağlamakta; dosyaların sınıflandırılması ve sıralanmasına yardımcı olmaktadır (UK Data Service, 2020b). Bu bölümde, klasörlerin düzenini içeren dizin yapısı, dosya adlandırılmasında dikkat edilecek hususlar, sürüm bilgisi kullanmanın yararları ve bunlara yardımcı araçlardan bahsedilecektir.

Dizin yapısı: Klasörler araştırma faaliyeti, veri veya içerik türüne göre belirli bir hiyerarşi içinde yapılandırılmalı, genel konular üst düzeyde olmak üzere gittikçe özele inen bir yapı kullanılmalıdır (The Graduate Institute, 2020). Önceden oluşturulmuş bir sınıflandırma şeması dizin yapısının oluşturulmasında yardımcı olacaktır. UK Data Archive tarafından proje dosyalarının yapılandırılmasında önerilen bir örneğe göre, en üst klasörde projeyi

nitelendirecek şekilde proje adı, proje kodu veya proje adının anlamlı bir kısaltılması kullanılabilir. Bir alt düzey, veriler ve dokümantasyon olarak ayrılabilir. Bunlar da kendi içinde veri türüne (metin, resim, ses, veri tabanı gibi) göre ayrılarak bir alt kademe oluşturulabilir. En alt klasör, veri toplama aracına göre (deney, anket, görüşmeler, odak grubu) bölümlenebilir. Ayrıca klasörlerin üç veya dört düzeyde bölümlendirilmesi ve bir klasörde ondan fazla öge bulunmaması önerilmektedir (UK Data Service, 2020b).

Dosya adlandırma: Araştırma sırasında zaman darlığından dolayı kullanılan veya indirilen dosyalar rastgele isimlerle kaydedilmekte, daha sonra düzeltilme ihtimaliyle öylece kalmaktadır. Geçen zamanda dosyalara erişim sağlayabilmek için anlamlı ve/veya tutarlı bir adlandırma kuralı benimsenmesi gerekmektedir.

Dosya adlandırılırken; içeriği tanımlayan ve tarih, proje adı/kodu, veri türü, koşullar, konum/mekânsal koordinatlar, araştırmacı ve sürüm bilgileri gibi öğeleri içeren bir adlandırma kuralı oluşturulmalıdır (University of Minnesota Libraries, 2020a). Yeni bir dosya adı oluşturulurken bu kurala uyulmalı ve tutarlı olunmalıdır. Adlandırma kuralları, kullanılan kısaltmaları veya kodları açıklayan bir benioku.txt dosyası oluşturulabilir (Stanford Libraries, t.y.). Dosya adlandırılırken ayrıca dikkat edilmesi gereken hususlar şu şekildedir (Mantra Research Data Management Training, 2020b; Stanford Libraries, t.y.) :

- İçerik hakkında bilgi verici anlamlı ama kısa adlar kullanılmalıdır
- Boşluk, nokta, Türkçe karakterler veya özel karakterler (@ # \$ % ^ & * () ` ; < > ? , [] {} gibi) kullanılmamalıdır
- Bir dosya adındaki öğeleri ayırmak için kısa çizgi (-), alt çizgi (_) veya büyük-küçük harf kullanılmalıdır
- Tarih atamaları için tek bir format tercih edilmeli, dosyaların kronolojik sıralanabilmesi için YYYYAAAGG veya YYAAGG formatı kullanılmalıdır
- Sürüm bilgisi kullanılmalı, sıralama yapabilmek için en az iki basamaklı sürüm numarası olmalıdır (1, 2 yerine 01,02 veya 001,002 gibi).

Dosya adlandırılırken oluşturma tarihi, proje numarası veya adının kısaltması, veri türü, yer bilgisi, dosyayı oluşturanların baş harfleri ve/veya sürüm bilgisi gibi öğeler kullanılarak tutarlı adlandırma yapılabilir (ELIXIR, 2021). Söz konusu öğeleri içeren bir adlandırmada örneği aşağıda sunulmuştur ELIXIR (2021):

- Bal arısı projesi (Honeybee), Helsinki'de yapılan deney 2, 2 Aralık 2020'de oluşturulan veri dosyası için
 - Dosya adı: 20201202_HB_EXP2_HEL_DATA_V03.xls
 - Açıklama:

Tarih_ProjeAdıKısaltması_DeneyNumarası_YerBilgisi_VeriTürü_SürümNo

Bazen farklı araştırmalar için aynı dosya adının kullanıldığı bilinmektedir. Bu tür genel isimlendirmeler verinin içeriği hakkında ipucu vermemektedir. Yetersiz olan dosya adlandırma örnekleri şu şekildedir:

- Manuscript_SON.doc
- Veriler.sav
- Grafik_ikon.xls

Sürüm bilgisi: Dosyaların farklı sürümlerini saklamak birçok açıdan faydalıdır. Hem çalışmanın ilerleyişini izleme ve önceki sürüme dönmede yardımcı hem de ortak veya birden fazla yerde çalışma yaparken kullanışlıdır (Briney, 2015, s. 77). Örneğin, yapılan hatalarda çalışmayı baştan yapmak yerine bir önceki sürüme dönme şansı sunarak veri ve zaman kaybı risklerini azaltır.

Yeni bir sürüm oluşturulurken dosyada yapılan değişiklikleri kaydetmek genellikle yararlıdır. İçeriğe ait değişiklikler beni oku dosyalarına kaydedilebildiği gibi özellikle iş birliği yapılan çalışmalarda hangi değişikliğin kim tarafından ne zaman yapıldığı bilgilerini içeren sürüm kontrol tabloları da kullanılabilir (The Graduate Institute, 2020).

Yardımcı Araçlar

Dosya adlandırma ve sürüm belirleme amacıyla araştırmacılara yardımcı bazı araçlar bulunmaktadır. Proje başlangıcında tutarlı bir dosya adlandırılması yapılmamışsa, tek tek dosya adı değiştirmek yerine bu işlemlerin toplu olarak yapılabilmesini sağlayan Renamer, Zamzar, Bulk Rename Utility, PSRenamer gibi bazı dosya adlandırma araçları mevcuttur (Stanford Libraries, t.y.). Bu uygulamaların; dosya adının önüne- sonuna kelime ekleme, büyük-küçük harf değiştirme, karakter silme-ekleme, dosya türü, dosya boyutu ekleme, Türkçe karakterleri İngilizce karakterlere dönüştürme, sıralama için standart formatta tarih ekleme, sayısal karakterler ekleme gibi birçok özelliği bulunmakta

ve bu uygulamalar sayesinde aynı anda birçok dosya için isim değişikliği yapabilmektedir (ReNamer, t.y.; Bulk Rename Utility, 2021).

Bazaar, Git, Mercurial gibi sürüm belirleme araçları ise, hangi tarihte hangi değişikliğin yapıldığının görülmesine, ortak çalışılan araştırmacının yaptığı düzeltmelerin izlenmesine ve iki ayrı kopyadaki tüm değişikliklerin bulunarak birleştirilmesine olanak sağlamaktadır.

3.2.2 Veri Dokümantasyonu

Veri yönetiminin en önemli öğelerinden biridir. Verinin nasıl oluşturulduğu, verinin bağlamının ne olduğu, veri dosyalarının düzeni, yapısı ve içeriği, herhangi bir manipülasyonun olup olmadığı, veri gizliliği, erişim ve kullanım koşulları gibi açıklayıcı bilgiler sunmaktadır (Briney, 2015, s. 44; Van den Eynden ve diğerleri, s. 9; The University of Arizona, 2020). Yeniden kullanımda, veriye yönelik bağlamsal anlam kritik öneme sahiptir. Verinin bağlamından ayrılması ve/veya bağlamın kaybı, yeniden kullanımı zorlaştırmakta veya imkânsız hale getirmektedir (Schneider, 2013). Diğer araştırmacılar bulguları anlamak ve doğrulamak, benzer bir çalışma tasarlamak, verileri analiz etmek, verileri kendi verileriyle karşılaştırmak veya yeniden kullanmak için bu tür verilere gereksinim duyabilmektedir (Mantra Research Data Management Training, 2020a). Gelecekte verinin hem oluşturan araştırmacı hem de diğer araştırmacılar tarafından anlaşılması ve yorumlanması iyi bir veri dokümantasyonu ile mümkündür. Dokümantasyonun açık ve ayrıntılı olması verinin kaliteli olmasını sağlamaktadır. Dokümantasyon, hem yapılan çalışma hem de veri için iki ayrı düzeyde oluşturulabilir (CESSDA Training Team, 2017 – 2020; Oregon State University, 2020):

- Proje düzeyinde dokümantasyon: Çalışmanın amacı, araştırma soruları, hipotezler, kullanılan yöntemler, verilerin ne olduğu, kim tarafından, nerede, nasıl, neden toplandığına ilişkin bilgiler içermektedir.
- Veri düzeyinde dokümantasyon: Veri ve veri setinin ayrıntılı bir biçimde tanımlanmasıdır. Veri türü (deneysel, gözlemsel, ham veya türetilmiş, vb.), kullanılan araçlar, veri toplama ayrıntıları, dosya türleri (.csv, .xlsx, .tiff vb.), veri

işleme yöntemleri, kullanılan yazılım, veri işleme komut dosyaları veya kodları, değişkenler, her değişkenin açıklaması gibi bilgiler içermektedir.

Kullanılan araştırma verisi çeşitli formatlarda (resim dosyası, veri tabanı, metin dosyası gibi) olabilmektedir. Kullanılan dosyaların içerikleri farklılık gösterdiğinden her bir dosya için ayrı dokümantasyon yapmak gerekecektir. Bazı dosya türleri (Access, Excel, SPSS gibi) veri hakkındaki bilgilerin aynı dosyada tutulmasına olanak sağlamaktadır. Bilgilerin aynı dosyaya kaydedilmesine elverişli olmayan dosyalar (ses dosyası, görüntü dosyası veya video gibi) için ayrı bir veri dosyası kullanmak gerekmektedir (CESSDA Training Team, 2017 – 2020).

3.2.3 Üst Veri ve Üst Veri Standartları

Veri yönetimi bağlamında üst veri (metadata) verilerin kim tarafından, nasıl, ne zaman, nerede toplandığı gibi veri hakkında bilgileri içeren standartlaştırılmış ve yapılandırılmış bir dokümantasyon biçimidir (UK Data Archive, 2011). Hem üst veri hem de dokümantasyon veri hakkında bilgi sunmaktadır ancak üst veri; makinece okunabilir, işlenebilir ve birlikte çalışabilir olma özellikleriyle dokümantasyondan ayrılmaktadır (Mantra Research Data Management Training, 2020a).

Veriler, doğru bir şekilde tanımlandığı ve belgelendiği sürece hem oluşturan kişi hem de diğer araştırmacılar tarafından çalışmanın önemli ayrıntılarının anlaşılması mümkün olur. Bununla birlikte, veri arşivlerinde depolanan verinin bulunabilir ve erişilebilir olması üst veri ile mümkündür (Cornell University Research Data Management Service Group, t.y.).

Araştırma verilerinin etkin bir şekilde düzenlenmesini sağlaması, keşfi kolaylaştırması, veri paylaşımını kolaylaştırması, veriler için dijital tanımlayıcılar sağlaması ve arşivlemeyi ve korumayı desteklemesi nedeniyle üst veri önemlidir (UWA Library, 2022).

NISO (National Information Standards Organization-Ulusal Bilgi Standartları Enstitüsü) üst verileri Tablo 4'te görüleceği üzere, dört ana kategori altında sınıflandırmıştır (Riley, 2017). Verinin tanımlanmasında kullanılan tanımlayıcı üst veri, verinin anlaşılabilmesi ve keşfedilebilmesi için kullanılmaktadır. Teknik, koruma ve haklar olmak üzere üç alt üst veri türünü içeren yönetimsel üst veri, verinin kullanımı için gerekli bilgiyi

sağlamaktadır. Yapısal üst veri kullanımıyla, verilerin diğer verilerle ve/veya kaynaklarla ilişkisini ortaya koymak amaçlanmaktadır. Son kategoride bulunan işaretleme dili ise, içerikteki diğer yapısal veya anlamsal özellikleri kullanarak üst verileri ve işaretleri entegre etmektedir.

Tablo 4. Üst veri türleri (Riley, 2017)

| Üst veri türleri | Kullanım amaçları |
|-------------------------|--|
| Tanımlayıcı üst veri | Bir kaynağı bulmak ve anlamak |
| Yönetimsel üst veri | |
| -Teknik üst veri | -Dosyaların kodunu çözmek ve işlemek |
| -Koruma üst verisi | -Verilerin uzun süreli yönetimini sağlamak |
| -Haklar üst verisi | -İçeriğe eklenen fikri mülkiyet haklarına dair bilgileri sağlamak |
| Yapısal üst veri | Verilerin diğer verilerle ve/veya kaynaklarla ilişkisini ortaya koymak |
| İşaretleme dili | İçerikteki diğer yapısal veya anlamsal özellikleri kullanımıyla üst verileri ve işaretleri entegre etmek |

Standartlar hayatın her alanında önemlidir. Özellikle farklı coğrafyalarda ortak yürütülen çalışmalarda ortak bir dilin olmasını sağlamaktadır. Bilimsel çalışmalarda standardın önemi, 1999 yılında kaybolan “Mars Climate Orbiter (Mars İklim Uydusu)” örneğiyle açıklanabilir. Mars'ın atmosferi, iklimi ve yüzeyi hakkında eşzamanlı araştırmalar yapmak üzere tasarlanan uydu aracının Mars yörüngesine yerleştirilmesi planlanmıştır. Yörünge aracının yazılımında metrik standart kullanılması ve Dünya'dan gelen komutların metrik standarda (Newton-saniye) dönüştürülmeden İngiliz ölçü biriminde gönderilmesi bir navigasyon hatasına neden olmuştur. Yörüngeye giremeyen “Mars Orbiter” atmosfere düşerek parçalanmış ve milyon dolarlık araştırma başarısızlığa uğramıştır (NASA, 2019).

Veri paylaşımının her aşamasında da önemli olan standartlar, alt yapı, yazılım gibi unsurların birlikte çalışabilir olması için ortam yaratmakta ve böylelikle veri alışverişini sağlamaktadır (Digital Archiving Consultancy, Bioinformatics, 2005, s. 8). Örneğin, YÖK Ulusal Tez Merkezindeki yayınlar, kayıtların standartlara ve birlikte çalışabilirlik protokollerine göre hazırlanmadığı için arama motorlarında taranamamakta, diğer arşivlerde harmanlanamamakta ve dolayısıyla ulusal ve uluslararası görünürlük azalmaktadır (Tonta ve Akbulut, 2019, s. 222).

Üst veride standartlar; verinin bulunabilir, erişilebilir ve yeniden kullanılabilir olmasını ve sistemlerin birlikte çalışabilirliğini sağlamaları nedeniyle önemlidir (bkz. [Bölüm 2.4.1 FAIR İlkeleri](#)). Yapılan bir çalışmaya göre, standart eksikliği önemli ölçüde üretkenlik kaybı ve verilerin daha az bulunabilir olduğu anlamına gelmektedir (Digital Archiving Consultancy, Bioinformatics, 2005, s. 8).

Genel konulu veya disipline özgü birçok üst veri standardı geliştirilmiştir. Dublin Core (DC) ve Data Documentation Initiative (DDI) yaygın olarak kullanılan genel konulu üst veri standartlarıdır. Konuya özgü üst veri standartları için RDA Üst Veri Standartları Dizini (RDA Metadata Standards Directory) incelenebilir. Bu dizinde (<http://rd-alliance.github.io/metadata-directory>) ayrıca üst veri oluşturma araçları hakkında da bilgi sunulmaktadır. Araştırma projeleri için üst veri oluşturmada yardımcı üst veri araçları (Annotare, CEDAR Workbench, ISA Creator, Morpho, OMERO, OntoMaton, RightField gibi) kullanılabilir (Stanford Libraries, 2022).

3.3 DEPOLAMA VE KORUMA

Verinin doğru formatta saklanması ve uzun süreli koruma için hazırlanması işlemlerini de içeren veri yönetimi veri kayıplarını önlemek için çok önemlidir. En çok insan hatası olmak üzere, donanım/yazılım hatası, sistem eskimesi, altyapı arızası, doğal afet gibi birçok nedenden dolayı veri kayıpları yaşanmaktadır (Peters, 2017). Kullanılan bilgisayarın çalınması nedeniyle verilerini kaybeden araştırmacı sayısı da azımsanmayacak kadar çoktur. Yedekleme yapmadığı için özgün verilerini kaybeden, emekleri boşa giden ve verileri için ödül koyan araştırmacılar bulunmaktadır (Kaiseer, 2008; Van Uffelen, 2019). Hırsızlık gibi olayları önlemek her zaman mümkün olmamakla birlikte uygun depolama ve yedekleme yöntemleriyle veri kayıplarına karşı koruma sağlanabilmektedir (Briney, 2015, s. 117). Koruma sürecinin planlanmasında; dosya formatları ve veri kalitesi, veri sahipliği, saklama süreleri, tercih edilen veri havuzları ve verilerin güvenli bir şekilde paylaşılması hususları dikkate alınmalıdır (James Cook University, 2020).

Bu bölümde araştırma süresince verilerin depolanması, yedeklenmesi, veri güvenliği ve araştırma bittikten sonra verilerin uzun süreli korunması konularına ilişkin bilgi verilmiştir.

3.3.1 Depolama ve Yedekleme

Depolama için birçok donanım seçeneği mevcut olmakla birlikte her bir donanımın avantajları olduğu gibi dezavantajları da bulunmaktadır. 2015 yılında yapılan bir çalışmaya göre, yaygın olarak kullanılan donanım seçeneklerinin kullanımına yönelik öneriler şu şekildedir (Briney, 2015, s. 118):

- Kişisel bilgisayarlar, diskler, yerel sürücüler ve manyetik bantlar başka bir depolama ortamıyla birlikte kullanılabilir
- CD/DVD kullanımı zahmetli ancak kabul edilebilir
- Bulut depolama hizmet şartları uygunsa kullanılabilir
- USB bellek sadece dosya transfer etmek amacıyla kullanılabilir.

Teknolojinin hızla ilerlediği günümüzde, veri depolama ortamları da değişmektedir. Bugün geçerli olan teknolojiyle kullanılan aygıtlar, gelecek zamanda kullanılamayabilir. Bunun en iyi örneği, floppy disklerdir. Şu an nasıl bu aygıtları çalıştıracak donanım bilgisayarlarımızda bulunmuyorsa, yakın zamanda CD/DVD için de aynı şey olacaktır. Bu yüzden veriler farklı ortamlarda yeni teknolojiler kullanılarak yedeklenmelidir.

Düzenli bir veri yedekleme stratejisi ile veri kayıplarını önlemek veya en aza indirmek mümkündür. Araştırma verilerinin yedeklenmesinde 3-2-1 kuralının uygulanması tavsiye edilmektedir. Bu kurala göre, en az iki farklı depolama ortamında ve biri dış ortamda (off-site) olmak üzere verilerin üç kopyasının kaydedilmesi gerekmektedir (Berkeley Library, t.y.). Örneğin, orijinal kopya kişisel bilgisayarda, yedek kopyalardan biri üniversite sunucusunda ve diğer yedek kopya ise buluta kaydedilebilir. USB belleklerin yedekleme için kullanılması önerilmemektedir.

Yedekleme zor ve pahalı bir süreç olduğundan, verinin boyutunun büyüklüğü hangi belgelerin, nerede, ne sıklıkla yedekleneceğini etkilemektedir. Yedeklemede temel kural önemli ve sık değiştirilen dosyaların düzenli olarak yedeklenmesidir (University of Cambridge, 2020a). Ancak kişisel, hassas veya gizli verilerin yedeklenmesinde 3-2-1

kuralının uygulanmaması, mümkün olduğunca az kopyanın (bir orijinal, bir yedek gibi) olması ve saklama için [Bölüm 3.3.3](#)'te bildirilen hususlara dikkat edilmesi gerekmektedir (UK Archive, 2011, s. 18)

3.3.2 Veri Güvenliği

Veri güvenliği; fiziksel güvenlik, ağ güvenliği, bilgisayar sistemlerinin ve dosyalarının güvenliğini içeren çok kapsamlı bir konudur (UK Data Service, 2020a). Ancak bu bölümde sadece, yetkisi olmayan kişilerin verilere erişiminin engellenmesinde önemli bir işlem olan şifrelemeden bahsedilecektir.

2015-16 yılları arasında Avustralya'da çalınan veya kaybolan dizüstü bilgisayarlar üzerine yapılan bir çalışma, veri güvenliğinin ve şifrelemenin önemine dikkat çekmektedir. Bu çalışmada, 33 dizüstü bilgisayarın yaklaşık yarısının (16) sağlık sektörüne mensup kişilere ait olduğu ve beş milyondan fazla insana ait verilerin ihlal edildiği; kişisel, özel ve gizli bilgiler için şifreleme kullanılan sadece bir bilgisayar olduğu belirtilmektedir. Çoğu bilgisayarda şifreleme kullanılmamış ve veri güvenliği sağlanamamıştır (Wakeling, Hannay ve Baig, 2017). Benzer şekilde, bir klinik araştırma projesine kayıtlı yaklaşık 2500 hastaya ait bilgileri şifrelenmemiş olarak saklayan ABD Ulusal Sağlık Enstitüsü (NIH) araştırmacısına ait bilgisayarın çalınmasıyla da veri güvenliği ihlali yaşanmıştır. Bu olay üzerine kuruluş, veri koruma politikalarını gözden geçirerek politikaların uygulanabilirliği artırmıştır (Kaiseer, 2008).

Yetkisi olmayan kişilerin verilere erişimini engellemek için şifreleme yöntemiyle olası risklerin azaltılması mümkündür. Şifreleme, verilerin sadece yetkili kişiler tarafından görüntülenebileceği bir kodlama işlemi olup tek tek dosyalar, klasörler, diskler ve USB depolama aygıtları şifrelenebilmektedir (UK Data Service, 2020a). Şifreleme için; VeraCrypt, AxCrypt, FileVault 2 ve BitLocker gibi yardımcı araçlar bulunmaktadır.

3.3.3 Uzun Süreli Koruma

Uzun süreli koruma, veri bütünlüğünün bozulmadan saklanması, depolanması ve veriye erişim sağlanması sürecidir.

Tablo 5. Saklama için tercih edilen formatlar (UK Data Service, 2020c).

| Veri Türü | Önerilen format | Kabul edilebilir format |
|--|--|---|
| Kapsamlı üst veriler içeren tablo verileri (Değişken etiketleri, kod etiketleri ve tanımlanmış eksik değerler) | Taşınabilir SPSS formatı (.por) Ayrılmış metin ve komut ('setup') dosyası (SPSS, Stata, SAS gibi) Üst veri bilgisi içeren yapılandırılmış metin ya da işaretleme dosyası (DDI XML dosyası gibi) | Tescilli istatistik programlarının formatları: SPSS (.sav), Stata (.dta), MS Access (.mdb/.accdb). |
| Minimum üst veri içeren tablo verileri (Sütun başlıkları, değişken isimleri gibi) | Virgülle ayrılmış değerler (.csv) Sekmeyle ayrılmış dosya (.tab) | Yaygın olarak kullanılan formatlar: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf), OpenDocument (.ods) |
| Metinsel veri | Zengin Metin Biçimi (.rtf) Düz metin, ASCII (.txt) Uygun bir Belge Türü Tanımına (Document Type Definition-DTD) veya şemaya göre Genişletilebilir İşaretleme Dili (eXtensible Mark-up Language) (.xml) metni | Hipermetin işaretleme dili (HTML) (.html) Yaygın olarak kullanılan formatlar: MS Word (.doc/.docx) Bazı özel yazılım programları: NUD*IST, NVivo and ATLAS.ti |
| Resim dosyaları | Sıkıştırılmamış TIFF versiyon 6 (.tif) | JPEG (.jpeg, .jpg, .jp2) (orijinali bu formatta yaratıldıysa) GIF (.gif) diğer TIFF versiyonları (.tif, .tiff) RAW (.raw) Photoshop dosyaları (.psd) BMP (.bmp) PNG (.png) Taşınabilir Doküman Formatı (PDF/A, PDF) (.pdf) |
| Ses dosyaları | FLAC (.flac) | MPEG-1 Audio Layer 3 (.mp3) (orijinali bu formatta yaratıldıysa) Audio Interchange File Format (.aif). Waveform Audio Format (.wav). |
| Video | MPEG-4 (.mp4) OGG video (.ogv, .ogg) motion JPEG 2000 (.mj2) | AVCHD video (.avchd). |
| Dokümantasyon ve komut dosyaları | Zengin Metin Biçimi (.rtf), PDF/UA, PDF/A veya PDF (.pdf), XHTML veya HTML (.xhtml, .htm) OpenDocument Metni (.odt) | Düz metin (.txt) Yaygın olarak kullanılan formatlar: MS Word (.doc/.docx), MS Excel (.xls/.xlsx) uygun DTD veya şemaya göre XML işaretli metin (.xml) |

Veri kayıplarının önlenmesi, uzun süreli erişimin ve yeniden kullanımın sağlanması, herhangi bir olumsuz durumda araştırma sonuçlarının doğrulanabilmesi açısından kritik öneme sahiptir (James Cook University, 2020). Araştırma verilerinin gelecekte kullanılabilir olmasını sağlamak için, araştırma sırasında dokümantasyon işleminin etkili bir şekilde yapılması ve korumaya hazırlık sürecinin iyi planlanması gerekmektedir.

Teknolojinin hızla gelişmesiyle hem verilerin kaydedildiği ortamlar hem de kullanılan yazılımlar değişmekte veya güncellenmektedir. Örneğin, önceleri floppy diskler depolama ortamı olarak kullanılırken şu an bilgisayarlarda bu disklerin yerleşeceği donanım bulunmamaktadır. Veya eski sürümle oluşturulan bir belge zaman içinde yeni sürümde açılmamakta, açılrsa bile içerikte bozulmalar olabilmektedir. Verilerin bozulmadan muhafaza edilebilmesi ve veri kayıplarının yaşanmaması için şu hususlara dikkat edilmelidir (University of Cambridge, 2020a; University of California Curation Center, 2020).

- Koruma sürecinin araştırmanın başında planlanması
- Birden fazla kopyanın farklı ortamlarda saklanması
- Verilerin belirli aralıklarla yeni depolama ortamlarına transfer edilmesi
- Verilerin yeni yazılım sürümlerine taşınması
- Saklama için tercih edilen formatların kullanılması (bkz. Tablo 5)
- Formatın tescilli olmaması
- Açık ve belgelenmiş standartlarda olması
- Araştırma topluluğu tarafından kullanılıyor olması
- ASCII, UTF-8 gibi standart karakter kodlarının kullanılması
- Sıkıştırılmamış olması

3.4 YAYIMLAMA VE PAYLAŞMA

Veri yayımlama; araştırmacıların kendi araştırma verilerini genel, disipline özgü veya kurumsal arşivlerde diğer araştırmacılar tarafından tekrar kullanılması amacıyla depolanması ve erişime sunmasıdır (Bote ve Termens, 2019, s. 329), Araştırma

verilerinin yönetimi, verilerin paylaşarak diğer arařtırmacıların erişimine ve kullanımına açıldığı sürece anlam kazanmaktadır.

Arařtırma başlamadan verilerin toplanması veya mevcut verilerin kullanılması, saklanması, depolanması ve paylaşılması aşamalarında etik ve yasal konulara dikkat ederek planlamaların yapılması gerekmektedir.

[Bölüm 2.4](#)'te veri paylaşmanın yararlarına yer verilmişti. Ancak bazı arařtırmacılar çeşitli nedenlerle veri paylaşımına olumsuz yaklaşmaktadır. Veri paylaşmama nedenleri řu şekildedir (Aydınođlu ve diđerleri, 2017, s. 279):

- Zaman ve kaynak eksikliği
- Bilgi talepleriyle ilgilenmek için zamanın olmaması
- Veri yönetiminde deneyim eksikliği veya bilgi yetersizliği
- Verinin anlaşılabilir olmayışı
- Yasal veya etik kısıtlamalar
- Verileri nerede arşivleyeceğini bilememe
- Verinin hatalı kullanılabileceđi veya yanlış yorumlanabileceđi endişesi
- Başkalarının bu verilere ihtiyacı olmadığı düşüncesi
- Belirli bir zorlayıcılığın olmayışı
- Belirli bir ödül sisteminin olmayışı

Yapılan bir çalışmada, bunların yanı sıra psikolojik ve sosyal engellerin de veri paylaşımına engel olabileceđi belirtilmiş ve cinsiyet, yaş ve unvana göre veri paylaşım durumlarında istatistiksel açıdan anlamlı farklılık olduğu saptanmıştır (Zhu, 2020). Erkekler, 35 yaş üstü arařtırmacılar ve profesörler diđer gruplara oranla verilerini veri arşivlerinde daha fazla paylaşmaktadırlar.

Arařtırmacılar, ayrıca veri setindeki bilgilerin yasal olarak korunması veya hassas veri olabilmesi nedeniyle veri paylaşımı konusunda endişe duyabilmektedir (University of Minnesota Libraries, 2020b).

Yasal olarak korunmada; kanunlar, yönetmelikler, fon sağlayıcı kuruluş veya mensup olunan kurumun politika belgeleri bağlayıcı niteliktedir. Türkiye'de kişisel veriler, 6698 sayılı Kişisel Verilerin Korunması Kanunu'na göre korunmaktadır. Bu kanunda:

- Kimliği belirli veya belirlenebilir gerçek kişiye ilişkin her türlü bilgi **kişisel veri**;
- Kişilere ait ırk, etnik köken, siyasi düşünce, felsefi inanç, din, mezhep veya diğer inançlar, kılık ve kıyafet, dernek, vakıf ya da sendika üyeliği, sağlık, cinsel hayat, ceza mahkûmiyeti ve güvenlik tedbirleri ile ilgili veriler ile biyometrik ve genetik veriler **özel nitelikli kişisel veri** olarak tanımlanmıştır.

Bu tür veriler, kişinin kimliğini doğrudan tanımlayabileceği gibi dolaylı olarak da tanımlayabilir. Doğrudan tanımlayıcı olmayan veriler, kombinasyon halinde tanımlayıcı hale gelebildikleri için paylaşım durumunda dikkat etmek gerekmektedir (PLoS One, 2019).

Kişisel verilerin yanı sıra; nesli tükenmekte veya tehdit altında olan bitki ve hayvan türlerini tehlikeye atacak ekolojik veriler, sit alanlarındaki saha çalışmalarından elde edilen veriler, arkeolojik alanların yer bilgileri, insanlara zarar verme ihtimali olan veriler ve ticari araştırma fonu anlaşmaları gereği gizli olan veriler de hassas veri olarak nitelendirilmektedir (Hampton ve diğerleri, 2015, s. 6; PLoS One, 2019; Mantra Research Data Management Training, 2020c).

Önceden hassas verilerin paylaşılmasının etik olmadığı veya çok zor olduğu anlayışı yaygınken bu anlayış yerini hassasiyetin kaldırılarak yani kimliği gizleme yöntemiyle koşullu paylaşmaya bırakmıştır (ANDS Guides, 2018). Örneğin, PLoS dergi politikasında hassas verilerin gerekli şartlar yerine getirilerek paylaşılmasının desteklendiği bildirilmiştir (PLoS One, 2019).

Hassasiyetin ortadan kaldırılması için anonimleştirme (anonymization) veya takma ad kullanma (pseudonymization) yöntemleriyle veriler paylaşılabilir hale getirilebilmektedir. Anonimleştirme kişisel olarak tanımlayıcı olan verilerin kalıcı olarak ve tamamen ortadan kaldırılması eylemidir (EDUCAUSE, 2020). Takma adlandırma ise tanımlayıcıların yerine araştırmacı tarafından oluşturulan takma adlar veya kodlar kullanılarak verinin gizlenmesi yöntemidir (The University of British Columbia, 2018). Ancak bu yöntemi kullanırken oluşturulan kodun kişisel bilgilerden türetilmemiş ve kişiyi tanımlayacak şekilde çözümlenemez olmasına dikkat edilmelidir (University of Minnesota, 2020). İki yöntem arasındaki fark ise; anonimleştirilmiş verilerde veri setleri arasında bağlantı kurulamazken, takma ad kullanılan verilerde veri setleri arasında bağlantı kurulabilmesidir (The University of British Columbia, 2018).

Veri paylaşımında dikkat edilmesi gerekenler maddeler halinde aşağıda özetlenmiştir (ANDS Guides, 2018):

- Çalışmaya başlamadan önce paylaşım planları yapılmalıdır
- Veri toplamadan önce gerekli Etik Komite izinleri alınmalıdır
- Paylaşım için katılımcılar bilgilendirilmeli ve rıza formları alınmalıdır
- Hâlihazırda mevcut olan veriler kullanılacaksa atıf verilmelidir
- Verinin sahibinin kim olduğu önemli olup veriyi yayımlama hakkına sahip olunmalıdır
- Hassas veriler varsa anonimleştirme, takma adlandırma gibi yöntemlerden uygun olanı kullanılarak hassasiyet ortadan kaldırılmalıdır
- Veriler tekrar kullanım ve kaynak olarak gösterilebilmesi için lisanslanmalıdır.

Veri lisanslama kullanım hakları belirlenerek izin verilip verilmeyen durumların netleşmesi açısından önem taşımaktadır. Uluslararası olması ve yaygın olarak kullanılması nedeniyle Tablo 6’da Creative Commons ve Open Data Commons lisansları ve açıklamaları sunulmuştur (Briney, 2015, s. 153; Holt, 2018).

Tablo 6. Lisanslar ve koşulları

| Lisans türü | Creative Commons | Open Data Commons | Lisans Açıklaması |
|--|-------------------------|--------------------------|--|
| Kamu malı | CC0 | PDDL | Kullanımında herhangi bir kısıtlama yoktur |
| Atıf | CC BY | ODC-By | Atıf verme koşuluyla kullanılabilir |
| Atıf - Aynı Lisansla Paylaş | CC BY-SA | ODC-ODbL | Aynı lisansla lisanslandığı ve atıf verildiği sürece tekrar kullanılabilir |
| Atıf - Gayri Ticari | CC BY-NC | | Ticari amaç dışında, atıf verildiği sürece tekrar kullanılabilir |
| Atıf - Türetilemez | CC BY-ND | | Orijinal çalışma yeniden kullanım sırasında değiştirilemez ve atıf verilmeli |
| Atıf - Gayri Ticari - Aynı Lisansla Paylaş | CC BY-NC-SA | | İçeriği yalnızca ticari olmayan amaçlarla yeniden kullanmak için atıfta bulunulmalı ve türetilmiş çalışmaları aynı lisansla paylaşılmalı |
| Atıf - Gayri Ticari – Türetilemez | CC BY-NC-ND | | Orijinali değiştiremez veya ticari olarak kullanılamaz ve atıf verilmeli |

Veri arşivlerinin listesini içeren Re3dataorg (Registry of Research Data Repositories), araştırma verilerinin depolanması ve erişime sunulması için konuya, veri türüne ve/veya ülkelere göre uygun veri arşivlerinin seçilmesinde kolaylık sağlamaktadır (re3data.org -

Registry of Research Data Repositories, t.y.). Lisanslanan veri/veri setleri seçilen veri arşivlerinde yayımlanarak diğer araştırmacıların kullanımına sunulmalıdır.

3.5 KEŞFETME VE YENİDEN KULLANMA

Veriler farklı işlevlerle tekrar kullanılabilir. Toplama (aggregation), entegrasyon (integration), meta-analiz (meta-analysis), kopyalayarak çoğaltma (replication), yeni bir amaçla tekrar kullanım (repurposing) ve sentez (synthesis) yeniden kullanımın farklı türleridir (Curty, 2015, s. 37). Bu aşamada verilerin keşfedilebilir olması için FAIR ilkelerine uygun bir biçimde veri arşivlerinde kullanıma sunulması önem taşımaktadır. Araştırma sırasında yapılan dokümantasyon ile üst veri işlemleri verinin keşfedilebilmesi ve anlaşılabilir olması için yapılmaktadır. Amaç sadece verileri saklamak değil araştırmanın bütünlüğü içinde açıklayıcı unsurlarla bir arada tutarak tekrar kullanıma sunmaktır.

Genel, kurumsal ve disipline özgü birçok veri arşivi bulunmakla birlikte, bu bölümde en yaygın olarak kullanılan genel arşivlerden kısaca aşağıda bahsedilmektedir.

- Figshare: Araştırmacılara verilerin saklanması, organize edilmesi, paylaşılması ve diğer araştırmaların keşfedilmesi için olanaklar sunmaktadır. Kullanıcıların herhangi bir dosya formatında veri yükleyebildikleri bu uygulama, sadece veri depolama amaçlı kullanılabileceği gibi özel bağlantı paylaşımı ile kısmi veya herkese açık veri paylaşımı yapılabilen bir platformdur. Veriler için DOI alma veya ayırtma imkânı da bulunmaktadır. Kurumsal ve yayıncı abonelikleri, FAIR ilkelerine dayalı ve meslektaşlar arası iş birliğini destekleyici özellikleriyle çalışmaların görünebilir ve atıf alabilir olmasını sağlamaktadır (Figshare, 2022.).
- Dryad: Kâr amacı gütmeyen, araştırma verilerinin açık erişilebilir, bilimsel literatürle bütünleşik ve tekrar kullanılabilir olduğu açık kaynak bir platformdur (Dryad, 2020).
- Zenodo: OpenAIRE projesi kapsamında araştırma programlarını desteklemek için CERN tarafından geliştirilen Zenodo, Avrupa Komisyonu tarafından finanse edilen araştırmalara ait verilerin ve sonuçların paylaşılması amacıyla dünyanın

her yerinden farklı disiplinlerdeki arařtırmacılara hizmet sunmaktadır (Zenodo, 2020).

Image Data Resource, European Genome-phenome Archive GenBank, Trace, SRA, SNP, GEO disipline özgü konu arřivlerine, APERTA ise kurum arřivlerine örnek olarak verilebilir.

Uygun arřivin bulunmasına yardımcı Re3dataorg kullanılarak ve/veya genel, konusal, kurumsal arřivler taranarak veriye erişim sağlamak mümkündür. Veri/veri setlerinin lisans şartları, dokümantasyon ve üstveri bilgilerinin yeterliliđi, veri saklama koşulları, onam derecesinin ve veri formatlarının uygunluđu kontrol edilerek yeniden kullanım gerçekleştirilmelidir (Aydınolu, 2019). Yeniden kullanımda mutlaka atıf yapılmalıdır. Verilere/veri setlerine uygun şekilde atıf yapılması verinin bulunmasını kolaylařtırma, arařtırmanın şeffaflıđını ve tekrarlanabilirliğini artırma, yeniden kullanıma teřvik etme, veri/veri setlerinin kalıcılıđını destekleme, veri üreticisinin tanınmasını, güvenilirliğini, görünürlüğünü artırma gibi hem veri üreticisi hem de veri kullanıcıları için fayda sağlamaktadır (Cornell University, t.y.).

Veri kaynaklarına atıf yapmak için yazar, başlık, yayın yılı, yayıncı ve/veya dađıtıcı ve kalıcı tanımlayıcı bilgisi olması gereken bileşenlerdir (IASSIST SIGDC, 2012). APA, MLA, Chicago gibi yardımcı atıf stilleri ve/veya Dataverse, Dryad, Figshare gibi veri arřivlerinin atıf kılavuzları kullanıcılara atıf verme biçimleriyle ilgili açıklayıcı bilgiler sunmaktadır (Michigan State University Libraries, 2021). Özel iletişim kanallarıyla temin edilen yayımlanmamıř verilerde de atıf ilkeleri uygulanmakta (Harvard Medical School, 2022) ancak entelektüel mülkiyet haklarının tanımlanması verinin yayımlanmasıyla sağlamaktadır (Curty, 2015, s. 29). Veri hakkında ne kadar bilgi sunulursa diđer kullanıcıların veriye ulaşması ve o veriyi kullanması o kadar kolaylařmaktadır.

4. BÖLÜM: BULGULAR VE DEĞERLENDİRME

Bu bölümde; Hacettepe, Orta Doğu Teknik, İstanbul, İstanbul Teknik ve Ankara üniversitelerinde 2014-2018 yılları arasında TÜBİTAK desteği ile yürütülen projelerde görev alan araştırmacılara uygulanan anket sonuçları değerlendirilmiştir.

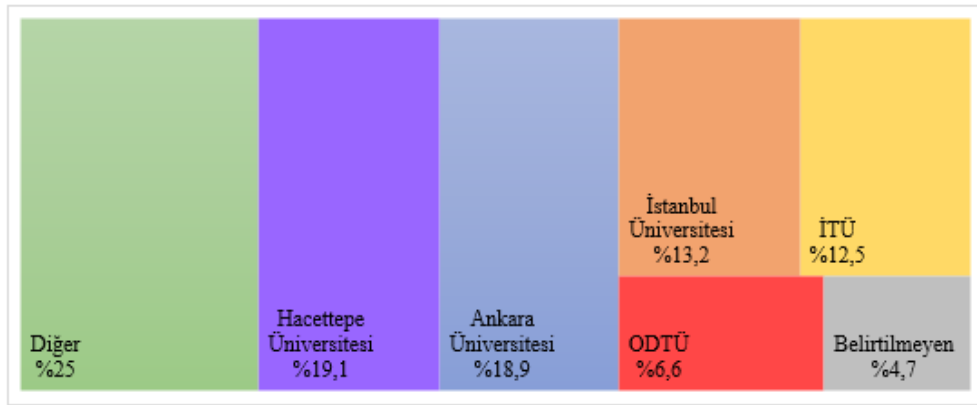
2014-2018 yılları arasında adı geçen üniversitelerde TÜBİTAK desteğiyle üretilen 689 proje olduğu, bu projelerde görev alan ve araştırmanın yapıldığı tarihlerde kurum ayrımı yapılmaksızın herhangi bir üniversitede çalışan araştırmacı sayısının 1572 olduğu belirlenmiştir. 26.02.2020-15.01.2021 tarihleri arasında, araştırmanın evrenini oluşturan 1572 araştırmacının hepsine anket gönderilmiş ancak 455 katılımcıdan yanıt alınmıştır. Bu yanıtlardan 13 tanesinin geçersiz olduğu saptanmıştır. Anket uygulanan kişi sayısı 442 olup anket yanıtlanma oranı %28,1'dir. Katılımcılardan 34 tanesi dijital araştırma verisi kullanmadığını belirttiğinden bu bölümdeki bulgular dijital araştırma verisi kullanan 408 katılımcıya aittir.

Anket üç bölümden oluşmaktadır. İlk bölümde araştırmacıların araştırma verilerinin yönetimi konusundaki farkındalıklarını ve uygulama eğilimlerini belirlemeyi amaçlayan 23 soru, ikinci bölümde veri paylaşım yaklaşımlarına yönelik altı soru, son bölümde ise demografik bilgileri kapsayan beş soru bulunmaktadır.

Ankette doğrudan sorulan “bölüm”, “yaş” ve “deneyim” soruları sonradan kategorilere ayrılmıştır. Disiplin kategorizasyonu yapılırken TÜBİTAK proje kodlarından yararlanılmış ve tıp, fen, mühendislik, sosyal bilimler ve ziraat olmak üzere beş kategori belirlenmiştir. Yaşa ve deneyime göre sınıflandırma yapılırken ise, saçılım grafiklerinden yararlanılmış ve her iki alan için de dört grup olması gerektiğine karar verilmiştir. Her bir gruptaki aralığın eşit olması için, değişim aralığı (en büyük değer ile en küçük değer farkı) grup sayısına bölünerek sınıf aralığı hesaplanmış ve buna göre gruplandırma yapılmıştır.

4.1 DEMOGRAFİK BİLGİLER

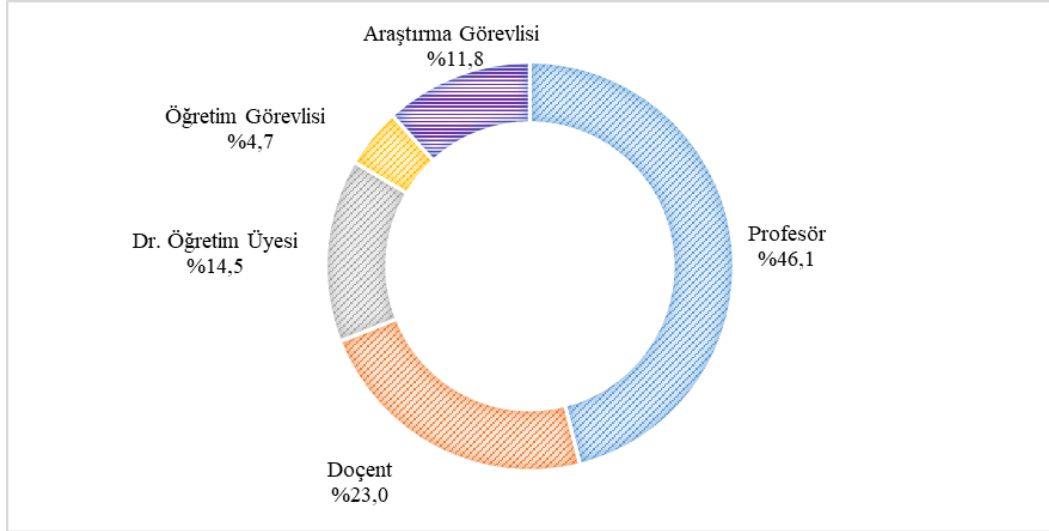
Araştırmaya 73 farklı üniversiteden toplam 408 araştırmacı katılmıştır. Araştırmacıların %19,1'i (n=78) Hacettepe Üniversitesi, %18,9'u (n=77) Ankara Üniversitesi, %13,2'si (n=54) İstanbul Üniversitesi, %12,5'i (n=51) İTÜ ve %6,6'sı (n=27) ODTÜ mensubudur (bkz. Şekil 5). Katılımcıların dörtte birini (n=102) oluşturan diğer üniversiteler, bir ile beş arasında değişen katılımcı sayısı ile 68 farklı üniversiteyi temsil etmektedir. Katılımcıların yaklaşık %5'i (n=19) ise kurum bilgisi sorusuna cevap vermemiştir.



Şekil 5. Katılımcıların kurumlara göre dağılımı

Katılımcıların %24'ü (n=98) tıp, %22,8'i (n=93) fen, %20,1'i (n=82) sosyal bilimler, %18,1'i (n=74) mühendislik ve %13'ü (n=53) ziraat alanına mensuptur. Disiplin bilgisi belirtmeyen 8 kişi (%2) bulunmaktadır. Bu nedenle disipline göre yapılan karşılaştırmalarda, oranlar 400 araştırmacı üzerinden hesaplanmıştır.

Unvanlara göre dağılım Şekil 6'da sunulmuş olup araştırmacıların %46,1'i (n=188) Profesör, %23'ü (n=94) Doçent, %14,5'i (n=59) Dr. Öğretim Üyesi, %11,8'i (n=48) araştırma görevlisi ve %4,7'si (n=19) öğretim görevlisidir.



Şekil 6. Katılımcıların unvanlara göre dağılımı

Katılımcıların yaş (bkz. Tablo 7) ve deneyim bilgileri (bkz. Tablo 8) incelendiğinde, yarısından fazlasının 50 yaşından küçük ve 24 yıldan az deneyime sahip oldukları görülmektedir.

Tablo 7. Katılımcıların yaş aralıkları

| Yaş aralığı | N | % |
|-----------------|------------|------------|
| 39 yaştan küçük | 103 | 25,3 |
| 39 – 49 | 160 | 39,3 |
| 50 – 60 | 112 | 27,5 |
| 61 yaş ve üzeri | 32 | 7,9 |
| Toplam | 407 | 100 |

60 yaşından büyük (%7,9) ve 34 yıldan fazla deneyime sahip olanların (%11,3) oranı en azdır. Sadece bir kişi yaş bilgisini paylaşmamıştır.

Tablo 8. Katılımcıların deneyim bilgileri

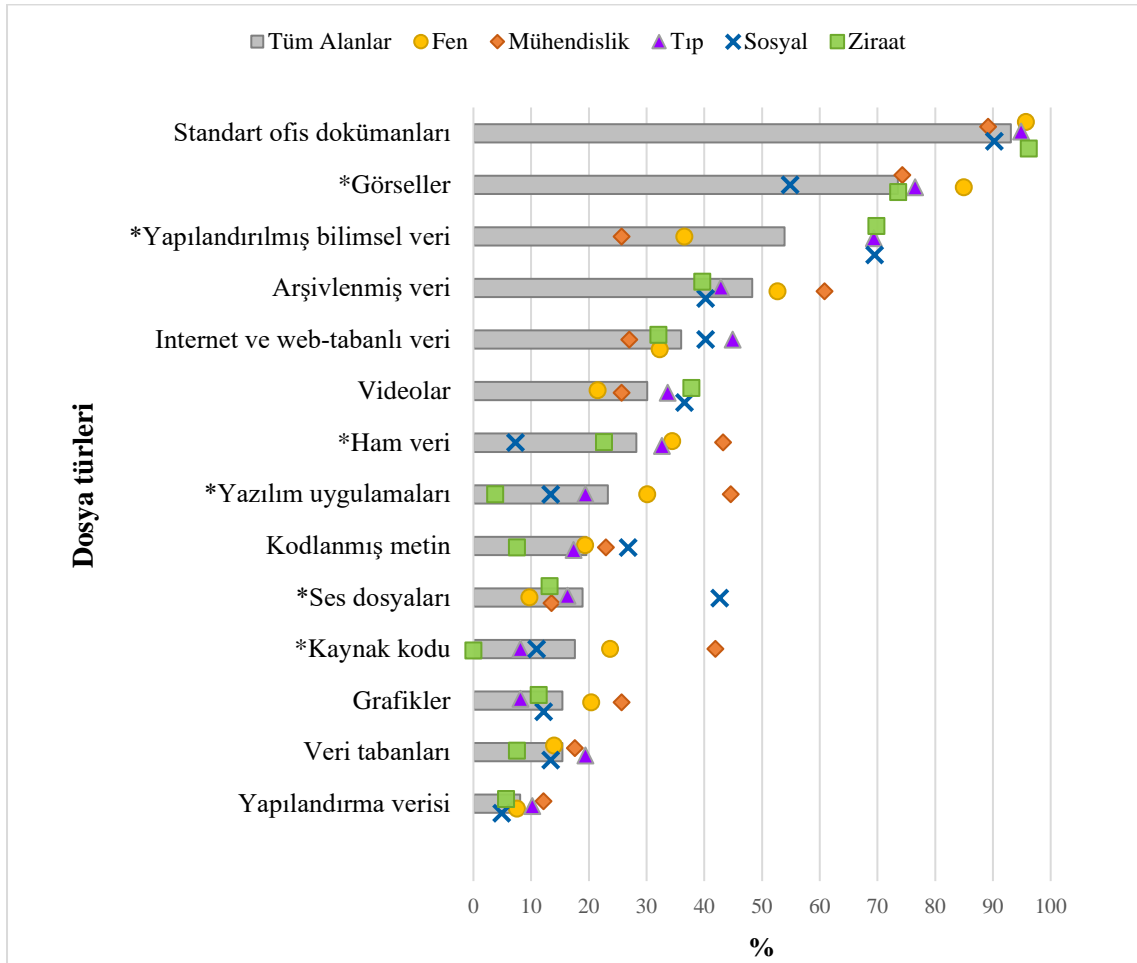
| Yıl | N | % |
|-----------------|------------|------------|
| 14 yıldan az | 118 | 28,9 |
| 14 – 23 | 136 | 33,3 |
| 24 – 34 | 108 | 26,5 |
| 34 yıldan fazla | 46 | 11,3 |
| Toplam | 408 | 100 |

4.2 ARAŞTIRMA VERİLERİ YÖNETİMİ KONUSUNDA FARKINDALIK DURUMU VE UYGULAMA EĞİLİMLERİ

4.2.1 Dosya Türleri

Araştırmacıların kullandıkları dosya türlerini belirlemek ve kullanılan dosya türlerinde disipline göre herhangi bir farklılık olup olmadığını saptamak amacıyla araştırmacılara birden fazla seçeneğin işaretlenebildiği bir soru yönlendirilmiştir. Standart ofis dokümanları gibi bazı dosya türleri, farklı tür/türlerde (metinsel, sayısal, görsel gibi) veriler içerebilmektedir. Alınan yanıtlar doğrultusunda, araştırmalarda en çok *standart ofis dokümanlarının* (%93,1; n=380) ve *görsellerin* (%73,5; n=300) kullanıldığı saptanmıştır. Araştırmacıların yaklaşık yarısı tarafından tercih edilen *yapılandırılmış bilimsel ve istatistiksel veri* (%53,9; n=220), kullanımda üçüncü sıradadır. Araştırmacılar tarafından en az (%8,1; n=33) işaretlenen dosya türü *yapılandırma verisidir* (parametre düzenlemeleri, loglar, kütüphane dosyaları gibi). *Veri tabanları* ve *grafiklerin* kullanım oranı ise aynı ve görece düşüktür (%15,4; n=63).

Disiplinlere göre araştırmalarda kullanılan dosya türleri incelendiğinde, hem benzerlikler hem de farklılıklar olduğu görülmektedir. Disipline göre kullanılan dosya türleri Şekil 7’de sunulmuş olup başında * işareti ile belirtilen seçenekler istatistiksel açıdan anlamlı bir farklılık olduğuna işaret etmektedir. Tüm disiplinlerde en çok *standart ofis dokümanları* kullanılmaktadır. Sosyal bilimler haricindeki diğer tüm disiplinlerde *görseller* en çok kullanılan ikinci dosya türüdür. Sosyal bilimlerde ikinci olarak en çok (%69,5; n=57) kullanılan dosya türü olan *yapılandırılmış bilimsel ve istatistiksel veri*, tıp (%69,4; n=68) ve ziraat (%69,8; n=37) alanlarında da benzer oranlarda kullanılmaktadır. *Arşivlenmiş veri*, *video* ve *yapılandırılmış grafik* kullanım oranları da sosyal bilimler, tıp ve ziraat alanlarında benzerlik göstermektedir.



Şekil 7. Disipline göre kullanılan dosya türleri (%). *p<0,05.

Veri türlerinden; *görseller* ($\chi^2_{(4)} = 21,203$; $p=0,000$), *yapılandırılmış bilimsel ve istatistiksel veri* ($\chi^2_{(4)} = 57,853$; $p=0,000$), *ham (makine tarafından üretilmiş) veri* ($\chi^2_{(4)} = 29,265$; $p=0,000$), *yazılım uygulamaları* ($\chi^2_{(4)} = 37,875$; $p=0,000$), *ses dosyaları* ($\chi^2_{(4)} = 37,799$; $p=0,000$) ve *kaynak kodu* ($\chi^2_{(4)} = 52,514$; $p=0,000$) kullanımını disiplinlere göre istatistiksel açıdan anlamlı bir farklılık göstermektedir. Diğer disiplinlere oranla:

- Mühendislik alanında; *arşivlenmiş veri* (%60,8; $n=45$), *ham veri* (%43,2; $n=32$), *yazılım uygulamaları* (%44,6; $n=33$), *kaynak kodu* (%41,9; $n=31$) ve *yapılandırılmış grafik* (%25,7; $n=19$) kullanımını daha fazla iken, *yapılandırılmış bilimsel ve istatistiksel veri* (%25,7; $n=19$) ile *internet ve web tabanlı veri* (%27; $n=20$) kullanımını daha azdır.

- Fen bilimlerinde *görsel* (%84,9; n=79) kullanımı daha çok ancak *video* (%21,5; n=20) ve *ses dosyaları* (%9,7; n=9) kullanımı daha azdır.
- Sosyal bilimler alanında *görseller* (%54,9; n=45) ve *ham veri* (%7,3; n=6) daha az kullanılırken *ses dosyaları* (%42,7; n=35) ve *kodlanmış metin* (%26,8; n=22) daha çok kullanılmaktadır.
- Ziraat alanında *yazılım uygulamaları* (%3,8; n=2), *kodlanmış metin* (%7,5; n=4) ve *veri tabanları* (%7,5; n=4) kullanımı en azdır. *Kaynak kodu* ise hiç kullanılmamaktadır.
- Tıp alanında *internet ve web tabanlı veri* (%44,9; n=44) kullanımı daha fazlayken *yapılandırılmış grafik* (%8,2; n=8) kullanımı azdır.

Diğer seçeneğine ekleme yapan araştırmacıların belirttikleri veri türleri şu şekildedir: *matrix veri yapıları (.mtx, .mat, veya diğer); netcdf (Network Common Data Form); fastq, bam, sam; analiz programları; fits formatta ham görüntü verileri; R ve RData; EEG verileri (değişik formatlarda); Nvivo nitel araştırma program dosyaları, nvp.*

Literatürdeki çalışmalar incelendiğinde, araştırmacıların en çok kullandıkları dosya türleri açısından bulgular arasında benzerlik ve farklılıklar görülmektedir. Diğer araştırmaların bulguları da en çok üretilen ve kullanılan veri türünün standart ofis dokümanları olduğunu göstermektedir (Vilar ve Zabukovec, 2019, s. 29; Ünal ve Kurbanoğlu, 2018, s. 294). İkinci olarak tercih edilen seçenek olan görseller, Vilar ve Zabukovec'in (2019, s. 29) çalışmayla benzerlik gösterirken Ünal ve Kurbanoğlu'nun (2018, s. 294) bulgularıyla farklılık göstermektedir. Farklılık bu çalışmada ikinci olarak en çok tercih edilen seçeneğin yapılandırılmış bilimsel ve istatistiksel veri dosyaları olmasından kaynaklanmaktadır. Aynı çalışmada, yapılandırılmış grafikler, ham veri, yazılım uygulamaları, kaynak kodu ve yapılandırma verisi gibi dosya türlerinin fen bilimlerinde diğer disiplinlere oranla daha fazla kullanıldığı görülmüştür (Ünal ve Kurbanoğlu, 2018, s. 294). Söz konusu çalışmada disiplin sınıflandırmasında mühendislik alt alanının fen alanı içinde değerlendirildiği göz önüne alındığında, sonuçların benzerlik gösterdiği söylenebilir.

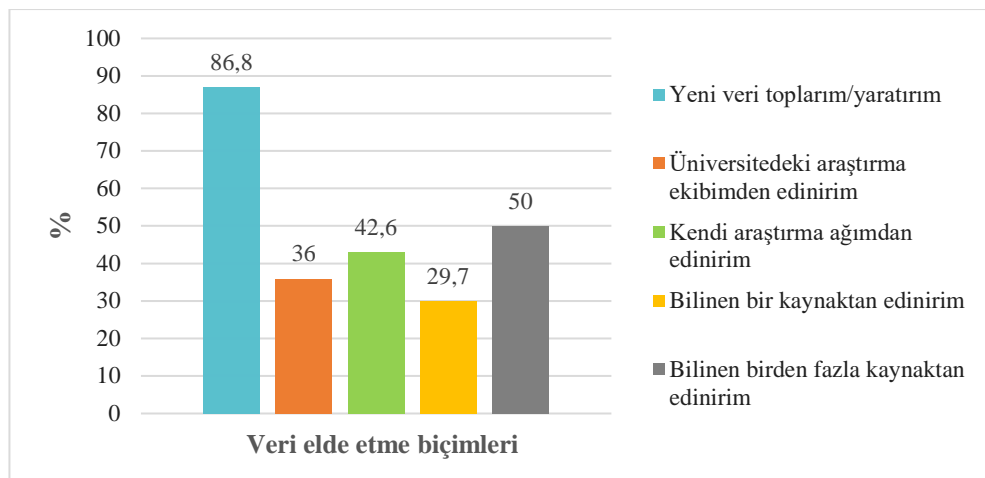
Bazı çalışmalarda ise, araştırmalarda en çok kullanılan dosya türlerinin metin dosyası ve hesap işlem tablosu olduğu ortaya çıkmıştır (Zencir, 2019; s. 144-145; Buys ve Shaw,

2015, s. 9). Standart ofis dokümanlarının bu dosya türlerini kapsadığı düşünüldüğünde, elde edilen bulgular literatürdeki çalışmaları desteklemektedir.

4.2.2 Veri Elde Etme Biçimleri

Araştırmacıların veri elde biçimlerini belirlemek için, araştırmacılara birden fazla seçeneğin işaretlenebildiği bir soru yönlendirilmiştir. Verilen yanıtlar doğrultusunda; katılımcıların çeşitli yöntemleri bir arada kullandığı, büyük çoğunluğunun genelde (%86,8; n=354) *yeni veri oluşturduğu*, yarısının (n=204) *bilinen birden fazla kaynaktan* da veri sağladığı, %42,6'sının (n=174) ise *kendi araştırma ağından* (kendi üniversitesi dışında yer alan araştırmacılardan) veri elde ettiği görülmektedir (bkz. Şekil 8). Diğer seçeneğini işaretleyen katılımcılardan bazıları verdiği yanıtlar nedeniyle, (Üç katılımcı deneysel çalışmalar sonucu ve bir katılımcı arazi çalışmaları) yeni veri oluşturanlar kategorisinde değerlendirilmiştir. Açıklama yapan diğer katılımcıların ifadeleri ise şu şekildedir: *araştırma grubumuzca hazırlanmış kaynak kodlarını içeren simülasyon sonuçları; daha önce araştırma ekibimizin ürettiği bilgisayar kodları; bilimsel yayınlar, teknik raporlar; araştırma konusuna göre çok değişkenlik gösterebilmektedir; geleneksel yöntemleri de kullanıyorum. Basılı "gerçek" kaynaklar ve dijital olmayan kütüphane; ülkemizdeki resmi kurumlardan, çoğunlukla ücret karşılığı temin ederim.*

Açıklamalar, bazı katılımcıların veri ve bilgi kavramlarını karıştırmış olabileceğini düşündürmektedir.



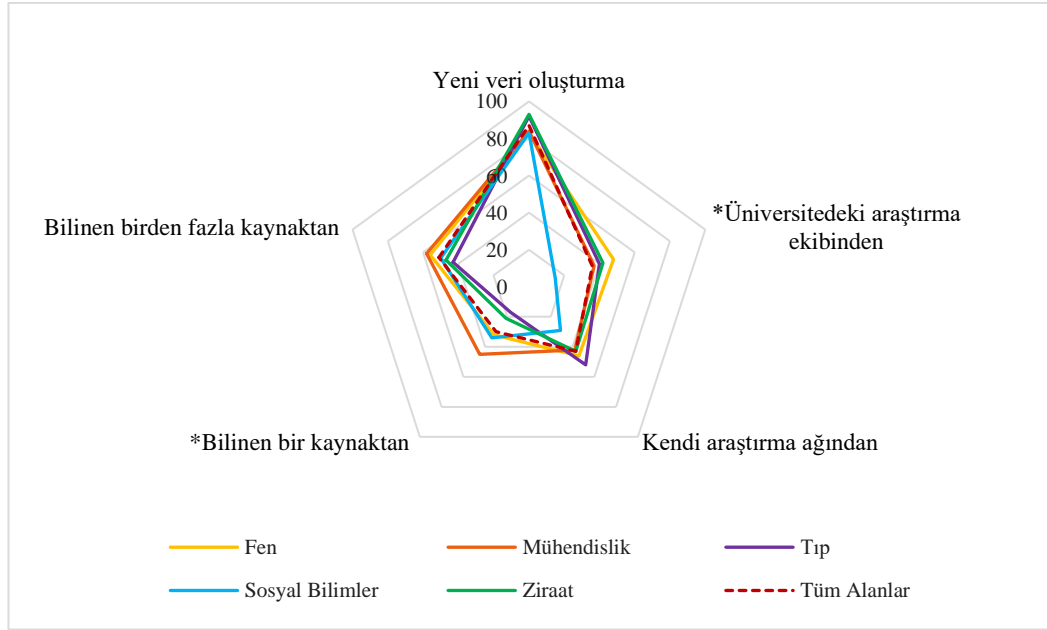
Şekil 8. Araştırmacıların veri elde etme biçimleri (%)

Slovenya'daki arařtırmacılar üzerine yapılan alıřmada, oęunluęun yeni veri oluřturduęu grlmektedir (Vilar ve Zabukovec, 2019, s. 30). Yapılan bařka bir alıřmada ise, katılımcılara verilerini ne sıklıkta ve hangi yollarla elde ettikleri sorulmuřtur. Katılımcıların byk bir oęunluęunun ara sıra (%95,3) ve oęunun (%76,9) sıklıkla yeni veri toplamıř veya oluřturmuř olduęu sonucuna ulařılmıřtır (Kennan ve Markauskaite, 2015, s. 80). nal ve Kurbanoęlu tarafından 2018 yılında lkemizdeki arařtırmacılar üzerine yapılan alıřmada, arařtırmacıların sırasıyla en ok birden fazla kaynaktan (%58), sahip olduęu arařtırma aęından (%56) ve kendi verisini oluřturma (%55) yoluyla veri elde ettięi grlmektedir (nal ve Kurbanoęlu, 2018, s. 295). Sıralaması farklı olmakla birlikte, veri elde etmede tercih edilen ilk  seenek arařtırma bulgularımızla rtşmektedir.

Arařtırmacıların veriyi elde etme biimlerinin demografik zelliklere gre farklılık gsterip gstermedięi sorgulanmıř, sadece disipline gre bazı veri elde etme biimleri arasında istatistiksel aıdan anlamlı farklılıklar bulunmuřtur. Disipline gre, *niversite arařtırma ekibinden* ($\chi^2_{(4)}=23,677$; $p=0,000$) ve *bilinen bir kaynaktan* ($\chi^2_{(4)}=18,106$; $p=0,01$) veri elde etme biimleri istatistiksel aıdan anlamlıdır. Bu farklılık, dięer disiplinlere oranla sosyal bilimlerde *niversite arařtırma ekibinin* daha az, mhendislik alanında ise *bilinen bir kaynaktan* veri kullanımının daha ok tercih edilmesinden kaynaklanmaktadır.

Ziraat (%92,5; n=49) ve tıp (%91,8; n=90) alanında biraz daha yksek olmakla birlikte, tm disiplinler iin yeni *veri toplama/oluřturma* en ok tercih edilen seenektir.

Tıp alanında alıřanlar *bilinen bir kaynaktan* veri kullanmayı en az (%17,3; n=17) tercih ederken, *kendi arařtırma aęından* daha ok (%52, n=51) veri saęlamaktadırlar. Bunun aksine, *kendi arařtırma aęından* veri elde etme sosyal bilimlerde en azken (%29,30; n=24) *bilinen bir kaynaktan* veri kullanma mhendislik alanında en fazladır (%44,60; n=33). Dięer disiplinlere gre, fen alanındaki arařtırmacılar *niversite arařtırma ekibinden* veri saęlamayı daha ok (%48,4; n=45), sosyal bilimlerdekiler ise daha az (%14,6; n=12) tercih etmektedir (bkz. řekil 9). Bu durum, fen bilimlerinde ekip alıřmasının, sosyal bilimlerde ise bireysel alıřmanın genelde daha yaygın olmasından kaynaklanıyor olabilir.



Şekil 9. Disipline göre katılımcıların veri elde etme biçimleri (%). * $p < 0,05$

Bilgi arama davranışlarının disipline göre farklılık göstermesi gibi, araştırmacıların veri elde etme biçimlerinin de disipline göre farklılık göstermesi beklenen bir sonuçtur. Çalışmamızda disiplinler arası farklılık çoğunlukla sosyal bilimlerden kaynaklanmakla birlikte, Türkiye'deki araştırmacılar üzerine yapılan başka bir çalışmada, beşerî bilimlerden kaynaklı bir farklılık saptanmıştır. Söz konusu çalışmada, diğer disiplinlere oranla beşerî bilimlerde *bilinen birden fazla kaynaktan* veri sağlama daha çok, *yeni veri yaratma* ve *üniversitedeki kendi araştırma ekibinden* veri elde etme daha az tercih edilmiştir (Ünal ve Kurbanoğlu, 2018, s. 295-296).

Yaş (39'dan küçük, 39-49, 50-60 ve 60'tan büyük) ve deneyim (14'ten az, 14-23, 24-34, 34'ten fazla) arttıkça veriyi kendi *üniversite araştırma ekibinden* sağlayanların oranının da arttığı (yaş için sırasıyla: %30,1; %35; %41,1; %43,8 deneyim için sırasıyla: %28; %33,1; %44,4; %45,7) görülmektedir. Bu farklılık, yaşı fazla olan kişilerin meslekte daha fazla süredir faaliyet göstermeleri ve bunun sonucunda mesleki ağlarının daha geniş, bir (hatta birden fazla) araştırma ekibi içinde yer alma olasılıklarının daha fazla olması ile açıklanabilir.

Deneyime göre veri elde etme biçimleri incelendiğinde, 34 yıldan fazla deneyime sahip olanlar ile 24-34 yıl arasında deneyimi olan araştırmacılar (%45,7 n=21; %44,4 n=48) diğer gruplara göre (14'ten az %28 n=33; 14-23 yıl %33,1 n=45) *üniversite araştırma*

ekibinden daha çok veri sağlamaktadır. 24-34 yıl arasında deneyime sahip araştırmacıların yarısından fazlası (%56,5; n=61) *kendi araştırma ağından* veri elde ederken diğer gruplarda bu oran daha azdır (14'ten az %36,4 n=43; 14-23 yıl %40,40 n=55; 34'ten fazla %32,6 n=15).

Araştırmacıların diğer kaynaklardan veri kullanıp kullanmadığını belirlemek ve dış kaynaklı verinin tekrar kullanımını için ne ölçüde çaba harcadıklarını saptamak amacıyla bir soru sorulmuştur. Birden fazla seçeneğin işaretlenebildiği bu soruya üç araştırmacı yanıt vermemiştir. Verilen yanıtlar doğrultusunda, araştırmacıların yaklaşık üçte birinin (%35,8; n=145) *diğer kaynaklardan* veri kullanmadığı görülmektedir. Dış kaynaklı veri kullanımının disiplin temelli dağılımı incelendiğinde, tıp alanındaki araştırmacıların yarısından fazlası (%55,2; n=53) *diğer kaynaklardan veri kullanırken* diğer disiplinlerde bu oran daha fazladır (bkz. Tablo 9).

Tablo 9. Disipline göre dış kaynaklı veri kullanma durumu

| | Diğer kaynaklardan veri kullananlar | | Diğer kaynaklardan veri kullanmayanlar | | Toplam N |
|-----------------|-------------------------------------|------|--|------|----------|
| | N | % | N | % | |
| Fen | 66 | 71,0 | 27 | 29,0 | 93 |
| Mühendislik | 53 | 71,6 | 21 | 28,4 | 74 |
| Tıp | 53 | 55,2 | 43 | 44,8 | 96 |
| Sosyal Bilimler | 55 | 67,1 | 27 | 32,9 | 82 |
| Ziraat | 31 | 59,6 | 21 | 40,4 | 52 |
| Toplam* | 258 | | 139 | | 397 |

* 3 katılımcı soruya yanıt vermemiş, ikisi diğer kaynaklardan veri kullanan 8 araştırmacı disiplin belirtmemiştir.

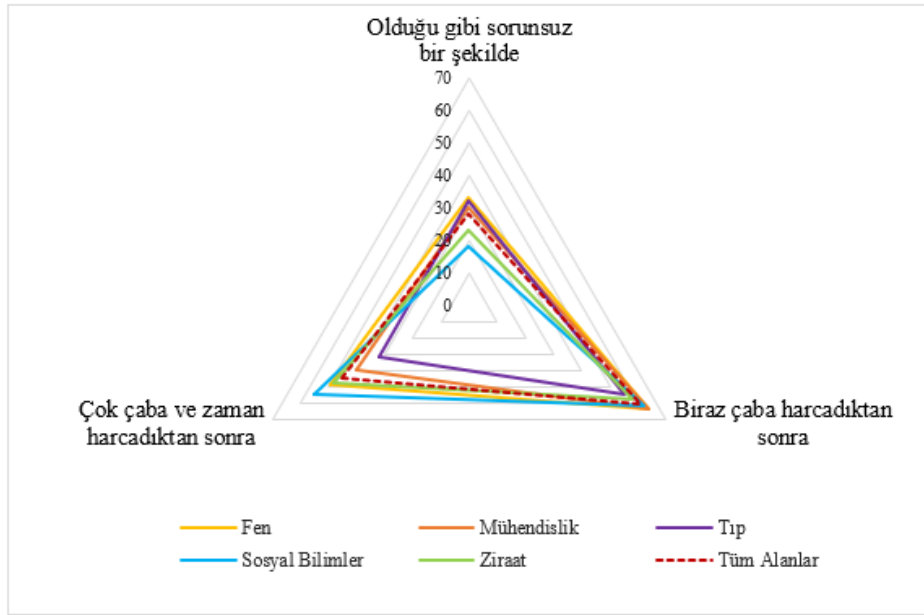
Yaş ile *diğer kaynaklardan* veri kullanım durumu arasında ters orantı bulunmaktadır. Yaş azaldıkça diğer kaynaklardan veri kullanım oranı artmaktadır (39'dan küçük %66,7; 39-49 %64,4; 50-60 %64 ve 60'tan büyük %54,8). Yaşı daha genç olanların henüz bir araştırma ekiplerinin ve yeterli birikimlerinin olmaması, hazır verinin hem zaman ve emekten tasarruf sağlaması hem de bir nevi güvence olması nedeniyle bu grupta diğer kaynaklardan veri kullanımını görece daha yüksek olabilir. Yaş ilerledikçe dış kaynaklardan veri kullanımının azalması da seçiciliğin artması ve başka çalışmalara olan güvenin azalmasıyla açıklanabilir. Bu grubun veri toplamada çalıştırabilecek bir ekibinin olması olasılığı da yeni veri oluşturmada avantaj olarak düşünülebilir.

Diğer kaynaklardan veri kullanan araştırmacıların (n=260) verinin yeniden kullanımı için ne ölçüde çaba harcadıkları (sorunsuz bir şekilde, biraz çaba ve/veya çok çaba) belirlenmeye çalışılmıştır. Dış kaynaklı veri kullanan araştırmacıların büyük çoğunluğu (%60,8; n=158) veriyi yeniden kullanırken *temizleme ve/veya değişiklikler için biraz çaba harcarken*, yaklaşık yarısı (%45; n=117) *çok çaba ve zaman* harcamaktadır. Dış kaynaklı veriyi *olduğu gibi sorunsuz şekilde* kullanan araştırmacıların oranı (%28,1; n=73) azdır.

Dış kaynaklı veri kullanan araştırmacıların tamamı en az iki seçenek, yaklaşık üçte biri (%33,1; n=86) ise en az üç seçenek işaretlemiştir. Başka bir deyişle, olduğu gibi sorunsuz bir şekilde dış kaynaklı veri kullanan bir araştırmacı aynı zamanda biraz çaba veya çok çaba harcayarak da dış veri kullandığını belirtmektedir. Bu durumda, her çalışma için ve farklı veri setleri için kendine özgü bir durum söz konusu olduğunu söylemek yanlış olmayacaktır. Diğer seçeneğini işaretleyerek katkıda bulunan araştırmacıların ifadeleri de aşağıda sunulmuş olup bunu doğrulamaktadır:

Çoğunlukla deneysel veya hesaplamalı yöntemlerle kendi hesaplamalarımıza ilişkin verileri kullandığımız için diğer kaynaklardan veri nadiren ve genellikle karşılaştırmalar için kullanılıyor; araştırma konusuna göre çok değişkenlik gösterebilmektedir; program aracılığıyla; sadece fikir edinmek amaçlı; diğer kaynaklardan sağladığım özellikle raw dataları (ham veri) kendi datalarımı analiz etmeden önce analiz akışını kontrol etmek için kullanırım; ilgili olan kısımları alıyorum; ham veriyi çeşitli algoritmalar ile işledikten sonra (örneğin uydu görüntüleri); ham verileri işleyerek projemde/araştırmamda kullanılacak yeni veriler/sonuçlar üretirim; diğer kaynaklarda bulunan verileri tartışma kısmında değerlendiririm; projede işbirliği yaptığım başka üniversitelerdeki öğretim üyelerinden uygun formda hazırlanmış olarak alıp kullanıyoruz

Disipline göre dış kaynaklı veri kullanım biçimleri Şekil 10'da sunulmuştur. İki araştırmacı disiplin belirtmemiştir. Mühendislik (%64,2; n=34), fen (%63,6; n=42) ve sosyal bilimlerde (%61,8; n=34) ağırlıklı olarak dış kaynaklı veriler *biraz çaba harcadıktan sonra* kullanılmaktadır.



Şekil 10. Disipline göre dış kaynaklı veri kullanım biçimleri (%)

Diğer kaynaklardan elde edilen verinin *çok çaba ve zaman harcadıktan sonra* kullanımı fen (%48,5 n=32), ziraat (%48,4 n=15) ve mühendislikte (%39,6 n=21) tıp alanına (%32,1 n=17) göre daha çok, sosyal bilimlere göre (%54,5 n=30) daha az işaretlenmiştir. Dış kaynaklı verinin *olduğu gibi sorunsuz bir şekilde* kullanımı en az (%18,2; n=10) sosyal bilimler alanındadır.

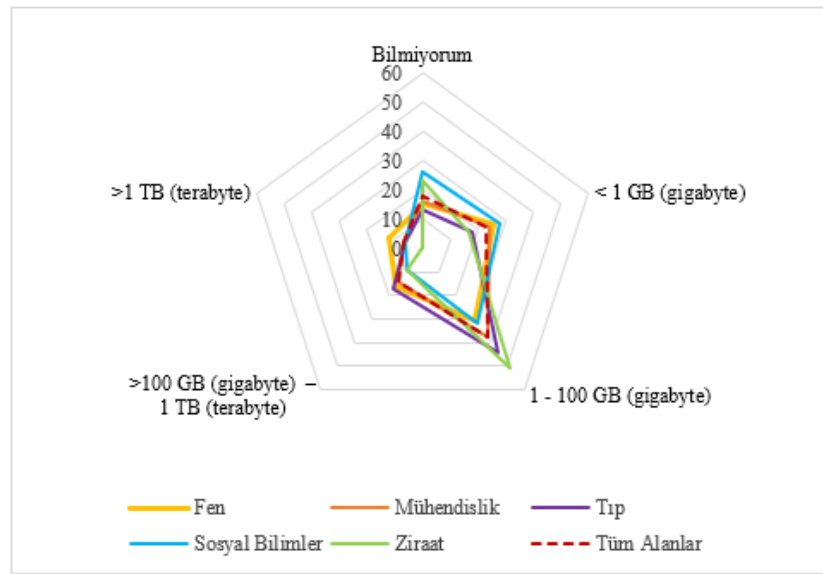
Dış kaynaklı veri kullanımının unvana göre dağılımı incelendiğinde ise, verinin yeniden kullanımı için *çok çaba ve zaman harcarım* seçeneğinin araştırma görevlileri tarafından daha çok (%59,4; n=19), Dr. Öğretim üyeleri tarafından daha az (%22,9; n=8) tercih edildiği görülmektedir.

4.2.3 Veri Miktarı

Araştırmalarda kullanılan veri miktarı, özellikle büyük verilerle çalışan araştırmacıların hem saklama/depolama alanlarının belirlenmesinde hem de yapılacak dokümantasyon işlemlerinin planlanmasında önemli rol oynamaktadır. Araştırmacıların kullandıkları veri miktarını belirlemek amacıyla yönlendirilen sorunun bulgularına göre, katılımcıların %17,6'sı (n=72) araştırmalarında kullanılan veri miktarını bilmemektedir. Bu durum, araştırmadan araştırmaya kullanılan veri miktarının değişmesinden kaynaklanabileceği

gibi, katılımcıların veri miktarı ölçü birimlerini bilememesinden de kaynaklanıyor olabilir. Katılımcıların çoğu (%38; n=155) *1 GB-100GB*, %22,8'i (n=93) *1 GB'den daha az*, %13,7'si (n=56) *1 TB ve daha az ama 100 GB'den fazla* veri kullanmaktadır. 1 TB'den fazla veri kullananların sayısı en düşüktür (%7,1; n=29). Kullanılan veri miktarı ile ilgili 3 araştırmacı, diğer seçeneğini işaretlemiş ve şu şekilde yanıt vermiştir:

Üç boyutlu uzayda zamanda ya da frekansta elektromanyetik alanlara ilişkin problemlerle uğraştığımız için verilerin her zaman saklanması mümkün olmuyor, çoğunlukla ara işlemlerde kullanılan veriler silindiği için net bir rakam vermek zor; değişiyor; sayısal ve görsel veriler, tahmini zor



Şekil 11. Disipline göre kullanılan veri miktarları (%)

Tüm disiplinlerde en çok tercih edilen veri miktarı aralığı *1 GB-100 GB* olmakla birlikte, istatistiksel açıdan anlamlı olmasa da disipline özgü küçük bazı farklılıklar gözlenmektedir (bkz. Şekil 11). Araştırmada kullandığı veri miktarını bilmeyenlerin oranı diğer disiplinlere göre sosyal bilimlerde daha çok (%25,6; n=21), tıp ve sağlık bilimlerinde daha azdır (%13,3; n=13). TÜBİTAK araştırmacıları üzerine yapılan başka bir çalışmada da benzer şekilde, üretilen ve kullanılan veri miktarının çoğunlukla *1 GB'den az* olduğu, veri miktarını *bilmeyenlerin* en çok sosyal bilimlerde, en az sağlık ve tıp bilimlerinde olduğu sonucuna ulaşılmıştır (Doğan ve diğerleri, 2021, s. 56).

1 TB'den fazla veri kullanımı fen alanında diğer disiplinlere oranla nispeten daha fazladır (%11,8; n=11). Türk akademisyenler üzerinde yapılan bir diğer çalışmada da genel

anlamda TB düzeyindeki verinin kullanım ve üretim oranının düşük olduğu ancak diğer bilim dallarına göre, fen bilimlerinde kullanım oranının (%11) bu çalışmayla benzer olduğu görülmektedir (Ünal ve Kurbanoglu, 2018, s. 294-95).

Yaş (39'dan küçük, 39-49, 50-60 ve 60'tan büyük) ve deneyim (14'ten az, 14-23, 24-34, 34'ten fazla) arttıkça kullandığı veri miktarını bilmeyenlerin oranı da artmaktadır (Yaş için sırasıyla: %14,6; %15,0; %22,3; %25,0 deneyim için sırasıyla: %14,4; %16,2; %17,6; %30,4. Bu farklılık, gençlerin günlük hayatta telefon kapasitesi, izledikleri video boyutu, internet kota takibi gibi nedenlerle kapasite ölçü birimlerine görece daha aşına olması ile açıklanabilir. Ayrıca, araştırmacıların büyük verilerle çalışmaması buna paralel olarak depolama sıkıntılarının olmaması nedeniyle veri miktarı hakkında net bir bilgiye sahip olmadıkları da düşünülebilir.

Genel anlamda kullanılan veri miktarlarına bakıldığında, araştırmacıların hacim olarak çok büyük verilerle çalışmadığı görülmektedir. Verinin değerinin giderek arttığı günümüzde, bilim insanlarının büyük hacimli verilerle çalışmaması düşündürücüdür. Bunun birçok nedeni olabileceği gibi, daha çok geleneksel yöntemlerle araştırma yapılması ve veriyi işlemek için gerekli olan teknolojiye sahip olmama veya söz konusu teknolojinin kullanımının yeteri kadar bilinmemesi de etkili olabilir. Veri sahipliği kadar önemli olan veri işleme, veriye katma değer kazandırmaktadır.

4.2.4 Depolama ve Saklama

Özellikle veri güvenliğinden kaynaklanan öneminden dolayı, verilerin depolandığı ve sakladığı ortamlar belirlenmeye çalışılmıştır (bkz. Tablo 10). Katılımcılar, araştırma verilerini araştırma sırasında en çok kişisel bilgisayarlarda (%94,4; n=385) ve ikinci olarak harici disk, flash bellek gibi taşınabilir aygıtlarda (%82,4; n=336) tutmaktadır. Literatürdeki çalışmalarda da kişisel bilgisayarlar en çok tercih edilen veri saklama ortamlarıdır (Chen ve Wu, 2017, s. 349; Ünal ve Kurbanoglu, 2018, s. 296; Doğan ve diğerleri, 2021, s. 57; Zencir, 2019 s. 150-151). Araştırmacıların büyük çoğunluğunun yeni veri oluşturduğu ve verilerin elektronik olduğu düşünüldüğünde, verilerin kullanılmasında gerekli donanım olan bilgisayarların saklama ortamı olarak en çok tercih edilmesi beklenen bir sonuçtur. Araştırmacıların çoğunun verilerini birden fazla

ortamda sakladıkları görülmektedir. Depolama ortamı olarak en az iki seçenek işaretleyenlerin oranı %92,9 (n=379) iken araştırmacıların yarısından fazlası (%57,6; n=235) en az üç seçenek işaretlemiştir.

Tablo 10. Veri saklama ortamları

| Veri saklama ortamları | N | % |
|---|-----|------|
| Kişisel bilgisayar(lar) | 385 | 94,4 |
| Taşınabilir aygıtlar (harici disk, flash bellek gibi) | 336 | 82,4 |
| Bulut hizmetleri (Dropbox, Google Drive, iCloud gibi) | 212 | 52,0 |
| E-posta hesabında | 149 | 36,5 |
| Bölüm/Kurum sunucusu | 35 | 8,6 |
| Diğer | 5 | 1,2 |

Araştırmacıların yaklaşık yarısı (%52; n=212), saklama ortamı olarak *bulut hizmetlerini* kullandıklarını belirtmişlerdir. Bir mekâna bağlı kalmaksızın veriye erişim imkânı sunan bulut hizmetlerinin kişisel bilgisayarlara ve taşınabilir aygıtlara oranla daha az tercih edilmesinin araştırma evrenindeki genç katılımcı oranının daha düşük ve buna bağlı olarak teknoloji yatkınlığının az olması ile ilgili olabileceği düşünülmektedir. Araştırmacıların yaklaşık %40'ı (n=149) araştırma verilerini *e-posta* yoluyla sakladıklarını belirtmişlerdir. Bu yöntem pratik bir uygulama olarak görülebilir ancak önerilmemektedir. Veri saklama ortamı olarak *kurum/bölüm sunucusu* kullanmak araştırmacılar tarafından pek tercih edilmemektedir (%8,6; n=35). Az tercih edilme nedeni böyle bir hizmetin üniversite tarafından sunulmaması veya kampüs dışından veriye erişim olanağının olmaması ile açıklanabilir. Diğer seçeneğini işaretleyen beş katılımcının ifadeleri şu şekildedir:

Araştırma grubu NAS depolama üniteleri; NAS- Synology kişisel sunucu (40 TB kapasite-yedekleme de sağlıyor); Projeme ait sunucuda depolarım. Aynı zamanda yedekleme ünitelerinde de bir kopyasını saklarım; distributed version control sites (bitbucket.org, github.com); E-posta hesabında bilgisayar veya flash içinde saklamadıklarımı tutuyorum.

Disipline, yaşa, deneyime ve unvana göre katılımcıların araştırma verilerini araştırma sırasında sakladıkları ortamlarda farklılık olup olmadığı sorgulanmış ve istatistiksel açıdan anlamlı farklılıklar içeren değerler Tablo 11'de sunulmuştur.

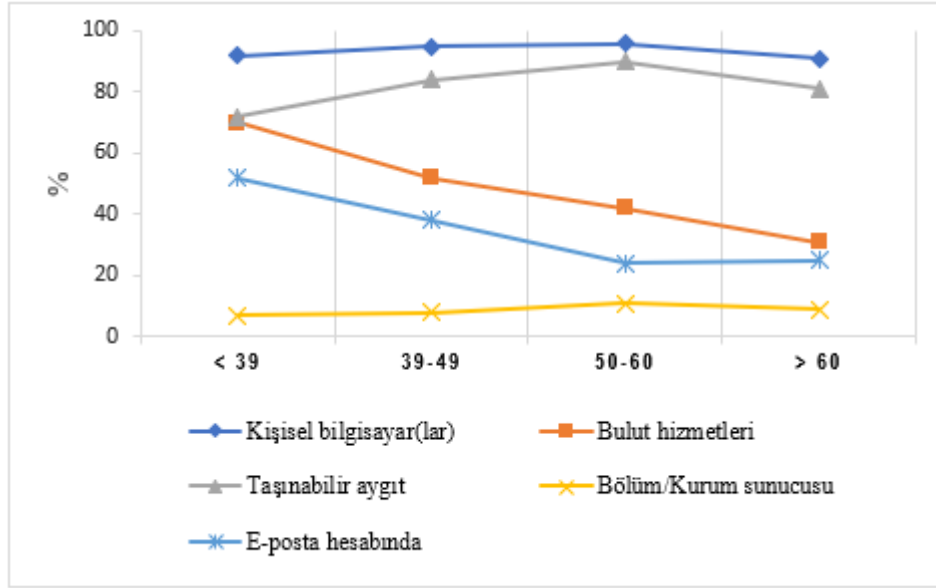
İstatistiksel açıdan anlamlı olmamakla birlikte, araştırma verilerinin saklandığı ortamlar disipline göre hem benzerlikler hem de farklılıklar göstermektedir. Tüm disiplinlerde veriler en çok *kişisel bilgisayarlarda* ve *taşınabilir aygıtlarda* depolanmaktadır. Sosyal bilimlerde *e-posta* kullanımı (%52,4; n=43) diğer disiplinlere oranla (ziraat %37,7; tıp %33,7; fen %32,3; mühendislik %27) daha fazladır. Ziraat alanında ise bulut hizmetlerinin kullanımı (%39,6; n=21) diğer disiplinlere oranla nispeten azdır (sosyal bilimler %58,5; tıp %54,1; fen %52,7; mühendislik %51,4).

Tablo 11. Demografik özelliklere göre veri saklama ortamları

| | Veri saklama ortamı | Değer |
|---------|---------------------|--------------------------------------|
| Yaş | Bulut | $(\chi^2_{(3)} = 23,269; p = 0,000)$ |
| | E-posta | $(\chi^2_{(3)} = 19,344; p = 0,000)$ |
| Deneyim | Bulut | $(\chi^2_{(3)} = 22,016; p = 0,000)$ |
| | E-posta | $(\chi^2_{(3)} = 18,985; p = 0,000)$ |
| Unvan | E-posta | $(\chi^2_{(4)} = 17,038; p = 0,002)$ |

Yaşa göre farklılıklar incelendiğinde, araştırma sırasında verilerin saklanmasında 38 yaş ve altı araştırmacılar, diğer araştırmacılara oranla *bulut hizmetlerini* (%69,9; n=72) ve *e-posta hesabını* (%51,5; n=53) daha çok kullanmaktadır. Bunun aksine, verileri saklama ortamı olarak 50-60 yaş aralığındaki araştırmacılar *taşınabilir aygıtları* daha çok (%90,2; n=101) ve *e-posta hesabını* (%24,1; n=27) daha az kullanmaktadırlar. Ayrıca yaş arttıkça (39'dan küçük, 39-49, 50-60 ve 60'tan büyük) *bulut hizmetlerinin* kullanım oranının azaldığı (sırasıyla: %69,9, %51,9, %42 ve %31,3) görülmektedir (bkz. Şekil 12).

Deneyim arttıkça saklama ortamı olarak *bulut hizmetleri* ve *e-posta* kullanım oranları azalmaktadır. Veriler yaş verileriyle uyumludur. Deneyim ve yaş genelde doğru orantılı arttığı için bu beklenen bir durumdur. 14 yıldan az deneyime sahip olanlar *bulut hizmetlerini* daha çok (%62,7; n=74), 34 yıldan daha fazla deneyime sahip olanlar daha az (%30,4; n=14) kullanmaktadır. 24-34 yıl deneyime sahip olanların büyük çoğunluğu (%90,7; n=98) *taşınabilir aygıtlarda* verilerini saklamaktadır. Yine bu grup, diğer gruplara göre *bölüm/kurum sunucusunu* (%13,9; n=15) daha çok kullanmaktadırlar.



Şekil 12. Yaşa göre verilerin saklanma ortamları

Unvana göre veri saklama ortamı olarak *taşınabilir aygıtların* kullanımı Profesörler tarafından en çok (%86, n=162) araştırma görevlileri tarafından en az işaretlenmiştir (%65, n=31). Literatürdeki benzer bir çalışmada, veri saklama ortamı olarak bulut kullanımı unvana göre istatistiksel açıdan farklılık göstermekte ve Profesör unvanına sahip araştırmacılar bulut hizmetlerini daha az kullanmaktadır (Doğan ve diğerleri, 2021, s. 57). Bu durum, taşınabilir aygıtların görece daha eski bir teknoloji olması nedeniyle profesörlerin alışkanlıklarını devam ettirmesi ve yeni teknolojik gelişmelerin sonucu olan bulut hizmetlerine gençlerin daha yatkın olmasıyla açıklanabilir.

4.2.5 Yedekleme

Veri güvenliği açısından önemli olan bir diğer konu da yedekleme olup araştırmacılara verilerini ne sıklıkta yedekledikleri sorulmuştur. Bulgulara göre, verilerin çoğunlukla *değişiklik yapıldığı her an yedeklendiği* (%36,3; n=148), verilerini *yedeklemeyenlerin* oranının ise oldukça düşük (%5,9; n=24) olduğu görülmektedir. Verileri yedekleme sıklığına ilişkin soruda işaretlenen diğer seçenekler sırasıyla; *yılda birkaç kez* (%27,2; n=111), *aylık* (%11,8; n=48), *haftalık* (%5,6; n=23), *yıllık* (%4,2; n=17), *birkaç yılda bir* (%2,2; n=9) şeklindedir. Diğer seçeneğini işaretleyen araştırmacılar ise; sistematik yedekleme yapmadıklarını, verilerini akıllarına geldikçe yedeklediklerini, önemli ve

üretilmesi zor olanları her an, daha kolay üretilebilir ve büyük boyutlu verileri daha nadir yedeklediklerini belirtmişlerdir.

Literatürdeki çalışmalara göre araştırmacıların veri yedekleme sıklığı bu çalışmada farklılık göstermektedir. Türkiye'deki araştırmacılar üzerine 2017 yılında yapılan bir çalışmada, düzenli yedekleme yapan araştırmacıların yaklaşık yarısı haftada bir verilerini yedeklerken, anlık yedekleme yapanların oranı (%26,7) daha azdır (Aydınoglu ve diğerleri, 2017, s. 280). Kaliforniya Politeknik Eyalet Üniversitesi'ne (California Polytechnic State University) mensup araştırmacılar üzerine yapılan çalışmanın sonuçlarına göre, haftalık yedekleme yapanların oranı (%27), herhangi bir değişiklik yapıldığında yedekleme yapanların oranından (%17) fazladır (Scaramozzino ve diğerleri, 2012, s. 358). Veri yedekleme sıklığının hem kişisel bir tercih hem de verinin niteliğine bağlı olması nedeniyle mevcut farklılıkların normal olduğu düşünülmektedir. Yaşanan teknolojik gelişmelerle birlikte, depolama biçimlerinin değişmesi ve olanakların artması da alışkanlıkların değişmesinde etkilidir.

Araştırma/proje bittikten sonra veri saklama süreleri incelendiğinde, araştırmacıların büyük çoğunluğunun (%73,8; n=301) araştırma verilerini *daima sakladıkları*, verilerini *1 yıldan az* (%0,7; n=3) ve *1-2 yıl* (%2,9; n=12) *saklayanların* oranının oldukça düşük olduğu, bir kişinin (%0,2) ise verilerini *hiç saklamadığı* saptanmıştır. Verilerini *2 yıldan fazla 5 yıldan az* saklayanlar (%11,8; n=48) ile *5 yıldan 10 yıla dek* saklayanların (%10,5; n=43) oranı birbirine yakındır. Katılımcıların büyük veri ile çalışmadıkları göz önüne alındığında, herhangi bir depolama alanı sıkıntısı yaşamadıkları ve verilerini sürekli saklamak için bir engelleri olmadığı söylenebilir. Literatürdeki diğer çalışmaların bulguları da araştırmacıların verilerini çoğunlukla süresiz olarak sakladığını göstermektedir (Zencir, 2019, s. 152; Buys ve Shaw, 2015, s. 12).

4.2.6 Uzun Süreli Koruma

Katılımcıların neredeyse tamamına yakını (%94,6; n=386) uzun süreli koruma için verilerini *kendi cihazlarında sakladıklarını*, yaklaşık yarısı ise (%47,1; n=192) *bulut hizmetlerini* de kullandıklarını belirtmişlerdir. Araştırmacıların çok azı (%1,5; n=6) verilerini *uzun süreli korumadıklarını* belirtmiştir. Uzun süreli korumada; *kurumsal*

arşivler (%7,8; n=32), konusal/kurum dışı arşivler (%2,9; n=12), açık erişimli bir veri havuzu (%1,5; n=6), kamuya açık bir web sitesi (%0,7; n=3) ve ticari veri havuzu (%0,1; n=1) kullanım oranı oldukça düşüktür³. Diğer seçeneğinde görüş bildiren araştırmacıların ifadeleri: *distributed version control sites* (dağıtık versiyon kontrol sistemleri – *bitbucket.org, github.com*); *projenin diğer araştırmacıları ile aynı anda kişisel bilgisayar ve harici depolamada; çok önemli kısımların çıktılarını saklıyorum* şeklindedir.

Bu görüşler ve uzun süreli korumada en çok kişisel bilgisayarların tercih edilmesi, uzun süreli korumanın çoğunlukla depolama ile karıştırıldığı şeklinde yorumlanabilir. Veri arşivlerinin az tercih edilmesi de araştırmacıların erişilebilirlik, anlaşılabilirlik, doğrulanabilirlik, hesap verilebilirlik gibi bir amaçları olmadığı şeklinde açıklanabileceği gibi yeterli sayıda veri arşivinin olmamasıyla da ilgili olabilir.

Araştırma verilerinin uzun süreli koruma ortamlarından bazıları yaşa göre istatistiksel olarak anlamlı bir farklılık göstermektedir. Depolamada olduğu gibi uzun süreli korumada da yaş arttıkça (39'dan küçük, 39-49, 50-60 ve 60'tan büyük) bulut hizmetlerinden yararlanma oranı (sırasıyla: %65, %45,6, %37,5 ve %31,3) azalmaktadır ($\chi^2_{(3)} = 20,822$; $p=0,000$). Ayrıca basılı olarak saklama biçimi (sırasıyla: %11,7, %20,6, %25,9 ve %40,6) artmaktadır ($\chi^2_{(3)} = 14,265$; $p=0,003$). İstatistiksel açıdan anlamlı olmamakla birlikte, uzun süreli koruma biçiminde, kurumsal arşivlerin kullanım oranı ile yaş arasında doğru orantı (artan sıralamayla: %5,8, %8,1, %8,9 ve %9,4) bulunmaktadır. Yaş arttıkça uzun süreli korumada kurumsal arşiv kullananların oranı da artmaktadır.

Ülkemizde yapılan diğer çalışmalarda da araştırma verilerin depolanması amacıyla kurumsal arşiv ve kurum dışı arşivlerin kullanımını düşük bulunmuştur (Aydinoğlu ve diğerleri, 2017, s. 279; Doğan ve diğerleri, 2021, s. 57; Ünal ve Kurbanoğlu, 2018, s. 296). Ülkemizde kurumsal arşivlerin genellikle YÖK'ün talebi üzerine kurulması, gelişme aşamasında olması ve çoğunlukla araştırma verilerinin değil araştırma çıktılarının kütüphanelerin çabalarıyla kurumsal akademik arşivlere yüklenmesi,

³Veri arşivleri standart üst veriler kullanılarak verilerin yayımlandığı herkese açık olan paylaşım ortamlarıdır. Veri havuzları ise belirli konularda standartlara sahip ve belirli araştırma gruplarına hitap eden depolardır.

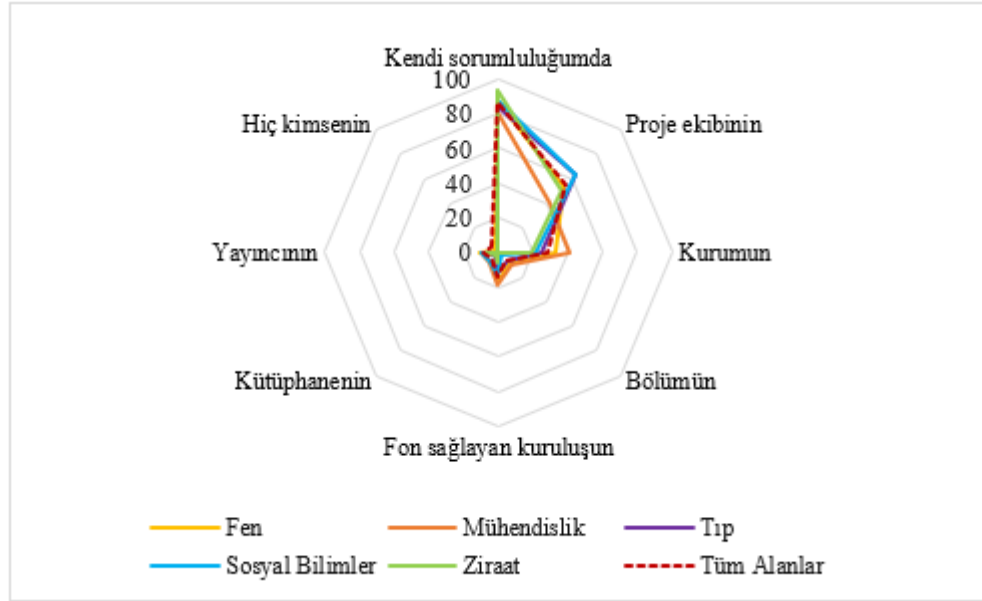
araştırmacıların henüz bu ortamları bilmemeleri, benimsememeleri uzun süreli koruma ortamı olarak kurumsal arşivlerin tercih edilmeme sebepleri olabilir.

Araştırmacılara uzun süreli korumanın kimin sorumluluğunda olduğu sorulmuş ve tüm disiplinlerde en çok (%86,5; n=353) *kendi sorumluluğunda* seçeneği işaretlenmiştir. Birden fazla seçeneğin işaretlenebildiği bu soruda, katılımcılar uzun süreli korumada birden fazla sorumlu olduğunu düşünmektedir. Katılımcıların yarısından fazlası (%54,9; n=224) sorumluluğun *proje ekibinde*, %27,7'si (n=113) *kurum/üniversitenin* ve %14,5'i (n=59) ise *fon sağlayıcı kuruluşun* sorumluluğunda olduğunu belirtmiştir. Sorumluluğun *yayıncıda* (%8,8; n=36), *bölümde* (%6,4; n=26) ve *kütüphanede* (%4,9; n=20) olduğunu düşünenlerin oranı oldukça azdır. Sadece iki kişi (%0,5), uzun süreli korumada *hiç kimsenin* sorumlu olmadığını düşünmektedir.

Ülkemizde kütüphanelerin çoğunlukla henüz araştırma verilerine yönelik bir misyonunun ve alt yapılarının olmaması nedeniyle kütüphanelerin uzun süreli korumada sorumlu olarak görülmediği düşünülmektedir.

Uzun süreli korumada sorumlu olduğu düşünülenlerin disipline göre dağılımı Şekil 13'te sunulmuştur. Diğer disiplinlere oranla, mühendislik alanındaki araştırmacılar uzun süreli korumada sorumlu olarak -yüzdeleri aynı olmakla birlikte (%40,5; n=30)- *proje ekibi* seçeneğini diğer disiplinlerdekilere göre daha az, *kurum/üniversite* seçeneğini ise daha fazla işaretlemişlerdir. Tıp (%63,3; n=62) ve sosyal bilimler alanlarında (%63,4; n=52) sorumluluğun *proje ekibinde* olduğunu düşünenler çoğunluktadır.

Diğer seçeneğini işaretleyen katılımcıların ifadeleri şu şekildedir: *ham veriler proje yürütücüsü; yayınlar/ ürünler fon sağlayan kuruluş tarafından korunmalıdır; kurum bence bu tür olanakların sağlanması ile ilişkili destek sağlanmalı, örneğin web üzerinde tutulan datanın güvenliği gibi; geçen zaman içinde herhangi bir kurumun koruma/saklama hizmetine güvenmiyorum; proje yürütücü ve ekibin sorumluluğunda ama kurumların maddi ve iş gücü desteği ile.*



Şekil 13. Uzun süreli korumada sorumluluk (%)

Literatürdeki bir çalışmada da uzun süreli korumada sorumluluğun kendisinde olduğunu düşünenler çoğunlukta (%95,1) olup bu çalışmayla benzerlik göstermektedir. Ancak en çok tercih edilen ikinci (hiç kimse - %11) ve üçüncü seçeneklerde (kütüphane/arşiv ve kampüs bilgi teknolojileri - %8,5) farklılıklar vardır (Scaramozzino ve diğerleri, 2012, s. 356). Başka bir çalışmada, proje sonrasında verilerin uzun süreli saklanması sorumluluğunun kendisi veya diğer proje/ekip üyelerinde olduğunu düşünenler yine çoğunlukta (%81,2) (Kennan ve Markauskaite, 2015, s. 85). Türkiye'deki araştırmacılar üzerine yapılan bir çalışmada ise, araştırma verilerinin hem depolanması ve erişime sunulmasındaki sorumluluk birimi olarak hem de uzun süreli depolama yeri olarak en çok üniversiteler tercih edilmiştir (Ünal ve Kurbanoglu, 2018, s. 304-305).

4.2.7 Veriyi Tanımlama: Üst Veri Kullanımı ve Dokümantasyon İşlemleri

Veriyi tanımlamanın gelecekte verinin ve bağlamının anlaşılabilirliği açısından taşıdığı önemden dolayı, araştırmacıların üst veri kullanım durumları ve veriye yönelik yaptıkları dokümantasyon işlemleri belirlenmeye çalışılmıştır.

Verilerin aranabilir, erişilebilir, keşfedilebilir ve tekrar kullanılabilir olması için üst veri vazgeçilmez bir unsurdur. Araştırmacıların üst veri kullanım durumları incelendiğinde,

yarısından fazlası (%59,8; n=244) *üst veri bilgilerini kaydederken*; %19,1'i (n=78) *gerekli bulmadıkları*, %13'ü (n=53) *zaman bulamadıkları* ve %6,9'u (n=28) *teknik desteğe ihtiyaç duymaları* nedeniyle üst veri bilgilerini kaydetmediklerini belirtmişlerdir. Diğer seçeneğini işaretleyen beş araştırmacı, üst verilerini ayrıca laboratuvar veya araştırma defterine yazdığını, üst veri bilgilerini düzenli olarak oluşturmadığını, mümkün olduğunca, yapabildiği ölçüde veya fırsat buldukça kaydedebildiğini belirttiği için üst veri bilgilerini kaydedenler kategorisinde değerlendirilmiştir. Bu seçenekte görüş bildiren diğer katılımcılar ise (%1,5; n=6) üst veri bilgilerine önem verdiğini ancak teknik desteğe ihtiyacı olduğunu, verileri depolarken bu bilgileri içerecek şekilde dosyalama yaptıklarını, kaydettikleri dosya adından anladıklarını ve iki kişi ise yayın aşamasında bu bilgilere yer verdiğini ifade etmişlerdir. Ankara Üniversitesi proje yürütücüleri üzerine yapılan bir çalışmada, üst veri oluşturmama nedenleri sırasıyla zaman bulamamak (%34,3), teknik desteğe ihtiyaç duymak (%28,6) ve gerekli bulmamak (%27,1) olarak ifade edilmiştir (Zencir, 2019, s. 154).

Tablo 12. Kullanılan üst veri standartları

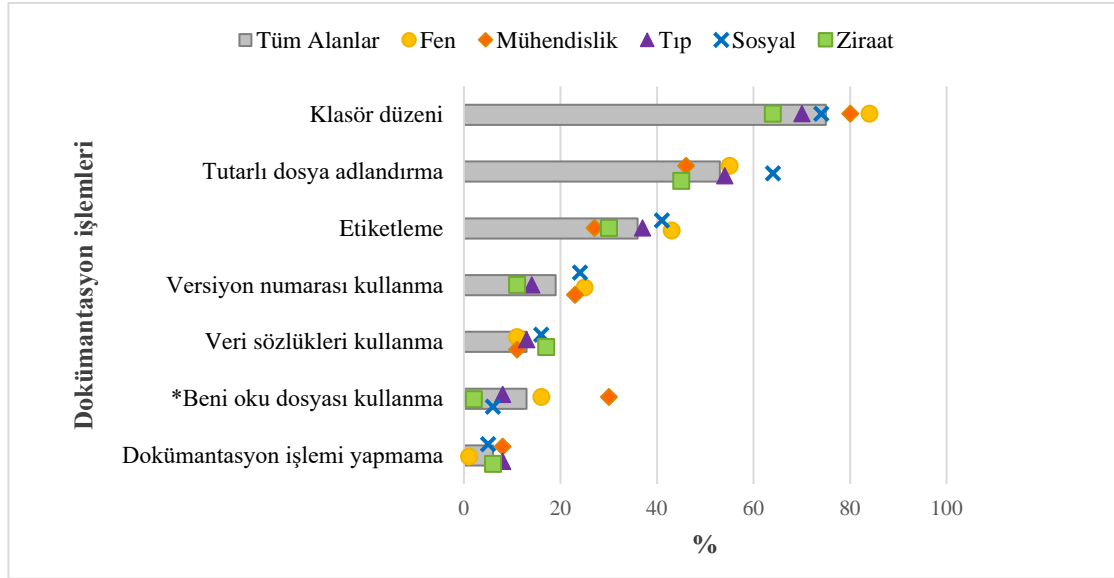
| | N | Üst veri standardı kullananlar içinde % | Üst veri kullananlar içinde % | Tüm katılımcılar % |
|--|----|---|-------------------------------|--------------------|
| Laboratuvarımda standartlaştırılmış üst veri | 25 | 65,8 | 10,2 | 6,1 |
| ISO (International Standards Organization) | 5 | 13,2 | 2,0 | 1,2 |
| NetCDF (Network Common Data Form) | 4 | 10,5 | 1,6 | 1,0 |
| Diğer | 4 | 10,5 | 1,6 | 1,0 |
| FITS (Flexible Image Transport System) | 3 | 7,9 | 1,2 | 0,7 |
| OGIS (Open GIS) | 3 | 7,9 | 1,2 | 0,7 |
| CIF (Crystallographic Information Framework) | 2 | 5,3 | 0,8 | 0,5 |
| AVM (Astronomy Visualization Metadata Standard) | 1 | 2,6 | 0,4 | 0,2 |
| CF (Climate and Forecast) | 1 | 2,6 | 0,4 | 0,2 |
| DIF (Directory Interchange Format) | 1 | 2,6 | 0,4 | 0,2 |
| EML (Ecological Metadata Language) | 1 | 2,6 | 0,4 | 0,2 |
| FGDC Metadata Standard (Federal Geographic Data Committee) | 1 | 2,6 | 0,4 | 0,2 |
| OME-XML (Open Microscopy Environment-XML) | 1 | 2,6 | 0,4 | 0,2 |

Araştırmacılara üst veriyi kaydederken herhangi bir üst veri standardı kullanıp kullanmadıkları sorulmuştur. Eldeki verilere göre, üst veri bilgilerini kaydeden araştırmacıların %15,6'sı (tüm katılımcıların %9,3'ü, n=38) üst veri standardı kullanmaktadır. Araştırmacılar tarafından tercih edilen üst veri standartları Tablo 12'de

sunulmuş olup en çok *laboratuvarımda standartlaştırılmış üst veri kullanıyorum* seçeneğinin işaretlendiği görülmüştür. Diğer seçeneğini işaretleyen katılımcıların açıklamaları şu şekildedir: *göz-izleme cihazı yazılımının hazır olarak sunduğu formatı kullanıyorum; ENDF; kendim oluşturduğum bir tablolama sistemi kullanıyorum; projeler, hem konu olarak hem de ekip olarak çok farklılık gösterebilmektedir. Bu nedenle proje süreci içinde ve önceki deneyimlerime göre kişisel bir yöntem geliştiriyorum.*

Diğer çalışmalar incelendiğinde, araştırmacıların çoğunlukla üst veri kavramına aşina olmadığı yönünde bulgular (Akers, 2013, s. 12; Zencir, 2019, s. 152) olduğu gibi, üst veri kullanımının üçte birin üzerinde (%36,4) olduğu yönünde bulgular da vardır (Aydınoglu ve diğerleri, 2017, s. 278). Üst veri kullanan araştırmacıların da çoğunlukla *çalıştığı laboratuvarında kullanılan ve/veya kendi oluşturduğu üst veri standardını* kullandığı görülmektedir (Aydınoglu ve diğerleri, 2017, s. 278; Whitmire, Boock ve Sutton, 2015, s. 390; Tenopir ve diğerleri, 2011, s. 6; Zencir, 2019, s. 154). ISO üst veri standardının en çok tercih edilen ikinci seçenek olması, 2017 yılında ülkemizdeki akademisyenler üzerinde yapılan çalışmanın bulgularıyla benzerlik göstermektedir (Aydınoglu ve diğerleri, 2017, s. 278). Bu çalışmada, üst veri kullanımının daha yüksek bulunması geçen zamanda yayıncı politikalarının uygulanması, AVP zorunluluğu ve buna bağlı olarak da üst veri farkındalığının artması ile açıklanabilir.

Araştırmalara ait verileri anlamlandırmak ve/veya aranılan dosyayı kolayca bulmak için yapılan dokümantasyon işlemlerine yönelik soruya verilen yanıtlar incelendiğinde, büyük çoğunluğun (%94,1; n=384) dokümantasyon işlemi yaptığı görülmektedir. Fen alanında (%83,9; n=78) en çok, ziraat alanında en az (%64,2; n=34) olmak üzere, katılımcılar en çok (%75) “*çalışmama ait klasörler belirli bir düzen içerisindedir (projeye göre, tarihe göre, veri türüne göre vb)*” seçeneğini işaretlemişlerdir.



Şekil 14. Yapılan dokümantasyon işlemleri (%). * $p < 0,05$

Beni oku dosyası kullanımı disiplinlere göre farklılık göstermekte ve bu fark istatistiksel açıdan anlamlı görülmektedir ($\chi^2_{(4)} = 30,871$; $p = 0,000$). *Beni oku dosyası* kullanımı en çok (%29,7; $n = 22$) mühendislikte iken, en az (%1,9; $n = 1$) ziraat alanındadır. Diğer disiplinlere oranla, fen alanında *aradığım dosyayı kolay bulmak için etiketleme yaparım* (%43; $n = 40$) ve *dosya adlandırma sırasında sürüm (versiyon) numarası kullanırım* (%24,7; $n = 23$) seçenekleri; sosyal bilimler alanında ise *dosya adlandırırken belirli bir kurala uyarak tutarlı adlandırma yaparım* (%64,2; $n = 52$) seçeneği daha fazla tercih edilmiştir (bkz. Şekil 14).

Kullanılan veri miktarına bağlı olarak yapılan dokümantasyon işlemlerinde bazı farklılıklar görülmektedir. Veri miktarı çok olan araştırmacılarda (1 TB'den fazla veri kullananlar - %24,1'i $n = 7$; >100 GB-1 TB veri kullananlar - %21,4'ü $n = 12$) *beni oku dosyası* kullanımı daha fazladır ($\chi^2_{(4)} = 11,488$; $p = 0,022$). Veri miktarı ile dosya etiketleme arasındaki farklılık istatistiksel açıdan anlamlıdır ($\chi^2_{(4)} = 17,549$; $p = 0,002$). Diğer veri miktarlarına oranla, 1 GB'den az verisi olanlar *aradıkları dosyayı kolay bulmak için etiketlemeyi* daha az (%19,4; $n = 18$), >100 GB-1 TB verisi olanlar daha çok (%48,2; $n = 27$) yapmaktadır.

4.2.8 Veri Yönetim Planı

Araştırma sürecinde ve sonrasında araştırma verilerinin nasıl yönetileceğine ilişkin bilgiler içeren Veri Yönetim Planı (VYP) ile ilgili bulgulara göre, araştırmacıların sadece %23,3'ünün (n=95) bir plan hazırladıkları saptanmıştır. Veri yönetim planı hazırladığını belirten dört araştırmacı mensup olduğu disiplini belirtmemiştir. Sosyal bilimler alanında VYP hazırladıklarını belirten araştırmacıların oranı (%31,7; n=26) diğer disiplinlere göre (fen %22,6 n=21; ziraat %22,6 n=12; tıp %21,4 n=21; mühendislik %14,9 n=11) daha fazladır.

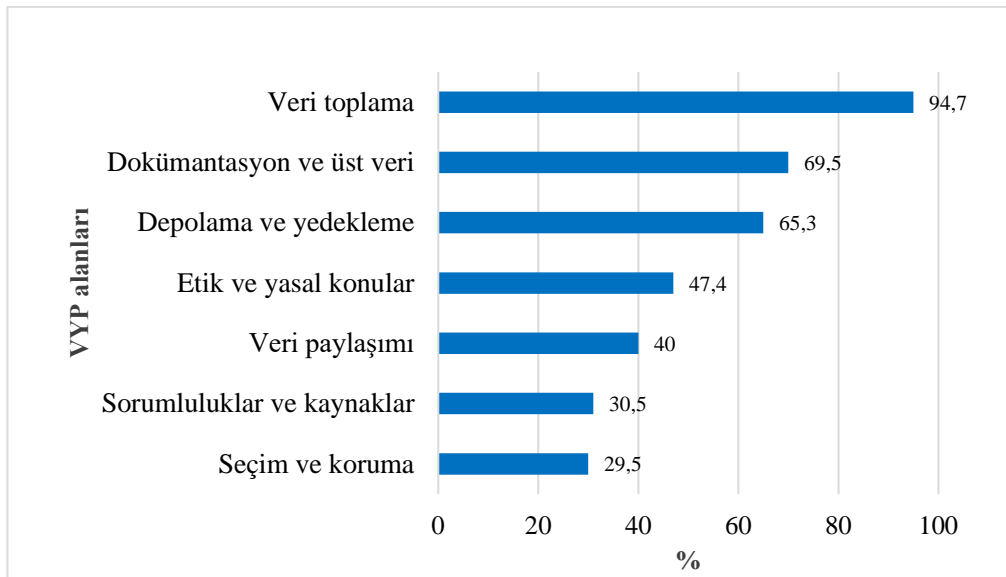
Türkiye'deki araştırmacılar üzerine yapılan bir çalışmada, veri yönetim planını bilenler ve/veya bir plan hazırlayanların oranı %9'dur (Doğan ve diğerleri, 2021, s. 57). Yine Türkiye'deki araştırmacılar üzerine yapılan başka bir çalışmaya göre, katılımcıların %13'ü daha önceki çalışmaları için, %16'sı ise mevcut projeleri için bir VYP hazırlamıştır (Ünal ve Kurbanoglu, 2018, s. 299). Ankara Üniversitesi mensubu akademisyenler üzerinde yapılan çalışmada da benzer bir oran (%16,5) saptanmıştır (Zencir, 2019, s. 146). Bu çalışmada, VYP hazırlayanların oranının diğer çalışmalara göre nispeten fazla çıkması TÜBİTAK projesinde yer alan bu araştırmacıların yurt dışı bağlantılı bir araştırma ekiplerinin de olabileceği ve geçen zamanda dünya genelinde AVY politikalarının uygulanmasıyla açıklanabilir. TÜBİTAK Açık Bilim Politikası 14.03.2019 tarihinde kabul edilmiş ve bu politika belgesiyle proje desteği başvuru sürecinde araştırma verileri yönetim planının hazırlanması önerilmiştir. Daha sonra, TÜBİTAK web sayfasında yayımlanan veri yönetim planı bilgi notuyla (https://tubitak.gov.tr/sites/default/files/20689/veri_yonetim_plani_bilgi_notu.pdf), proje başvurularında veri yönetim planının diğer başvuru belgeleriyle birlikte sisteme yüklenmesi gerektiği belirtilmiştir. Bu araştırma için anketin uygulandığı tarihlerde projelerde VYP hazırlanması yönünde bir zorunluluk olmadığı için toplanan verilerle saptanan mevcut durumun değişmiş olacağı göz önüne alınmalıdır.

Araştırmacıların VYP hazırlama nedenleri incelendiğinde, araştırmacılar en çok sırasıyla *kişisel tercih* (%83,2; n=79), *araştırma ekibinin tercihi* (%31,6; n=30) ve *fön sağlayıcı kurumun zorunlu tutması* (%21,1; n=20) nedeniyle plan hazırladıklarını bildirmişlerdir. *Fön sağlayıcı kurumun tavsiye etmesi* (%9,5; n=9), *çalışılan kurumun zorunlu tutması*

(%3,2; n=3) ve *tavsiye etmesi* (%3,2; n=3) nedeniyle VYP hazırlayanlar olmakla birlikte oranları azdır. Diğer seçeneğini işaretleyip katkıda bulunan araştırmacıların ifadeleri şu şekildedir: *Etik Kurul Şartları, KVKK, Aldığım bir eğitimde hazırlamam gerektiğini öğrendim.*

Araştırma verilerinin yönetiminin bir bütün olarak düşünülmesi ve politika belgelerinde de hem plan hem de üst veriden bahsedilmesi nedeniyle, VYP hazırlama durumu ile üst veri bilgilerini kaydetme arasında ilişki olup olmadığı sorgulanmıştır. Yapılan ki-kare analizine göre istatistiksel olarak anlamlı bir farklılık bulunmuştur ($\chi^2_{(1)} = 12,006$; $p = 0,001$). Veri yönetimi planı hazırlayanların büyük çoğunluğu (%73,7; n=70) aynı zamanda üst veri bilgilerini kaydederken, üst veri bilgilerini kaydetmeyenlerin büyük çoğunluğu (%85,3; n=145) VYP hazırlamamaktadır.

Başka bir çalışmada da VYP kullanım durumunun ülkelere göre farklılık gösterdiği, Birleşik Krallıkta VYP kullanımının daha fazla olduğu görülmektedir (Ünal, Chowdhury, Kurbanoglu, Boustany ve Walton, 2019, s. 21). Bu da politika belgelerinin önemini ortaya koymaktadır. Politika belgelerinde VYP'nin zorunlu tutulmasının ve düzenlenen eğitimlerde planın öneminden bahsedilmesinin plan hazırlamada olumlu sonuçlar doğurduğu söylenebilir. Veri yönetim planıyla veri dokümantasyonu adım adım düşünülmemekte ve bu plan üst veri kullanımına zemin hazırlamaktadır.



Şekil 15. VYP hazırlayan araştırmacıların kullandıkları alanlar (%)

Herhangi bir VYP hazırladığını belirten katılımcılara, hazırladıkları plan içinde hangi alanların olduğu sorulmuştur. Birden fazla seçeneğin işaretlenebildiği bu soruya verilen cevaplara göre araştırmacıların planlarında en çok (VYP hazırlayanların %94,7'si, tüm katılımcıların %22,1'i n=90) veri türü, formatı, büyüklüğü, dosya yapısı, isimlendirme gibi bilgileri içeren *veri toplama* alanına; en az ise *seçim ve koruma* (VYP hazırlayanların %29,5'i, tüm katılımcıların %6,9'u n=28) alanına yer verdikleri görülmektedir (bkz. Şekil 15).

Katılımcıların büyük çoğunluğunun (%92,4; n=377) VYP hazırlamada kolaylık sağlayan DMPTool, DMPOnline gibi bazı araçları bilmediği saptanmıştır. Bu araçları bilen ama kullanmayanların oranı (%6,1; n=25) bilen ve kullananlara (%1,5; n=6) göre daha yüksektir. Veri yönetim planı hazırlayanların sayısının (n=95) düşük olması nedeniyle bu araçların yeterince bilinmediği düşünülmektedir.

Araştırmacıların yaklaşık üçte biri (%35,3; n=144) veri yönetim planı hazırlamanın araştırma verilerinin yönetiminde yardımcı olduğunu düşünürken, büyük çoğunluğun (%62,7 n=256) bu konu hakkında fikri yoktur. Veri yönetim planının araştırmacılara yardımcı olmadığını düşünenlerin oranı (%2, n=8) ise çok düşüktür. Bu soruya yönelik plan hazırlayan ve hazırlamayan araştırmacılardan alınan yanıtlar arasında istatistiksel açıdan anlamlı bir farklılık vardır ($\chi^2_{(2)} = 99,749$; $p=0,000$). VYP hazırlayan araştırmacıların büyük çoğunluğu (%77,9 n=74) araştırma planının yardımcı bir doküman olduğunu düşünmekteyken, plan hazırlamayan araştırmacılar büyük çoğunlukla (%75,7 n=237) bu konu hakkında fikir sahibi değildir.

TÜBİTAK Açık Bilim Politika Belgesine Yönelik Farkındalık Durumu

14.03.2019 tarihinde yürürlüğe giren TÜBİTAK Açık Bilim Politikası hakkında farkındalıklarını belirlemek amacıyla katılımcılara bir soru yöneltilmiştir. Araştırmacıların yarıdan fazlasının (%62; n=253) bu politika belgesinin *farkında olmadığı* ortaya çıkmıştır. Araştırmacıların sadece %10,3'ü (n=42) *farkındayım ve okudum* seçeneğini işaretlerken, *farkındayım ama okumadım* diyenlerin oranı %27,7'dir (n=113).

Tablo 13. TÜBİTAK Açık Bilim Politikası farkındalık durumu

| Kurum Adı | Farkındayım ve okudum | | Farkındayım ama okumadım | | Farkında değilim | | Toplam |
|------------------------|-----------------------|------|--------------------------|------|------------------|------|--------|
| | N | % | N | % | N | % | |
| Ankara Üniversitesi | 4 | 5,2 | 20 | 26,0 | 53 | 68,8 | 77 |
| Hacettepe Üniversitesi | 8 | 10,3 | 20 | 25,6 | 50 | 64,1 | 78 |
| İTÜ | 2 | 3,9 | 25 | 49,0 | 24 | 47,1 | 51 |
| İstanbul Üniversitesi | 8 | 14,8 | 11 | 20,4 | 35 | 64,8 | 54 |
| ODTÜ | 5 | 18,5 | 9 | 33,3 | 13 | 48,1 | 27 |
| Diğer | 15 | 14,7 | 24 | 23,5 | 63 | 61,8 | 102 |
| Belirtmeyenler | 0 | 0 | 4 | 21,1 | 15 | 78,9 | 19 |
| Toplam | 42 | 10,3 | 113 | 27,7 | 253 | 62 | 408 |

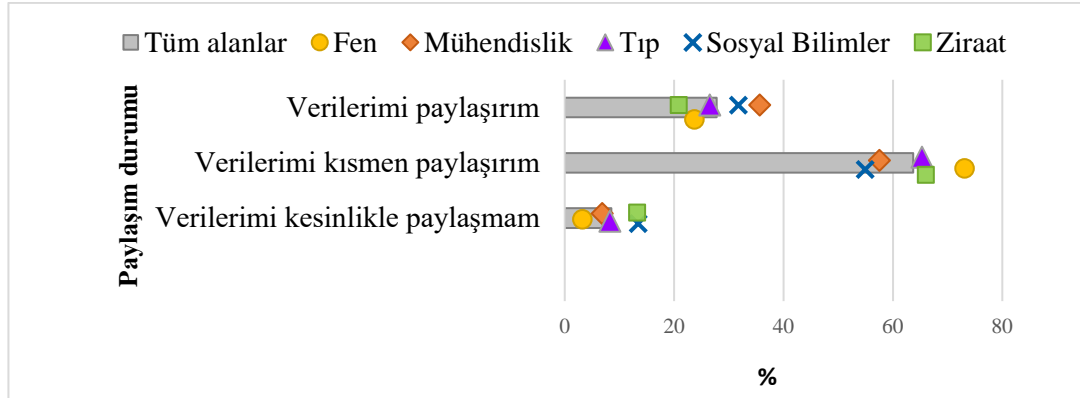
Açık bilim politikası belgesine yönelik farkındalık durumunun kuruma göre dağılımı Tablo 13'te sunulmuştur. İTÜ ve ODTÜ mensubu araştırmacıların yaklaşık yarısı TÜBİTAK Açık Bilim Politikası'nın *farkında değilken*, diğer üniversitelerde bu oran nispeten daha yüksektir. Politika belgesinin farkında olduğunu belirten katılımcılar arasında bu belgeyi *okuyanların* oranı en yüksek ODTÜ mensupları arasındayken (%18,5; n=5), *okumayanların* oranı en yüksek İTÜ mensupları arasındadır (%49; n=25). Bu durum kurumun yapısı, işleyişi, birimlerin koordineli çalışması ve araştırmacının pozisyonu gibi faktörlerden kaynaklanıyor olabilir.

4.3 VERİ PAYLAŞIMI

Araştırmacıların veri paylaşımı konusundaki yaklaşımlarını belirlemek amacıyla verilerini başkalarıyla paylaşıp paylaşmadıkları, paylaşıyorlarsa kimlerle ve hangi yollarla paylaşım yapıldığı, paylaşmıyorlarsa neden paylaşım yapmadıkları sorulmuştur.

4.3.1 Veri Paylaşım Durumları

“Verilerinizi başkalarıyla paylaşır mısınız?” sorusuna, araştırmacıların %27,5'i (n=112) *verilerimi paylaşırım*, %63,7'si (n=260) *verilerimi kısmen paylaşırım (sadece bazı bölümlerini)*, %8,6'sı (n=35) ise *verilerimi kesinlikle paylaşmam* seçeneğini işaretleyerek yanıt vermişlerdir. Bir araştırmacı ise veri üretmediğini belirtmiştir.



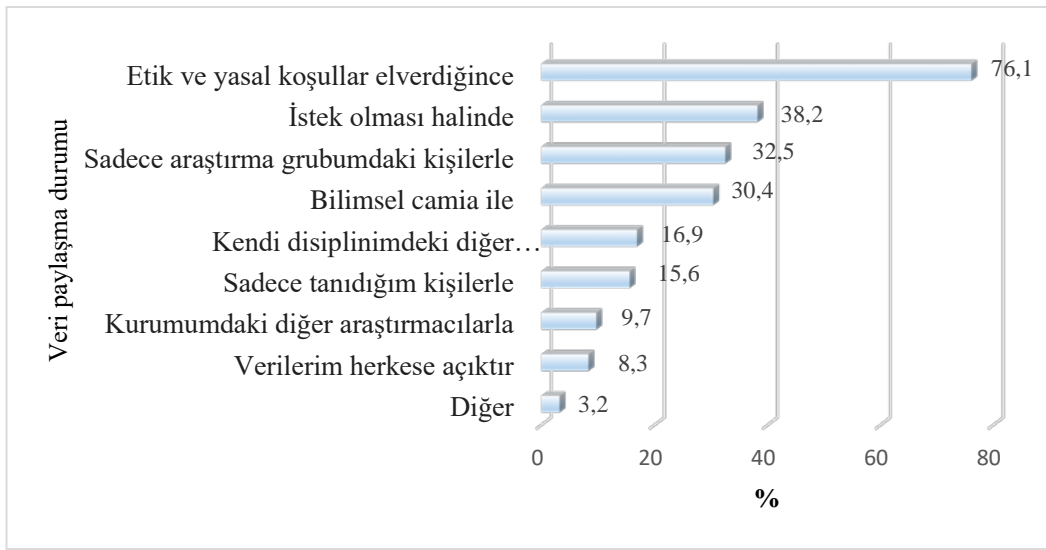
Şekil 16. Katılımcıların disipline göre veri paylaşım durumları (%)

Araştırmacıların veri paylaşım durumunun disipline göre dağılımı Şekil 16'da sunulmuştur. İstatistiksel açıdan anlamlı olmamakla birlikte, araştırmacıların veri paylaşım yaklaşımlarında disipline özgü küçük farklılıklar bulunmaktadır. Mühendislik alanında verilerini paylaşanların oranı diğer disiplinlere göre daha fazlayken (%35,6; n=26), ziraat alanında daha azdır (%20,8; n=11). Fen alanında, verilerinin sadece bazı bölümlerini paylaştıkları için *verilerimi kısmen paylaşırım* seçeneğini işaretleyen araştırmacılar en fazlayken (%73,1; n=68), *verilerimi kesinlikle paylaşmam* seçeneğini işaretleyenler en azdır (%3,2; n=3). *Verilerini kesinlikle paylaşmayan* araştırmacılar en çok sosyal bilimler (%13,4; n=11) ve ziraat alanında (%13,2; n=7) olup oranları birbirine yakındır.

Verilerini paylaşan araştırmacıların (%91,2; n=372) veri paylaşımına yönelik yaklaşımlarını belirlemek amacıyla, araştırmacılardan kendilerine sunulan ifadelerden uygun olan seçenek/seçenekleri işaretlemeleri istenmiştir (bkz. Şekil 17). Elde edilen bulgulara göre, veri paylaşan katılımcıların yaklaşık dörtte üçü (%76,1; n=283) *verilerimi etik ve yasal koşullar elverdiğince paylaşırım*, yaklaşık beşte ikisi ise (%38,2; n=142) *istek olması halinde paylaşırım* ifadesini işaretlemişlerdir. Türkiye'deki araştırmacılar üzerinde yapılan bir diğer çalışmada da araştırmacılar verilerinin en çok (%44) "istek üzerine başkalarının kullanımına ve erişimine açık" olduğunu belirtmiştir (Ünal ve Kurbanoğlu, 2018, s. 297).

Veri paylaşım durumlarında diğer tercih edilen seçeneklere göre, araştırmacıların verilerini daha çok sadece *araştırma grubundaki kişilerle paylaştığı* (%32,5; n=121) ve

verilerini bilimsel camia ile paylaştığı (%30,4; n=113) görülmektedir. Verilerini kendi disiplinindeki diğer araştırmacılarla (%16,9; n=63), sadece tanıdığı kişilerle/araştırmacılarla (%15,6, n=58) ve çalıştıkları kurumdaki diğer araştırmacılarla (%9,7; n=36) paylaşanların oranı nispeten azdır. Verilerinin herkese açık olduğunu belirten araştırmacıların oranı ise sadece %8,3'tür (n=31). Ünal ve Kurbanoğlu'nun (2018) çalışmasında, araştırmacıların en çok (%76) kendi ekibindeki diğer araştırmacılarla, ikinci olarak (%45) mensup olunan kurumdaki diğer araştırmacılarla iş birliği yaptığı görülmektedir (Ünal ve Kurbanoğlu, 2018, s. 297).



Şekil 17. Katılımcıların veri paylaşma koşulları (%)

Veri paylaşım durumuyla ilgili görüş bildiren katılımcıların ifadeleri şu şekildedir:

- *Deneyisel çalışıyorum ve deneyisel çalışmalardan elde edilen verilerin olduğu gibi herkesle paylaşılması uygun değildir. Deney koşulları, adımları, süreç vs gibi bilgiler olmadan verilerin yorumlanması yanlış sonuçlar çıkarılmasına sebep olur. Bunun yanı sıra yayınlamış olduğum tüm verilerimi isteyen herkesle paylaşırım. Doktora öğrencisi iken daha yayınlamadığım özgün bilgilerimi araştırmacı arkadaşımınla paylaşmış, daha sonra onun sunumunda kendi çizimlerimi özgün işi olarak izlemek zorunda kaldım. Sonuç olarak, veri paylaştığınız herkesin verileri kullanma ve yayma konusunda etik davranacağından emin olamıyorsunuz. Buna çözüm olarak yayınladıklarımı paylaşma kuralını benimsedim.*

- *Elde ettiğim veriler konuyla ilgili herkese açıktır ve projeyi kaynak göstererek veya atıf yaparak (bana değil proje destekçileri beni değil projeyi destekliyor) bulgu ve yorumlarımı etik ve hukuka uygun olarak kullanılabilir.*
- *Karşılıklılık esasını gözetken diğer araştırmacılarla paylaşıyorum.*
- *Ortak bilimsel yayın üretecek bilim insanları ile paylaşıyorum*
- *Verilerimi kendi araştırma sorumla ilgili kısmı yayınladıktan sonra tüm bilim camiasıyla paylaşabilirim.*
- *Yayın ya da patent başvurusu yapana kadar sadece birlikte çalıştığım kişilerle, sonrasında herkesle paylaşıyorum*
- *Sadece gerekli gördüğüm kadarını paylaşıyorum*
- *Uluslararası dergilerde makale olarak basıldığı için ilgilenen herkes görebilir.*

Paylaşım ile ilgili araştırmacıların endişeleri olduğu ve bazı araştırmacıların araştırma verisi kavramını yanlış yorumladığı görülmektedir.

Veri paylaşım bulgularıyla birlikte, uzun süreli depolamada en çok kişisel bilgisayarların tercih edilmesi, üst veri bilgilerinin oluşturulmasında standartların kullanım oranının az olması ve VYP hazırlayanların oranının az olması sonuçları da dikkate alındığında paylaşımaya yönelik olumlu düşüncelerin fazla eyleme geçemediğini, verinin tekrar kullanılabilir olması amacıyla bir paylaşım yapılmadığını söylemek mümkündür. Belirli bir kesim ile paylaşılan verilerin ise ne kadar anlaşılır ve tekrar kullanılabilir olduğu belirsizdir.

Araştırmacıların veri paylaşım durumlarında disipline özgü bazı küçük farklılıklar görülmektedir. Fen alanında (%76,7; n=69) *etik ve yasal koşullar elverdiğince* paylaşım durumu, mühendislik (%64,7; n=44) ve tıp (%73,3; n=66) alanına göre daha yükseken, sosyal bilimler (%84,5; n=60) ve ziraat (%84,8; n=39) alanına göre daha düşüktür ($\chi^2_{(4)} = 9,932$; $p=0,042$). Verilerini sadece araştırma grubundaki kişilerle paylaşım durumu incelendiğinde, disipline göre istatistiksel açıdan anlamlı farklılıklar görülmektedir ($\chi^2_{(4)} = 12,441$; $p=0,014$). Diğer disiplinlere oranla, en çok tıp (%43, n=39) ve en az ziraat (%17, n=8) alanındaki araştırmacıların, verilerini *sadece araştırma grubundaki kişilerle paylaştığı* görülmektedir. Verilerini *istek halinde paylaşanların* oranı mühendislik (%44, n=30) ve fen (%43, n=39) alanında birbirine yakinken *herkesle paylaşanların* oranı

mühendislikte (%13, n=9) nispeten daha fazladır. Verilerini *bilimsel camia ile paylaşılanlar* en çok fen alanında (%38,9; n= 35) iken, en az sosyal bilimlerde (%21,1; n=15). Mühendislik (%30,9; n=21), tıp (%30; n=27) ve ziraat (%28,3; n=13) alanında bilimsel camia ile paylaşım yapanların oranı ise birbirine yakındır.

Türkiye’deki araştırmacılar üzerine yapılan bir çalışmada, katılımcıların veri paylaşımına yönelik düşünceleri sorgulanmış ve ilgili soruyu yanıtlayanların büyük çoğunluğunun (%62,4) veri paylaşımına olumlu baktığı, diğerlerinin ise (%37,4) veri paylaşımına yönelik olumsuz düşüncelerinin olduğu görülmüştür (Aydınoglu ve diğerleri, 2017, s. 278). Benzer bir şekilde Ankara Üniversitesi proje yürütücülerinin çoğunun (%60,4) “araştırma verilerinin paylaşılmasının önemli ve gerekli olduğu” düşüncesine katıldıkları görülmektedir (Zencir, 2019, s. 156). Başka bir çalışmada da katılımcıların çoğunluğu (ilgili soruya yanıt verenlerin %76’sı) kamu kaynaklarıyla üretilen verinin kamu malı olduğu ve söz konusu verinin açık erişimli olması gerektiğini düşünmektedir (Doğan ve diğerleri, 2021, s.58).

4.3.2 Veri Paylaşım Yolları

Verilerini paylaşan araştırmacıların (n=372) verilerini hangi yollarla paylaştığını belirlemek amacıyla katılımcılara birden fazla seçeneğin işaretlenebildiği bir soru yöneltilmiştir. Araştırmacılar birden fazla paylaşım yolunu tercih etmekle birlikte, en çok *e-posta* %78,2 (n=291), ikinci olarak *dergi makalesine ek bağlantı* (%37,4; n=139) seçeneklerinin işaretlendiği görülmektedir. Veri paylaşımında *veri havuzları* (%21; n=78), *kişisel web sayfası* (%7,3; n=27) ve *üniversite web sitesi* (%6,7; n=25) kullanımı nispeten daha düşüktür. Diğer seçeneğini işaretleyenlerin çoğunlukla istek halinde paylaşım yaptıklarını söylemek mümkün olup araştırmacıların ifadeleri şu şekildedir: *bulut (2 kişi), bulut depolama üstünden şifrelenmiş olarak; Google Drive üzerinden erişime açma; Researchgate gibi siteler (2), mendeley, Disk, DVD vs. transferi; harici disk aracılığıyla aktarma; boyut büyükse wetransfer web sayfası ya da harici disk/flash bellek; eeg verisi ya da spss-excel dosyaları halinde flash bellekle; wetransfer gibi hacimli veri transferi sağlayıcılar; dernek; yüz yüze; paylaşmam gerekirse belli şartlar altında özel olarak ulaştırırım; talebe bağlı olarak özel linkler aracılığı ile; isteme bağlı*

olarak her türlü; çalışmamızın yayınlandığı dergilerin yönlendirdiği veri tabanlarında paylaşıyorum.

Disipline göre veri paylaşım yolları incelendiğinde bazı anlamlı farklılıklar görülmektedir. Fen alanında verilerini paylaşan araştırmacıların yarısı (n=45), mühendislik (n=26) ve tıp (n=34) alanındakilerin %38'i *dergi makalesine ek bağlantı* yoluyla veri paylaşırken, sosyal bilimlerde bu oran %22,5 (n=16) ile en düşüktür. Tıp alanında *veri havuzları* kullanım oranı (%22,2; n=20), mühendislik (%27,9; n=19) ve fen (%26,7; n=24) alanlarına göre daha az, sosyal bilimler (%16,9; n=12) ve ziraat (%6,5; n=3) alanlarına göre daha fazladır.

Veri paylaşım yollarından, *dergi makalesine ek bağlantı* ($\chi^2_{(4)}=15,362$; p=0,004) ve *veri havuzları* ($\chi^2_{(4)}=25,069$; p=0,000) kullanımı veri miktarına göre istatistiksel açıdan anlamlı bir farklılık göstermektedir. Veri miktarı fazla olanların verilerini *dergi makalesine ek bağlantı* olarak paylaşım oranları (>100 GB – 1 TB %59,6 n=31; >1 TB %47,8; n=11) verisi nispeten az olanlara (1 GB'den az %37,3 n=31; >1GB – 100 GB %31; n=45) ve veri miktarını bilmeyenlere (%31,8; n=21) göre daha fazladır. Veri miktarı arttıkça veri havuzlarını kullananların oranı da artmaktadır (veri miktarını bilmeyenler %9,1 n=6; 1 GB'den az %9,6 n=8; >1GB – 100 GB %23,4 n=34; >100 GB – 1 TB %36,5 n=19; >1 TB %39,1 n=9).

Araştırmacıların veri paylaşım yaklaşımlarını belirlemek amacıyla yayıncı firma Wiley tarafından (2014) farklı disiplinlerdeki 2250'den fazla araştırmacıya uygulanan ankete göre, katılımcıların %52'si verilerini kamuya açık olarak paylaşmakta ve çoğunlukla (%67) dergiye ek materyal olarak paylaşım yapmaktadır. Bu araştırmada, dergi makalesine ek bağlantı yoluyla veri paylaşanlarının oranının daha fazla olmasının yayıncı politikalarıyla ilgili olduğu düşünülmektedir. Başka bir çalışmanın bulgularına göre ise, araştırmacılar verilerini en çok kişisel web siteleri aracılığıyla paylaşmaktadır (Van Loon, Akers, Hudson ve Sarkozy, 2017, s. 101-102),

4.3.3 Veri Paylaşmama Nedenleri

Birden fazla seçeneğin işaretlenebildiği bir soruyla, verilerini paylaşmayan araştırmacıların başkalarıyla paylaşım yapmama nedenleri saptanmaya çalışılmıştır (bkz. Şekil 18). Verilerini kesinlikle paylaşmayan araştırmacıların (n=35) yaklaşık yarısı (%48,6; n=17) *yasal ve etik sorunlar*, %34,3'ü (n=12) *verilerini kamuya açma hakkının bulunmaması* ve %28,6'sı (n=10) *verinin hatalı kullanılabilceği endişesi* nedenleriyle verilerini paylaşmamaktadır. *Verilerimin açıklama olmadan anlaşılacak durumda olmaması* (%25,7; n=9), *verinin yanlış yorumlanabileceği endişesi* (%22,9; n=8), *başkalarının benim verilerime ihtiyacı olmadığını düşünüyorum* (%17,1; n=6) ve *verilerimin temizlenmemiş olması* (%14,3; n=5) da veri paylaşmama nedenleri arasındadır. Paylaşmama nedenlerinden *bilimsel gücü kaybetme korkusu* ve *bu konuda zorlayıcı olmaması (fon veren kuruluş, çalıştığım kuruluş gibi)* üçer katılımcı (%8,6); *verilerimi kamuya açacak teknik bilgi yetersizliği*, *verilerimi sunacak yer olmaması* ve *üst veri standardı yetersizliği* de sadece birer araştırmacı (%2,9) tarafından işaretlenmiştir.



Şekil 18. Veri paylaşmama nedenleri (%)

Diğer seçeneğini işaretleyerek katkı sağlayan araştırmacıların görüşleri şu şekildedir:

- *Başkaları tarafından verilerin kullanılarak haksız yazarlık yapılması*
- *Kendi topladığım veri Etik Kurul Onayı ile katılımcılardan Onam alarak topladığımız veriler olduğundan paylaşılması uygun değil...*
- *Bugüne kadar yaptığım, fonlanan ya da fonlanmayan tüm araştırmalarımnda kişisel katkılarım vardır. Bu katkılar zamansal ve parasaldır. Bana ait olan bu katma değeri neden paylaşayım ki?*

Başka bir çalışmada, veri paylaşımıyla ilgili olarak araştırmacıların en çok (%47) “yasal ve etik sorunlar” konusunda endişe duyması bu çalışmayla benzerlik göstermektedir (Ünal ve Kurbanoglu, 2018, s. 298). Ancak Zencir’in (2019) çalışmasında, veri paylaşımını etkileyen en önemli nedenin “yayın fırsatını kaybetme korkusu” (%62,9) olduğu ortaya çıkmıştır (Zencir, 2019, s.159).

Katılımcıların genellikle veri paylaşımına yönelik olumlu düşünceleri olduğunu, ancak verilerini erişilebilir ortamlarda sunmadıklarını ve verilerin başkası tarafından istenmesi ve etik ve yasal koşulların elvermesi halinde paylaştıklarını söylemek mümkündür.

Açık uçlu bir soruyla, araştırmacılarından varsa veri yüklemek ve indirmek için kullandıkları (Zenodo, DRYAD, Figshare gibi) veri arşivlerini belirtmeleri istenmiştir. Görüş bildiren katılımcıların birçoğu (n=51) bu arşivleri bilmediğini veya kullanmadığını belirtmiştir. Bazı katılımcılar veri depolama veya aktarım için kullanılan uygulamaların (Google Drive, Dropbox) veya dosya transfer programlarının (wetransfer – transfernow) adlarını yazmışlardır. İki katılımcı ise bu soruya geçerli bir yanıt (web of science, scopus vb.; .rar .zip) vermemiştir. Veri arşivi kavramının tam anlaşılmadığı görülmektedir. Bu soruya geçerli bir yanıt vererek katkı sağlayan araştırmacıların kullandıkları arşivler şu şekildedir:

Zenodo (4 kişi); Mendeley datasets (2 kişi); Github; NCBI; UCI, Kaggle; NOAA, PANGEA; DRYAD, NIST Libraries, NCBI GenBank, UniProt; Endfb; Figshare; Tür kayıtlarımın bulunduğu özel arşivler (örn. observation.org); Turkoscore veri tabanı (Bilkent ve Ankara Üniversitesi ortak veri tabanı); KOERİ; Aspera Connect; NASA, JPL,

USGS, Earthdata; NASA'nın ftp siteleri; OECD Stat, TCMB EVDS, IMF, Dünya Bankası, Quandl, FRED; ICPSR

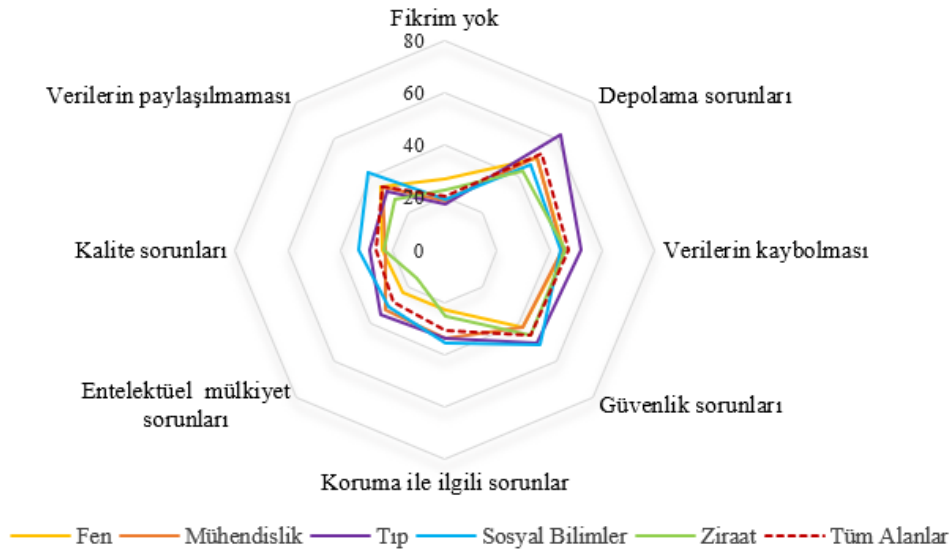
4.3.4 Araştırma Verilerinin Yönetimindeki Temel Problemler

Veri yönetimindeki temel problemlere yönelik düşüncelerini belirlemek amacıyla, araştırmacılara birden fazla seçeneğin işaretlenebildiği bir soru yönlendirilmiştir. Temel problemlerin anlaşılmasının araştırmacılara sunulacak hizmetlerin belirlenmesinde yardımcı olacağı düşünülmektedir.

Araştırmacıların yaklaşık yarısı (%51,8; n=211) *depolama sorunlarını* araştırma verileri yönetimindeki temel problemler olarak görmektedir. *Verilerin kaybolması* (%46,9; n=191) ve *güvenlik sorunları* (%46,2; n=188) ise en çok tercih edilen diğer seçeneklerdir. Araştırmacıların büyük miktarda veriler kullanmamalarına rağmen depolamayı sorun olarak görmeleri, sorunun verinin boyutuyla değil depolama alanının güvenliğiyle ilgili olabileceğini düşündürmektedir. Ayrıca, araştırmacıların tümü veri yönetiminde depolama aşamasını bizzat gerçekleştirdiği için bu konuda daha çok deneyime sahip olmaları ve sorunların daha çok farkında olmaları olasılığı da vardır. Literatürdeki bir çalışmada da temel problem olarak görülen ilk üç seçenek bu çalışmayla benzerlik göstermektedir (Chen ve Wu, 2017, s. 350). Katılımcıların yaklaşık üçte biri (%33,9; n=138) *verilerin paylaşılmaması*, %31'i (n=126) *koruma ile ilgili sorunlar*, %28'i (n=114) *entelektüel mülkiyet sorunları*, %26,5'i (n=108) *kalite sorunları* seçeneğini işaretlemişlerdir. Bu konuda fikri olmayanların oranı ise %20,6'dır (n=84). Bir katılımcı soruya yanıt vermemiş olup diğer seçeneğinde görüş bildiren üç araştırmacının görüşleri şu şekildedir: *uzun süreli depolama sorunları; maalesef bende dâhil ülkemizdeki birçok araştırmacının arşivleme kültürünün olmaması oldukça önemli bir problemdir; araştırma verilerinin orijinal halde tutulmaması*. Araştırmacıların kullandıkları veri miktarı az olmasına rağmen depolama sorunlarını problem olarak görmeleri, veri depolama ortamı olarak kendi bilgisayarlarını güvenli bulmadıklarından kaynaklanıyor olabilir.

AVY konusundaki temel problemlerin disiplin temelli dağılımı incelendiğinde, istatistiksel açıdan anlamlı olmamakla birlikte disipline özgü bazı farklılıklar bulunduğu

görülmektedir (bkz. Şekil 19). Tıp (%62,2; n=61), mühendislik (%50; n=37) ve fen (%49,5; n=46) alanlarında *depolama sorunları*, sosyal bilimlerde *güvenlik* (%51,2; n=42), ziraat alanında ise *verilerin kaybolması* (%46,2; n=24) ve *güvenlik* (%46,2; n=24) en çok tercih edilen temel problemler olarak görülmektedir. AVY konusundaki temel problemler hakkında fikri olmayanların oranı (%26,9; n=25) fen alanında en fazladır.



Şekil 19. AVY alanında temel problemler

Diğer disiplinlere oranla, *verilerin paylaşılmaması* (%41,5; n=34), *güvenlik sorunları* (%51,2; n=42) ve *kalite sorunları* (%32,9; n=27), *koruma ile ilgili sorunlar* (%35,4; n=29) sosyal bilimlerde; *verilerin kaybolması* (%52; n=51), *entelektüel mülkiyet sorunları* (%34,7; n=34) ise tıp alanında daha fazla tercih edilmiştir.

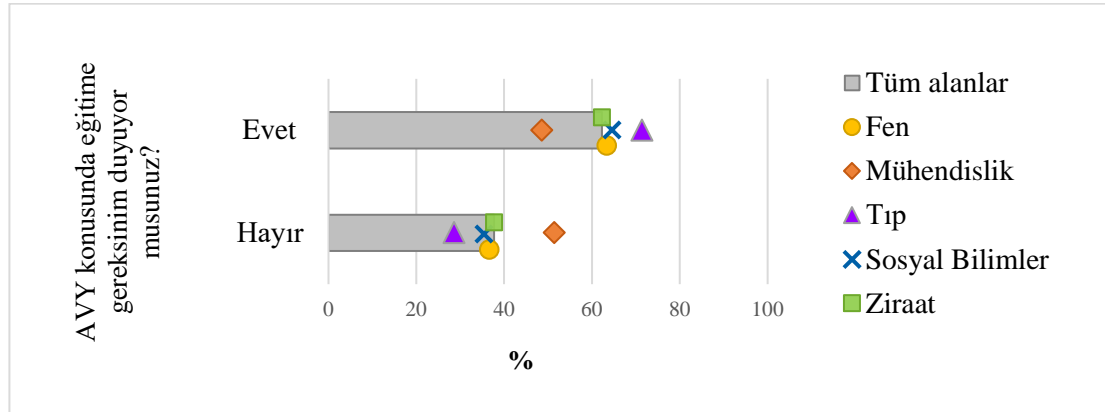
Ziraat alanında ise, *depolama sorunları* (%42,3; n=22), *entelektüel mülkiyet sorunları* (%15,4; n=8) ve *verilerin paylaşılmaması* (%26,9; n=14) diğer disiplinlere oranla daha az işaretilenmiştir.

4.4 AVY KONUSUNDA EĞİTİM GEREKSİNİMİ

Bu bölümde; araştırmacıların AVY konusunda eğitim gereksinimi olup olmadığını belirlemek, eğer eğitim gereksinimi varsa hangi konularda eğitim almak istediklerini saptamak ve TÜBİTAK tarafından sunulan Araştırma Verileri Yönetimi Eğitim Portalı

hakkındaki farkındalık durumlarını ortaya çıkarmak amaçlarıyla yöneltilen sorulara ait bulgulara yer verilmiştir.

Katılımcıların %37,7'si (n=154) AVY konusunda eğitime gereksinim duymadığını belirtmiştir. Eğitime gereksinim duyan katılımcıların (%62,3; n=254) ise birden fazla konuda eğitim almak istedikleri görülmektedir.



Şekil 20. AVY konusunda eğitim gereksinimi (%)

Disipline göre, "Araştırma verilerinin yönetimi konusunda eğitim gereksinimi duyuyor musunuz?" sorusuna verilen yanıtlar Şekil 20'de sunulmuştur. İstatistiksel açıdan anlamlı olmamakla birlikte disipline özgü bazı farklılıklar bulunmaktadır ($\chi^2_{(4)}=9,602$; $p=0,048$). Eğitime en çok gereksinim duyan disiplin tıp (%71,4; n=70) alanırken, mühendislik (%48,6; n=36) en az eğitim ihtiyacı olan bilim dalıdır. Sosyal bilimler (%64,6; n=53), fen (%63,4; n=59) ve ziraat (%62,3; n=33) alanında eğitime ihtiyaç duyan araştırmacıların oranı ise birbirine yakındır.

Eğitim gereksinimi yaşa göre büyük bir farklılık göstermemekle birlikte yaş arttıkça eğitim gereksiniminin bir miktar azaldığı görülmektedir. Yaşı 39'dan küçük araştırmacıların %64,1'i (n=66), 39-49 yaş grubundakilerin %62,5'i (n=100), 50-60 yaş grubundakilerin %61,6'i (n=66) ve 60 yaş üstü araştırmacıların %59,4'ü (n=19) eğitime gereksinimi olduğunu belirtmiştir. Unvana göre eğitim gereksinimleri incelendiğinde ise Profesör (%59,6; n=112) ve Doçent (%60,6; n=57) unvanına sahip araştırmacıların öğretim görevlisi (%68,4; n=13), Dr. öğretim üyesi (%67,8; n=40) ve araştırma görevlisi

(%66,7; n=32) arařtırmacılara oranla grece eđitime daha az gereksinimi olduđu grlmektedir.

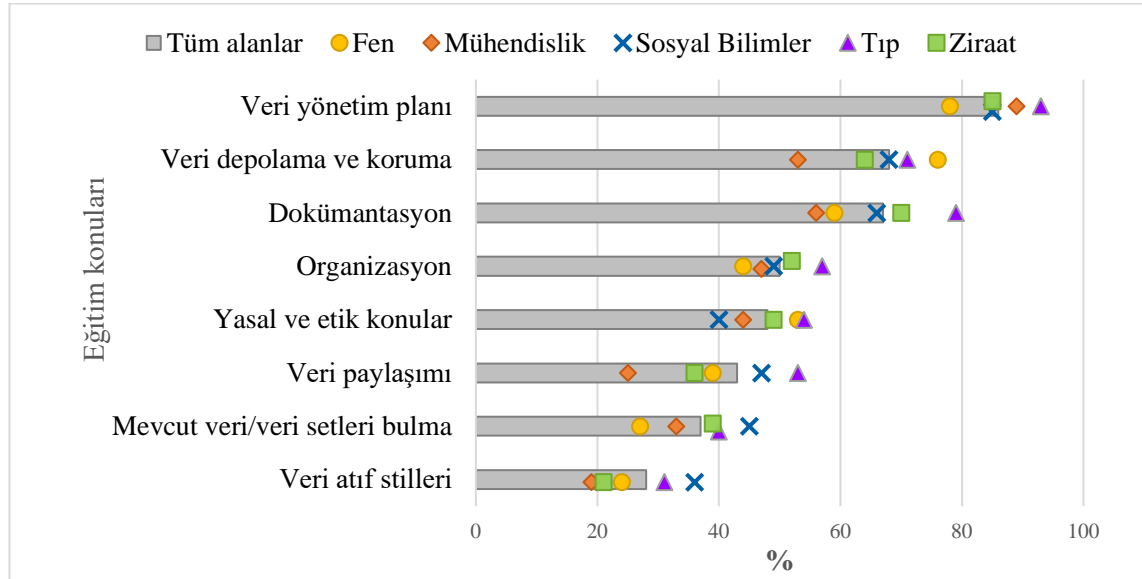
Kuruma gre incelendiđinde, İstanbul niversitesi'ndeki arařtırmacıların daha ok (%68,5; n=37), İT'deki arařtırmacıların ise daha az (%47,1; n=24) eđitim gereksinimi duyduđu saptanmıřtır. Diđer niversitelere mensup (%62,7; n=64) arařtırmacılar, ODT mensuplarına (%59,3; n=16) gre daha fazla, Ankara (%66,2; n=51) ve Hacettepe niversitesi mensuplarına (%65,4; n=51) gre daha az eđitim ihtiyacı hissetmektedir. Arařtırma verilerinin ynetimi konusunda hizmet ve eđitime gereksinimi olmayanların oranı (%33,8), Ankara niversitesi proje yrtcleri zerine yapılan alıřmayla benzerlik (%34) gstermektedir (Zencir, 2019, s. 164). Trkiye'deki arařtırmacılar zerinde yapılan bařka bir alıřmada ise, katılımcıların %67'sinin TBİTAK tarafından dzenlenecek eđitimsel katılmaya ve hazırlanan aık eriřimli kaynakları kullanmaya istekli oldukları grlmřtr (Dođan ve diđerleri, 2021, s. 58).

4.4.1 Eđitim Gereksinimi Duyulan Konular

Eđitim gereksinimi duyulan konulara iliřkin dađılım Őekil 21'de sunulmuř olup en ok iřaretlenen seenek (tm katılımcıların %53,4', eđitim gereksinimi olanların %85,8'i, n=218) veri ynetim planıdır. VYP hazırlayan arařtırmacıların oranının dřk olması bu konunun yeteri kadar bilinmediđini gstermektedir. Veri ynetim planının hem bilinmemesi hem de fon sađlayıcı kurum/kuruluřların planı zorunlu tutması nedeniyle en ok bu konuda eđitim gereksinimi hissedildiđini sylemek mmkndr. VYP veri yařam dngsndeki tm unsurları kapsadıđı iin de bu seenek daha ok iřaretlenmiř olabilir.

Sıklıkla tercih edilen diđer seenekler *veri depolama ve koruma* (%68, n=173) ile *dokmantasyon* (%67, n=170), en az tercih edilen seenek ise *veri atıř stilleri* (%28, n=70) olmuřtur. Veri atıř stillerinin az tercih edilme nedeni, katılımcıların ođunlukla yeni veri oluřturması ve dıř kaynaklı veri kullanım oranının dřk olması ile aıklanabilir. Diđer seeneđini  kiři iřaretlemiř ve iki kiři eđitim iin zamanı olmadıđını belirtmiřtir. Bir arařtırmacı ise veri madenciliđinde veri keřfine ynelik yntem ve programlar hakkında eđitime gereksinim duyduđunu belirtmiřtir.

Literatürdeki çalışmalar incelendiğinde, araştırmacıların en çok eğitim ihtiyacı duydukları konular uzun süreli erişim ve koruma ile araştırma sırasında verilerin depolanması ve yedeklenmesidir (Buys ve Shaw, 2015, s. 15-18; Renwick, Winter ve Gill, 2017, s. 58). Türkiye’deki araştırmacılar üzerine yapılan bir çalışmada VYP hazırlama, başka bir çalışmada ise üst veri oluşturma en çok eğitim gereksinimi duyulan konulardır (Ünal ve Kurbanoğlu, 2018, s. 305; Zencir, 2019, s. 165).

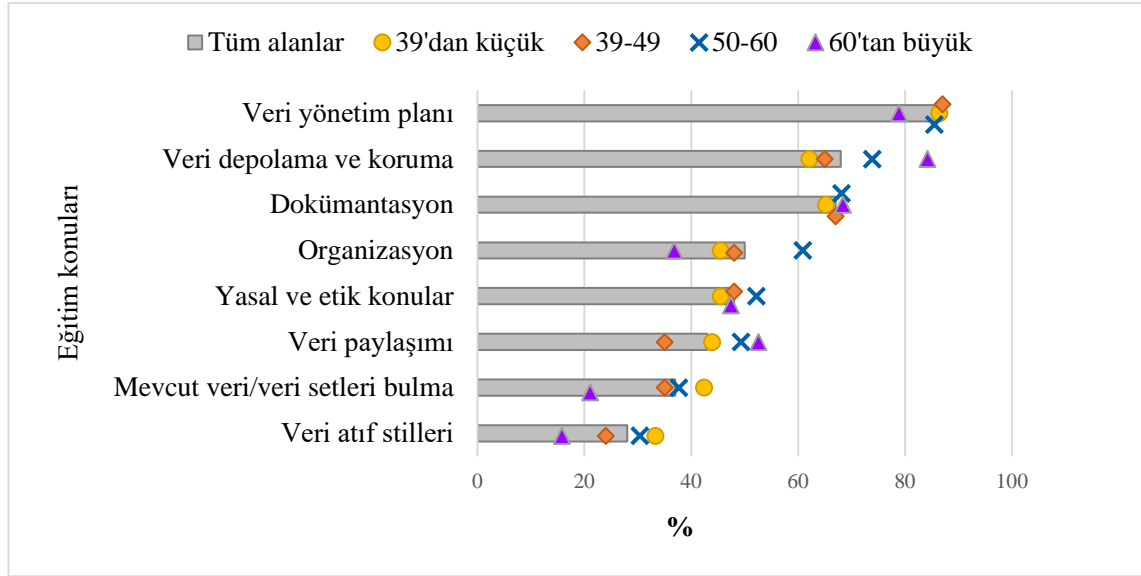


Şekil 21. Disipline göre eğitim gereksinimleri (%)

Disipline göre eğitim ihtiyaçları incelendiğinde, tıp alanındaki araştırmacıların *veri yönetim planı* (%92,9; n=65), *dokümantasyon* (%78,6; n=55), *organizasyon* (%57,1; n=40), *yasal ve etik konular* (%54,3; n=38) ve *veri paylaşımı* (%52,9; n=37) konularında diğer disiplinlere oranla nispeten daha fazla eğitime ihtiyaç duyduğu görülmektedir. Fen alanında diğer disiplinlere oranla, veri depolama ve koruma konularına yönelik eğitim gereksiniminin daha fazla (%76,3; n=45) olmasının bu alanda kullanılan veri hacminin daha büyük olmasıyla ilgili olduğu düşünülmektedir.

Yaş arttıkça (39’dan küçük, 39-49, 50-60 ve 60’tan büyük), *dokümantasyon* (sırasıyla: %65,2 n=43; %67 n=67; %68,1 n=47; %68,4 n=13) ve *veri depolama* (sırasıyla: %62,1 n=41; %65 n=65; %73,9 n=51; %84,2 n=16) konularında eğitim ihtiyacının da arttığı görülmektedir (bkz. Şekil 22). Dokümantasyon konusunda eğitim ihtiyacında yaş grupları arasında büyük bir fark olmamakla birlikte, veri depolamada özellikle 50 yaş üstü

araştırmacıların eğitim ihtiyacının fazla olması depolama seçenekleri ile ilgili sonuçları desteklemektedir. Görece yeni teknoloji olan bulut hizmetlerinin kullanımı yaş arttıkça azalmakta olduğundan yaş fazla olan araştırmacıların bu alandaki eksikliklerinin farkında olduğunu söylemek mümkündür.

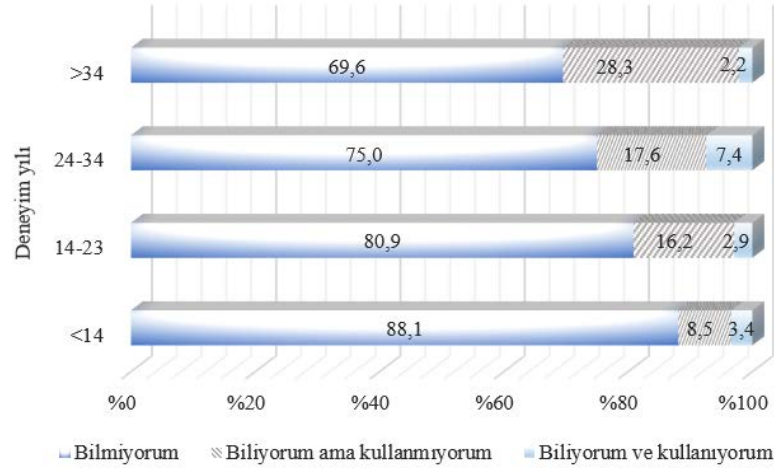


Şekil 22. Yaşa göre eğitim gereksinimleri (%)

Deneyime göre veri atıf stilleri konusunda eğitim ihtiyacında ise ters bir orantı vardır. Deneyim arttıkça (14'ten az, 14-23, 24-34, 34'ten fazla) veri atıf stilleri konusunda eğitim ihtiyacı azalmaktadır (sırasıyla: %36,7 n=29; %25,6 n=20; %22,2 n=16; %20 n=5). Deneyimi fazla olan araştırmacıların bu konuda belli bir birikimleri olması nedeniyle eğitim ihtiyacının az olduğu söylenebilir.

TÜBİTAK Araştırma Verileri Yönetimi Eğitim Portalı Farkındalığı

TÜBİTAK tarafından hizmete sunulan Araştırma Verileri Yönetimi Eğitim Portalı'nın, bu konuya yönelik eğitim ihtiyacını karşılayacağı düşüncesiyle araştırmacıların portal hakkındaki farkındalık durumu belirlenmeye çalışılmıştır. Katılımcıların büyük çoğunluğunun (%80, n=327) bu portalı bilmedikleri sonucuna ulaşılmıştır. "Biliyorum ve kullanıyorum" (%4, n=17) seçeneğini işaretleyenler, "biliyorum ama kullanmıyorum" (%16, n=64) seçeneğini işaretleyenlere göre nispeten daha azdır.



Şekil 23. Deneyime göre TÜBİTAK AVY Eğitim Portalı farkındalık durumu (%)

Deneyim yılına göre, TÜBİTAK tarafından sunulan eğitim portalı hakkındaki farkındalık durumu Şekil 23'te sunulmuştur. Deneyim azaldıkça *Araştırma Verileri Yönetimi Eğitim Portalını* bilmeme durumu artmaktadır. 34 yıldan fazla deneyime sahip olanların diğer gruplara oranla bu eğitim portalını daha çok bildikleri ama kullanmadıkları (%28,3; n=13), görülmektedir. Benzer bir ters orantı yaş verileri için de geçerlidir. Yaşı 39'dan küçük araştırmacıların %84,5'i (n=87), 39-49 yaş grubundakilerin %80,6'sı (n=129), 50-60 yaş grubundakilerin %79,5'i (n=89) ve 60 yaş üstü araştırmacıların %65,6'sı (n=32) bu eğitim portalını bilmemektedir. Yaşı ve deneyimi fazla olan araştırmacıların eğitim portalı hakkında farkındalıklarının görece fazla olmasının nedeni, bu araştırmacıların sorumlulukları gereği en büyük fon sağlayıcı olan TÜBİTAK'ın sunduğu hizmetleri takip ediyor olmaları ile açıklanabilir.

4.5 KATILIMCI GÖRÜŞLERİ

Katılımcıların araştırma verileri yönetimi ile ilgili katkı sağlamak istediği hususlar varsa açık uçlu olarak belirtmeleri istenmiş ve 24 katılımcı düşüncelerini paylaşarak çalışmaya katkı sağlamıştır.

Öncelikle, anket sorularıyla ilgili olarak görüş bildirenler bulunmaktadır. Çalıştığı veri miktarını belirten bir araştırmacıyla soruların öğretici olduğunu ve konu hakkında farkındalık yarattığını belirten iki araştırmacı vardır. Bir araştırmacı, genel olarak bilimsel

iletişimle ilgi sorunlardan, birbirinden habersiz benzer çalışmaların yapıldığından bahsetmiştir.

Açık erişime karşı olduğunu belirten bir araştırmacı ile yayın süreci tamamlanmadan verilerin ve proje raporlarının erişime açılmasına karşı olduğunu belirten iki araştırmacı bulunmaktadır. Bir araştırmacı verilerini internet üzerinden kesinlikle paylaşmadığını, e-postayla paylaşımı güvenli bulmadığını, sadece taşınabilir bellek kullandığını belirtmiştir.

Verinin kaliteli veya güvenilir olmaması, ticari sır içermesi ve etik sorunlar nedeniyle veri paylaşımının sakıncalı olabileceği durumlar konusunda katkı sağlayan üç araştırmacı vardır. Bir araştırmacı verinin herkes için aynı şeyi ifade etmeyebileceğini belirtmiş, bir diğeri soruların teknik bilimlere uymadığını ve veri kavramının disipline göre farklılaşabileceğini ifade etmiştir.

Bir katılımcı kamusal kuruluşlar tarafından fonlanan araştırmalarda elde edilen tüm verilerin herkese açık olması ve bunun için bilim politikaları oluşturulması gerektiğini düşünmektedir. Başka bir katılımcı verinin tekrar kullanılabilmesi için veri yönetiminin önemli olduğunu, bu durumu AB projelerinde fark ettiğini, bu aşamada fon sağlayıcı kuruluşlara ve araştırma performansı değerlendirme sorumlularına da büyük görev düştüğünü belirtmiştir.

Dört katılımcı konuyla ilgili yapılması gerekenler konusunda önerilerde bulunmuştur. Önerileri, TÜBİTAK tarafından etkinlikler düzenlenmesi, bilgilendirme yapılması, ulusal çalıştaylar yapılması gerektiği yönündedir. Bir katılımcı ise bu konuda düzenlenecek her etkinliğe katılacağını bildirmiştir.

Veri yönetiminin eğitim gerektirmeyecek kadar basit olması gerektiğini düşünen bir araştırmacı varken, başka bir araştırmacı bu konuda öğreneceği çok şey olduğunu belirtmiştir. Bir diğer araştırmacı, konu hakkında daha fazla bilgi edinirse kendi bölümündeki öğrencilere aktarabileceğini, bir başkası ise etik ve yasal sorunlar yaşanmaması için araştırmacılara kişisel verilerin korunması kanunu ve verilerin anonimleştirilmesi hakkında eğitimler düzenlenebileceğini bildirmiştir.

Veri paylaşımının ve açık erişimin öneminin farkında olan bir arařtırmacı, veri yönetim planının AB H2020 çerçevesinde oldukça iyi tanımlanmış bir doküman olduğunu, TÜBİTAK'ın bu planı ülkemize kazandırarak güzel bir iş yaptığını ancak kullanıcılara ulaşmada ne yazık ki başarılı olmadığını ifade etmiştir.

Arařtırmacıların katkıları genel olarak değerlendirildiğinde, bazılarının gerçekten veri yönetiminin ve paylaşımının öneminin farkında olduğu, bazılarının ise anketle birlikte konuya yönelik farkındalık kazandığını söylemek mümkündür. Ayrıca arařtırmacılar veri paylaşımında endişe duydukları hususları ve yukarıda yer verilen bazı önerilerini paylaşmışlardır.

5. BÖLÜM: SONUÇ VE ÖNERİLER

Araştırma verilerinin yönetimi, planlamayla başlayan ve verilerin toplanması, düzenlenmesi, saklanması, korunması, yayımlanması ve paylaşılması aşamalarını kapsayan bir süreçtir. Bu sürecin iyi yönetilmesi hem araştırmanın tekrarlanarak doğrulanmasında hem de araştırma verilerinin paylaşılarak tekrar kullanılmasında fayda sağlamaktadır. Ekonomik kalkınma ve kamu yararı sağlanması açısından veri paylaşımı giderek önem kazanmaktadır. Verilerin tekrar tekrar toplanması maliyet, zaman ve emek kaybına neden olmakta, paylaşımıyla, kaynak ve zaman tasarrufu sağlanmaktadır. Araştırma verilerinin erişilebilir ve keşfedilebilir ortamlarda sunulması araştırmanın hem görünürlüğünü hem de etkisini artırmakta, bu da atıf sayılarına olumlu olarak yansımaktadır. Aynı zamanda, tekrarlanabilirlik özelliğiyle, araştırmanın doğrulanması ve araştırmaya olan güvenin artmasına olanak sunmaktadır. Doğrulanabilirlik, hesap verebilirlik, görünürlük, geçerlilik, verimlilik, tekrarlanabilirlik ve kaynak tasarrufu özellikleri nedeniyle veri paylaşımının önemini fark eden fon sağlayıcı kuruluşlar hibe başvurularında araştırmacılardan veri yönetim planı sunmalarını istemeye başlamıştır. Ülkemizde bilimsel araştırmalara en büyük fon sağlayıcı kuruluş olan TÜBİTAK 14.03.2019 tarihinde Açık Bilim Politikasını yayımlamış ve veri yönetim planını proje başvuru belgeleri arasına dâhil etmiştir.

Araştırma verilerinin yönetimine ilişkin mevcut durumun bilinmesi, varsa sorunların ve eksikliklerin saptanması ve bu alanda gelişim sağlanabilmesi için uygulanacak yol ve yöntemlerin belirlenmesi açısından önem taşımaktadır. Bu çalışmanın amacı 2018-2019 URAP sıralamasında ilk beşte yer alan Hacettepe, Orta Doğu Teknik, İstanbul, İstanbul Teknik ve Ankara üniversitelerinde 2014-2018 yılları arasında TÜBİTAK desteği ile yürütülen projelerde görev alan araştırmacıların araştırma verilerinin yönetimi konusundaki tutumlarını ve uygulamada kullandıkları yaklaşımları belirlemektir. URAP sıralamasında önde gelen üniversitelerde çalışmış ve TÜBİTAK'tan proje desteği almış olan araştırmacıların araştırma verilerinin yönetimi konusunda ortalamının üzerinde bir profile sahip olacakları düşünülmüştür. Bu amaçla 26.02.2020-15.01.2021 tarihleri arasında dijital veri kullanan 408 katılımcıya anket uygulanarak veri toplanmıştır. Bu bölümde bulgular ışığında elde edilen sonuçlara ve önerilere yer verilmektedir.

5.1 SONUÇLAR

Araştırmaya katılan 408 katılımcı 73 farklı üniversiteyi temsil etmektedir⁴. Katılım oranları artan sıralamayla ODTÜ, İTÜ, İstanbul, Hacettepe ve Ankara Üniversitesinden olup katılımın ancak dörtte birini diğer üniversiteler oluşturmaktadır. En fazla katılım tıp (%24) ve fen (%23) alanındayken katılımın beşte biri (%20) sosyal bilimler alanındaki araştırmacılardan oluşmuştur. Mühendislik (%18) ve ziraat (%13) alanları görece daha az temsil edilmiştir. Katılımcıların yaklaşık yarısı Profesör unvanına sahiptir. Yaşa göre en fazla katılımı yaklaşık beşte ikilik oranla 39-49 yaş aralığındakiler, deneyime göre ise yaklaşık üçte birlik bir oranla 14 ile 23 yıl arasında deneyime sahip araştırmacılar oluşturmaktadır.

Anketin ilk bölümünde sorulara geçilmeden önce, araştırma verisinin tanımı verilmiş ve nelerin araştırma verisi olduğu örneklendirilmiştir. Herhangi bir kavram kargaşasına yer vermemek için bu açıklamalar yapılmış ancak bazı açıklamalar ve yanıtlar veri kavramının bazı katılımcılar tarafından yine de farklı algılandığını göstermiştir.

TÜBİTAK araştırmacılarının, araştırma verilerinin yönetimi sürecini tüm aşamalarıyla (planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) gerçekleştirip gerçekleştirmediklerini belirlemeye yönelik bulgulardan elde edilen sonuçlar aşağıda yer almaktadır.

Planlama: Araştırmacıların yaklaşık dörtte biri veri yönetim planı hazırlamıştır. Sosyal bilimlerde VYP hazırlayanların oranı diğer disiplinlere göre nispeten fazladır. Araştırmacılar en çok kişisel tercih nedeniyle veri yönetim planı hazırlamaktadır. Bu araştırmacıların büyük çoğunluğu aynı zamanda üst veri bilgilerini de kaydetmektedir. Üst veri bilgilerini kaydetme ve veri yönetim planı hazırlama arasındaki ilişki istatistiksel açıdan anlamlıdır. VYP hazırlayan araştırmacıların hazırladıkları plan içeriğinde en çok veri toplama, ikinci olarak dokümantasyon ve üst veri, üçüncü olarak ise depolama ve yedekleme alanları yer almaktadır. VYP hazırlamada kolaylık sağlayan araçlar (DMPTool, DMPOnline gibi) araştırmacılar tarafından pek bilinmemektedir.

⁴ Bu durum araştırmacıların çalıştıkları üniversiteden ayrılıp zaman içinde başka üniversitelerde görev almalarının bir sonucudur.

Veri yönetim planı hazırlamanın araştırma verilerinin yönetiminde yardımcı olup olmadığı konusunda, araştırmacıların büyük çoğunluğu fikir sahibi değildir.

Araştırmacıların büyük çoğunluğunun veri yönetim planı kullanmaması, VYP hazırlama araçlarını bilmemesi ve veri yönetim planının AVY sürecine sağladığı fayda hakkında fikirlerinin olmaması araştırmacıların planlama sürecini uygulamadıklarını göstermektedir. Bulgular **“Araştırmacılar AVY sürecinin planlama aşamasını gerçekleştirmemektedir”** şeklinde oluşturulan hipotezin doğrulandığını göstermektedir. Ancak, geçen sürede TÜBİTAK hibe gereksinimi olarak ARDEB 1001 projelerinde veri yönetim planı hazırlamayı zorunlu tutmuştur. Bugün veri yönetim planı hazırlamadan hibe başvurusu yapmak mümkün olmadığı için TÜBİTAK araştırmacıları için mevcut durumun değiştiğini kabul etmek gerekir.

Veri Toplama ve Düzenleme: Araştırmalarda en çok kullanılan dosya türü standart ofis dokümanları, görseller ve yapılandırılmış bilimsel ve istatistiksel veridir. Araştırmalarda kullanılan dosya türleri disipline göre hem benzerlik hem de farklılık göstermektedir. Görseller, yapılandırılmış bilimsel ve istatistiksel veri, ham (makine tarafından üretilmiş) veri, yazılım uygulamaları, ses dosyaları ve kaynak kodu disiplinlere göre anlamlı farklılık gösteren dosya türleridir.

Araştırmacıların veri elde etme biçimleri incelendiğinde, tüm disiplinlerde büyük çoğunluğun yeni veri topladığı/oluşturduğu görülmektedir. Veri elde etme biçimlerinden bazıları disipline göre istatistiksel açıdan anlamlı farklılıklar göstermektedir. Diğer disiplinlere oranla, mühendislik alanındakiler *bilinen bir kaynaktan*, fen bilimleri araştırmacıları ise *üniversite araştırma ekibinden* daha çok veri sağlamaktadır. *Kendi araştırma ağından* veri elde edenler en çok 24-34 yıl arasında deneyime sahip araştırmacılarıdır. Yaş ve deneyim arttıkça, üniversite araştırma ekibinden veri sağlayanların oranı da artmaktadır. Üniversite araştırma ekibinden en az veri elde edenler 14 yıldan az deneyime sahip gruptur. 14-23 yıl arası deneyime sahip olanların yaklaşık üçte biri, diğer iki gruptaki (24-34 yıl ve 34 yıldan fazla deneyimi olan) araştırmacıların ise yaklaşık yarısı üniversite araştırma ekibinden veri elde etmektedir. Demografik özelliklere göre farklılık sadece disipline göre bazı veri elde etme biçimleri için geçerlidir. Araştırmacıların veri elde etme biçimleri yaşa, deneyime ve unvana göre farklılık göstermemektedir.

Araştırmacıların yaklaşık üçte biri dış kaynaklı veri kullanmamaktadır. Diğer kaynaklardan veri kullananların yarısından fazlası temizleme ve/veya değişiklikler için biraz çaba, yarısına yakını ise kullanılabilir hale getirmek için çok çaba ve zaman harcamaktadır. Dış kaynaklı veriyi çok çaba ve zaman harcadıktan sonra kullandığını belirtenler disipline göre en çok sosyal bilimciler, unvana göre ise araştırma görevlileridir. Yaş azaldıkça diğer kaynaklardan veri kullanım oranı artmaktadır.

Katılımcılar büyük verilerle çalışmamaktadır. Kullanılan veri miktarı tüm disiplinlerde en çok 1 GB-100 GB iken, araştırmacıların yaklaşık beşte biri kullandıkları veri miktarını bilmemektedir. Araştırmacıların yaş ve deneyimi arttıkça veri miktarı bilmeme durumu artmaktadır. Fen alanında 1 TB'den fazla veri kullanımını diğer disiplinlere oranla daha fazladır.

Katılımcıların tamamına yakını araştırma verilerini anlamlandırmak ve/veya aranılan dosyayı kolayca bulmak için dokümantasyon işlemleri yapmaktadır. Araştırmacıların dörtte üçü çalışmalarına ait klasörleri projeye, tarihe, veri türüne vb. göre belirli bir düzen içerisinde tutmakta ve yaklaşık yarısı dosya adlandırırken belirli bir kurala uyarak tutarlı adlandırma yapmaktadır. Tutarlı dosya adlandırma sosyal bilimciler tarafından daha çok yapılmaktadır. Beni oku dosyası kullanımını disiplinlere göre anlamlı bir farklılık göstermekte ve en çok mühendisler tarafından kullanılmaktadır. Ziraat alanındakilerin tamamına yakını beni oku dosyası kullanmamaktadır. Fen alanında, diğer disiplinlere oranla klasör düzeni, etiketleme ve dosya adlandırmada sürüm (versiyon) numarası daha çok kullanılmaktadır.

Araştırmacıların yarısından fazlası üst veri bilgilerini kaydetmektedir ancak üst veri bilgilerini kaydederken herhangi bir standart kullanan araştırmacıların oranı azdır. Bu katılımcılar en çok laboratuvarlarında standartlaştırılmış üst veri, ikinci olarak ise ISO (International Standards Organization) standardı kullanmaktadır.

Veri miktarı ile hem beni oku dosyası kullanımını hem de dosya etiketleme arasında istatistiksel açıdan anlamlı bir ilişki vardır. Veri miktarı fazla olan araştırmacılarda (>100 GB-1 TB ve 1 TB'den fazla) beni oku dosyası kullanımını, >100 GB-1 TB verisi olanlarda ise dosya etiketleme daha fazladır.

Her arařtırmacının kendine özgü dzenleme davranıřı olabilmektedir. Ancak hem ortak yurütülen alıřmalarda hem de gelecekte verinin herkes tarafından anlařılabilmesi iin belirli bir dzen ierisinde, kurallara ve standartlara uygun ortak bir dil kullanılması nemlidir. Arařtırmacıların oğunluęu, verilerini anlamlandırmak iin dokümantasyon iřlemleri yapmakta, alıřma klasrlerini belirli bir dzen ierisinde tutmakta ve dosyalarını tutarlı bir řekilde adlandırmaktadır. Ancak arařtırmacıların ok azı tarafından srm numarası, beni oku dosyası ve veri szlkleri kullanılmaktadır. Ayrıca verilerin tanımlanması iin st veri bilgileri kaydedilirken st veri standartlarından yararlanılmamaktadır. Bu bulgular ıřıęında, “**Arařtırmacılar AVY srecinin dzenleme ařamasını gerekleřtirmemektedir**” hipotezi reddedilmiřtir. Arařtırmacıların verilerini dzenledikleri ancak bunu standartlara gre deęil kendilerine uygun bir řekilde yaptıkları grlmektedir. Bu tr bir dzenleme yaklařımının verinin eriřimi ve bařkaları tarafından kullanımında problem yaratabileceęi grlmektedir.

Depolama ve Koruma: Arařtırmacıların tamamına yakını verilerini en az iki, yarısından fazlası en az  farklı depolama ortamında saklamaktadır. Arařtırmacıların byk oğunluęu verilerini kiřisel bilgisayarlarda saklamakta, yaklařık yarısı ise depolama ortamı olarak bulut hizmetlerinden faydalanmaktadır.

Yař arttıka depolamada ve uzun sreli korumada *bulut hizmetleri* kullanma oranı azalmaktadır. Deneyim arttıka veri saklama ortamı olarak hem *bulut hizmetleri* hem de *e-posta* kullanım oranları azalmaktadır. Profesrler saklama ortamlarından *e-posta hesabını*, arařtırma grevlileri ise *tařınabilir aygıtları* dięer unvana sahip arařtırmacılara gre daha az kullanmaktadır.

Arařtırmacıların te birinden fazlası deęiřiklik yaptıkları her an ve yaklařık drtte biri yılda birka kez arařtırma verilerini yedeklemektedir. Verilerini yedeklemeyenlerin oranı ise olduka dřktr. Arařtırma/proje bittikten sonra, arařtırmacıların byk oğunluęu verilerini sresiz saklamaktadır.

Arařtırmacıların tamamına yakını uzun sreli korumada verilerini kendi cihazlarında, yarısına yakını ise bulutta saklamaktadır. Yař arttıka uzun sreli korumada bulut hizmetlerinin kullanım oranı azalmakta; basılı olarak saklama ve kurumsal arřiv kullanım oranı artmaktadır. Kurumsal arřivler, konusal/kurum dıřı arřivler, aık eriřimli bir veri

havuzu, kamuya açık bir web sitesi ve ticari veri havuzu uzun süreli korumada çok tercih edilmemektedir.

Katılımcılar uzun süreli korumada sorumluluğun en çok kendilerinde, ikinci olarak ise proje ekibinde olması gerektiğini düşünmektedir. Yayıncılar, bağlı buldukları bölüm ve kütüphaneler uzun süreli korumada araştırmacılarca çoğunlukla sorumlu olarak görülmemektedir.

Araştırmacıların yarısından fazlası verilerini araştırma sırasında en az üç depolama alanında saklamakta, düzenli yedekleme yapmaktadır. Ayrıca araştırmacılar belirli bir veri saklama süresi stratejisine sahip olmayıp verilerini süresiz saklamaktadır. Uzun süreli korumada ise, saklama ortamı olarak en çok kişisel cihazlar tercih edilmekte ve araştırma verileri uzun süreli korumaya hazırlanmamaktadır. Bu sonuçlara göre **“Araştırmacılar AVY sürecinin saklama ve koruma aşamasını gerçekleştirmemektedir”** şeklinde oluşturulan hipotez kısmen doğrulanmıştır. Araştırmacılar saklama aşamasını gerçekleştirmekte ve verilerini süresiz olarak saklama eğilimi göstermektedir. Ancak söz konusu saklama yaklaşımını uzun süreli koruma başlığı altında değerlendirmek mümkün değildir.

Yayımlama ve Paylaşma: Araştırmacıların büyük çoğunluğu kısmen (sadece bazı bölümlerini) veya tamamen verilerini paylaşmaktadır. Verilerini kesinlikle paylaşmayan araştırmacıların oranı düşüktür. Veri paylaşmayanlar diğer disiplinlere oranla sosyal bilimler ve ziraat alanında daha fazladır. Veri paylaşan katılımcıların dörtte üçü, verilerini etik ve yasal koşullar elverdiğince paylaşmaktadır. Bu seçenek, sosyal bilimler ve ziraat alanında diğer disiplinlere göre daha fazla tercih edilmiştir. Mühendislikte görece fazla olmak üzere, araştırma verileri herkese açık olan araştırmacıların oranı azdır. Araştırmacıların üçte birinden fazlası verilerini istek olması halinde, yaklaşık üçte biri sadece araştırma grubundaki kişilerle ve bilimsel camia ile paylaşmaktadır. Diğer disiplinlere oranla, verilerini istek halinde paylaşanlar mühendislik ve fen alanında, sadece araştırma grubundaki kişilerle paylaşanlar tıp alanında, bilimsel camia ile paylaşanlar fen alanında daha fazladır.

Araştırmacılar verilerini en çok e-posta, ikinci olarak dergi makalesine ek bağlantı yoluyla paylaşmaktadır. Veri paylaşımında veri arşivleri, kişisel web sayfası ve üniversite

web sitesi kullanımı daha az tercih edilmektedir. Veri paylaşım yollarından, dergi makalesine ek bağlantı ve veri havuzları kullanımı hem disipline hem de veri miktarına göre istatistiksel açıdan anlamlı bir farklılık göstermektedir. Fen alanında verilerini paylaşan araştırmacıların yarısı dergi makalesine ek bağlantı yoluyla veri paylaşırken veri havuzları kullanımı mühendislik ve fen alanında fazladır. Veri miktarına göre ise hem dergi makalesine ek bağlantı hem de veri havuzlarının kullanımı veri miktarı fazla olanlarda daha fazladır.

Verilerini kesinlikle paylaşmayan araştırmacıların yaklaşık yarısı yasal ve etik sorunlar, üçte biri verilerini kamuya açma hakkının bulunmaması nedeniyle verilerini paylaşmamaktadır. Araştırmacıların çok azı veri yüklemek ve indirmek için veri arşivlerini kullanmaktadır

Araştırmacılar depolama sorunlarını, verilerin kaybolmasını ve güvenlik sorunlarını araştırma verileri yönetimindeki temel problemler olarak görmektedir. Disipline göre tıp, mühendislik ve fen alanlarında depolama sorunları; sosyal bilimlerde güvenlik ve ziraat alanında ise verilerin kaybolması ve güvenlik AVY konusunda en çok işaretlenen temel problemlerdir.

Araştırmacıların verilerini kısmen ve sadece belirli bir kesimle (istek olması halinde, araştırma grubundakiler, bilimsel camia gibi) paylaşması, en çok e-posta yoluyla paylaşım yapılması, verilerin herkese açık olarak sunulmaması (veri arşivleri, kişisel web sayfası ve üniversite web sitesi kullanımının az olması) araştırmacıların veri paylaşım yaklaşımlarının olumlu olduğunu ancak verilerini yayımlamadıklarını göstermektedir. Yayımlama ve paylaşma aşamasıyla ilgili hipotez (*araştırmacılar AVY sürecinin yayımlama ve paylaşma aşamasını gerçekleştirmemektedir*) kısmen doğrulanmıştır.

Elde edilen sonuçlar ışığında, **“TÜBİTAK tarafından desteklenen projelerde görev alan araştırmacılar, araştırma verilerinin yönetimi sürecini içerdiği tüm aşamalarla (planlama, veri toplama ve düzenleme, depolama ve koruma, yayımlama ve paylaşma) bir bütün olarak gerçekleştirmemektedir”** şeklinde oluşturulan temel hipotez doğrulanmıştır. Uygulamada özellikle verilerin uzun süreli korunması ve yayımlanması aşamalarının gerçekleşmediği görülmektedir. URAP sıralamasında önde gelen üniversitelerden TÜBİTAK fon desteği almış araştırmacılarla ilgili bulguların veri

yönetimi yaşam döngüsünde bazı eksikliklere işaret etmesi ülkemizde araştırma verilerinin yönetimi uygulamalarının gelişmeye ihtiyacı olduğunu göstermektedir.

Araştırma yönetiminin planlama ile yayımlama ve paylaşma aşamalarının gerçekleştirilmesinde demografik özelliklere göre farklılık bulunmaması; veri toplama ve düzenleme ile depolama ve koruma aşamalarında ise sadece bazı unsurlar ile bazı demografik özellikler arasında farklılık olması nedeniyle, ***araştırmacıların AVY sürecindeki mevcut uygulamaları demografik özelliklere göre farklılık göstermektedir*** hipotezi tamamiyle doğrulanamamıştır.

Eğitim Gereksinimi ve Diğer Konular: Katılımcıların büyük çoğunluğu AVY konusunda eğitime gereksinim duymaktadır. Diğer disiplinlere oranla tıp alanında eğitim ihtiyacı daha fazla, mühendislikte ise en azdır. Eğitim gereksinimi yaş arttıkça bir miktar düşmekle birlikte, yaş grupları arasında ihtiyaca yönelik büyük bir farklılık yoktur. Diğer üniversitelere oranla, İstanbul Üniversitesi'ndeki araştırmacılar daha çok, İTÜ araştırmacıları ise eğitime daha az gereksinim duymaktadır.

Katılımcılar sırasıyla en çok veri yönetim planı, veri depolama ve koruma ile dokümantasyon konularına yönelik eğitim ihtiyacı hissetmektedir. En az eğitim ihtiyacı duyulan konu ise veri atıf stilleridir. Diğer disiplinlere oranla tıp alanında veri yönetim planı, dokümantasyon, organizasyon, yasal ve etik konular ve veri paylaşımı konularında; fen alanında ise veri depolama ve koruma konularına yönelik eğitim ihtiyacı daha fazladır. Yaş arttıkça dokümantasyon ve veri depolama konularında eğitim ihtiyacı da artmaktadır. Deneyim arttıkça veri atıf stilleri konusunda eğitim ihtiyacı azalmaktadır.

Katılımcıların büyük çoğunluğu TÜBİTAK tarafından hizmete sunulan Araştırma Verileri Yönetimi Eğitim Portalı'nı bilmemektedir.

Araştırmacıların yarısından fazlası TÜBİTAK Açık Bilim Politikasının farkında değildir. İTÜ ve ODTÜ mensuplarının açık bilim politika belgesi farkındalık durumları (politika belgesini okuyanlar + okumayanlar) diğer üniversitelere göre daha fazladır. TÜBİTAK Açık Bilim Politikası'nın tek başına yeterli olmadığı, açık bilimin devlet politikaları ve kurum/kuruluşların (üniversiteler, YÖK gibi) politikalarıyla desteklenmesi gerektiği

düşünülmektedir. Ulusal düzeyde uygulanacak politikaların üniversiteler üzerinde önemli bir etkisi olacağı muhakkaktır.

5.2 ÖNERİLER

Bu bölümde, önce elde edilen sonuçlara dayalı öneriler sunulmakta, daha sonra araştırma verilerinin yönetilmesinde TÜBİTAK bünyesinde ve araştırmacılar tarafından yapılabilecek şeyler konusunda önerilere yer verilmiştir.

Bulgular veri kavramının araştırmacılar tarafından yanlış anlaşılabilirliğini göstermektedir. Veri kavramı ile ilgili karmaşasının büyük ölçüde veri nesnesi ile verinin karıştırılmasından oluştuğu düşünülmektedir. Örneğin, bir insandan alınan doku örneği veri nesnesidir ancak o dokuya ait ölçümler, analizler, değerler, fotoğraflar vb. o nesneye ait veridir. Bazı disiplinlerde veri ifadesinin kullanılmamasının veya veri kavramının başka bir şey ifade etmesinin de kafa karışıklığına yol açtığı görülmektedir. Örneğin tarih alanında arşiv malzemeleri, birincil kaynaklar araştırma verisiyken matematik, felsefe gibi teorik araştırma yürütülen alanlarda çalışma sonucunu destekleyen kuramlar, notlar, açıklamalar veridir. Bu nedenle araştırma verilerinin yönetimi konusunda yapılacak çalışmalarda disipline özgü örneklerle veri kavramını açıklığa kavuşturmak mümkün olduğundan görüşme tekniğiyle veri toplanması önerilebilir. Böylelikle yanlış anlamaların ve kavram karmaşasının önüne geçmek mümkün olacaktır. Ayrıca disiplin spesifik sorular içeren anketler ve belli bir disiplinle sınırlı çalışmalar da yararlı olabilir.

Bulgular araştırmacıların yaklaşık dörtte birinin daha çok kişisel tercihe dayalı olarak bir veri yönetim planı hazırladığını göstermektedir. Geçen sürede TÜBİTAK hibe gereksinimi olarak ARDEB 1001 projelerinde veri yönetim planı hazırlamayı zorunlu tutmuş ve plan oluşturmada yardımcı olmak üzere DMPRoadmap aracını kullanıma açmıştır (<https://vyp.ulakbim.gov.tr/>). Söz konusu değişiklik ışığında veri planı hazırlama aşamasının daha detaylı olarak yeniden araştırılması faydalı olacaktır. TÜBİTAK projelerinde veri yönetim planı hazırlamanın zorunlu hale geldiği düşünüldüğünde, planlama aşamasında araştırmacılara yol göstermesi açısından örnek VYP şablonunda her alana (veri toplama, dokümantasyon ve üst veri, depolama ve yedekleme, veri paylaşımı gibi) özgü eğitici bağlantılar veya yönlendirmeler sunulması faydalı olabilir.

Arařtırmacılar çoęunlukla yeni veri toplamakta, aynı zamanda dıř kaynaklı veri de kullanmakta ancak sadece belirli kesimlerden veri elde etmektedir. Verinin tekrar kullanımı zaman ve emek tasarrufu saęlamaktadır. Toplanacak verinin daha önce üretilip üretilmedięi veri arřivleri, konu arřivleri gibi veri indirme ve yükleme imkânı sunan ortamlardan mutlaka kontrol edilmelidir. Arařtırmacılara hazır verinin nasıl keřfedileceęine dair eęitim verilip arařtırmacılar hazır verileri kullanmaya özendirilebilir.

Arařtırma sırasında veriler en çok kiřisel bilgisayarlarda olmak üzere birden fazla ortamda saklanmaktadır. Arařtırmacıların tamamına yakını depolama ortamı olarak iki, yarısından fazlası üç seęenek iřaretlemiřtir. Veri kayıplarını önlemek veya en aza indirmek için yedeklemede 3-2-1 kuralına uyularak düzenli bir veri yedekleme stratejisi uygulanabilir. Arařtırma sırasında verilerin bulutta saklanması kiřisel bilgisayarlar veya taşınabilir aygıtlar kadar tercih edilmemektedir. Ücretsiz bulut hizmetleri olmakla birlikte bunlarda kapasite sınırı bulunmaktadır. Bulut hizmetlerinden yararlanmayan kurumlar, bu hizmetleri kullanmaya bařlayarak arařtırmacılarına kampüs dıřından eriřim, saklama ve altyapı imkânı sunabilir. Ayrıca yař ilerledikçe görece yeni teknolojilerin kullanım oranı azaldıęı için belirli yař gruplarına yönelik kurum içi eęitimler düzenlenebilir.

Uzun süreli korumada (daha çok uzun süreli saklama řeklinde geręekleřmektedir) en çok kiřisel bilgisayarlar tercih edilmektedir. Arařtırma verilerinin kiřisel bilgisayarlarda depolanması bir kayıp olup bu verilerden en yüksek fayda saęlamak için fon saęlayıcı kuruluř, kurum ve kütüphanenin arařtırmacıyı uzun süreli koruma konusunda desteklemesi gerekmektedir. Bunun için öncelikle gerekli teknik altyapı oluřturulmalı, kütüphaneciler ve dięer ilgili personelin bu konuda eęitildikten sonra arařtırmacılara destek verilmesi önerilir. Böylelikle arařtırma verisinin kiřisel saklama ortamlarından çıkarılıp herkesin eriřebileceęi bir ortamda sunulması, keřfedilebilir olması hedeflenebilir.

Veriler uzun süreli korumada uygun olan dosya formatlarına göre saklanmalıdır. Teknoloji eskimesi durumunda yeni ortamlara dosya transferi yapılmalıdır. Söz konusu transfer de yine kurumsallařmıř bir yapı içinde geręekleřtirilerek sürdürülebilirlięi saęlanmalıdır.

Uzun süreli koruma ortamı olarak kurumsal arşivler, konusal arşivler, açık erişimli veri havuzları ve kamuya açık web siteleri araştırmacılar tarafından pek tercih edilmemektedir. Araştırmacıların verilerini yükleyebilecekleri veri arşivlerinin sayısı artırılabilir. Veri paylaşımının sağladığı katma değer ve veriyi tekrar kullanmanın önemi hakkında araştırmacılara farkındalık kazandırmak amacıyla hem ulusal hem de kurumsal çapta etkinlikler, toplantılar, eğitimler düzenlenebilir. TÜBİTAK projeleri için uzun süreli korumada APERTA kullanılabilir. Ayrıca akademik kurumsal arşivlerin amacına uygun kullanılabilmesi için öncelikle çeşitli veri türlerini destekleyen bir veri arşivi alt yapısı oluşturulmalıdır. Bu arşiv, üniversite performans sistemine entegre edilerek değerlendirmelerin ve akademik yükseltmelerin arşivdeki araştırma verileri de göz önüne alınarak yapılmasına olanak sağlanabilir.

Araştırmacılar, uzun süreli korumada sorumlu olarak kendilerini görmektedir. Kütüphaneler bu konuda araştırmacıların gözünde arka planda kalmaktadır. Uzun süreli koruma profesyonellik gerektiren bir süreçtir. Dünyadaki uygulamalar incelendiğinde, kütüphaneler bu süreçte aktif rol oynamaktadır. Kütüphanelerin araştırma verileri yönetiminin her aşamasında araştırmacıya destek olan bir duruma evrilmesi gerekir. Uzun süreli korumada veri arşivlerinin tercih edilmesi depolama alanı, veri güvenliği, görünürlük ve erişim gibi birçok açıdan araştırmacılara olanak sunmaktadır. Araştırma çıktılarının yanı sıra araştırmanın dayanağı olan verilerin de kurumun hafızası olan kurumsal arşivlerde yer alması kütüphanenin sorumluluğunda olabilir. Mevcut durumda, kütüphaneler bir şekilde yayıncının izin verdiği ölçüde araştırma çıktılarının tam metinleri veya sadece üst verilerini sisteme yüklemekte veya yüklenmesini sağlamaktadır. Kütüphaneciler yayına bir şekilde ulaşabilir ancak veriye kendi gayretleriyle ulaşmaları mümkün değildir. Verilerin kurumsal arşivlere yüklenmesi için araştırmacı ve kütüphaneci koordineli bir biçimde, iş birliği içinde çalışmalıdır. Söz konusu iş birliği üst yönetimden gelecek destek ve yaptırım ile mümkün olabilir. Öncelikle kütüphanelerde gerekli iş gücüne, diğer bir deyişle konuya hâkim eğitilmiş kütüphanecilere ihtiyaç vardır. YÖK tarafından veya YÖK-TÜBİTAK iş birliğiyle konunun uzmanları tarafından kütüphaneciler için eğitimler düzenlenebilir. Araştırma verileri yönetimi konusunda belirli bir yol kat etmiş olan üniversite kütüphanelerinin deneyiminden yararlanılabilir. Örneğin, Koç Üniversitesi Kütüphanesi'nin uygulamaları yol gösterici olabilir. Ayrıca kütüphaneler hedef kitlelerinin ihtiyaçlarını belirlemek için

çeşitli çalışmalar yürütebilir ve bulgularla belirlenen ihtiyaçlara yönelik hizmetler geliştirebileceği gibi eğitimler de düzenleyebilir.

Üst veri bilgilerini kaydeden araştırmacılar çoğunlukla herhangi bir standart kullanmamaktadır. Standart kullananlar ise çoğunlukla -sadece kısıtlı bir gruba hitap eden- laboratuvarlarında standartlaştırılmış üst veriyi tercih etmektedir. Üst veride standart kullanımı; araştırma verileri hakkındaki bilgilerin makinece okunabilir olmasını, erişilebilir ve keşfedilebilir olmasını ve karşılıklı işlerliği sağlamaktadır. Aynı zamanda, özellikle disiplinler arası çalışmalarda ortak bir dilde buluşmayı sağlar. Üst veride standardın yaygınlaşması için disipline özgü, mesleki gelişim bazlı bir eğitim düzenlenebileceği gibi disiplinler arası çalışmanın yaygınlaştığı günümüzde daha kapsayıcı genel bir üst veri standardı eğitimi de düzenlenebilir. Veya konu uzmanı kütüphaneciler, standart üst veri kullanımı konusunda araştırmacılara yardımcı olabilir.

Araştırma verilerinin yönetilmesinde kolaylık sağlayan ve araştırmacıların bir kısmı tarafından kullanılan sürüm (versiyon) bilgisi, veri sözlükleri ve beni oku dosyalarının kullanımı teşvik edilebilir. Herhangi bir hatada bir önceki dosyaya geçilebilmesi zaman ve emek kaybını önleyeceği için dosya adlandırma sürüm numarası kullanılması önerilmektedir. Sürüm numarası ekleme, klasik yöntemlerle yapılabileceği gibi değişikliklerin izlenmesine olanak sağlayan Google Drive veya versiyon siteleri de kullanılabilir. Kullanılan kodların anlamlarını içeren veri sözlükleri ve araştırmanın yöntemi, verilerin düzeni, yapısı gibi bilgileri içeren beni oku dosyası da mevcut düzenin anlaşılması için yararlıdır. Ayrıca veri paylaşımı durumunda diğer araştırmacılara yol gösterici bir rehberdir. Veri arşivlerine, yükleme yapılabilmesi için sürüm bilgisi eklenmesi, veri sözlükleri kullanımı, beni oku dosyası oluşturulması ön şart olabilir ve bunları hazırlama sürecinde kütüphaneciler araştırmacılara rehberlik edebilir.

Araştırmacıların yarısından fazlası TÜBİTAK Açık Bilim Politikasının farkında değildir. Politika belgesi farkındalığı geçen zamanda artmış olabilir. Ancak eldeki verilere göre, duyuruların ve/veya hizmet tanıtımlarının yetersiz kaldığı düşünülebilir. Bir katılımcı tarafından yapılan yorumda, bilgilendirmenin sadece yazı göndererek yapıldığı belirtilmekte ve daha bilgilendirici toplantıların konu uzmanları tarafından yapılması önerilmektedir. Bilgilendirmelerin farklı yollarla daha etkin bir biçimde yapılması,

araştırmacıya ulaşılabilecek her mecranın kullanılması önerilebilir. Ulusal politikalarla açık bilim desteklenmeli, henüz bir veri paylaşım ve yönetim politikasına sahip olmayan üniversiteler gerekli düzenlemeleri yapmalıdır. Araştırmacılar veri paylaşımına yönelik olumlu düşüncelere sahip olmakla birlikte, çoğunlukla verilerini etik ve yasal koşullar elverdiğince paylaşmaktadır. Kişisel veriler veya hassas verilerde anonimleştirme (anonymization) veya takma ad (pseudonymization) gibi yöntemler kullanılarak hassasiyet ortadan kaldırılabilceği konusunda bilgilendirme ve eğitimler faydalı olabilir.

Araştırmacılar verilerini belirli bir kesimle (istek olması halinde, araştırma grubundakiler, bilimsel camia gibi) paylaşmakta, herkese açık olarak sunmamaktadır. Kamuya açık bir veri paylaşımı söz konusu olmadığından öncelik araştırmacıları verilerini paylaşmaya teşvik etmektir. Araştırmacılar, veri paylaşımının önemini, sağlayacağı yararları bildiği ve buna inandığı sürece verilerini paylaşacaktır. Verilerin lisanslanarak herkese açık bir şekilde yayımlanması için araştırmacı teşvik edilmelidir. Ayrıca, örneğin fon sağlamanın şartı olarak araştırma verilerini açma zorunluluğu getirilebileceği gibi akademik yükseltmelerde de benzer bir uygulamaya gidilebilir.

Araştırmacılar depolama sorunlarını, verilerin kaybolmasını ve güvenlik sorunlarını araştırma verileri yönetimindeki temel problemler olarak görmektedir. Bu sorunlar veri arşivlerinin kullanılmasıyla çözülebilir. Depolama sorununu araştırmacı düşünmemeli, fon sağlayıcı kuruluş ve/veya üniversite depolama ve uzun süreli koruma için gerekli teknik alt yapıyı sağlamalıdır.

Araştırmacılar çoğunlukla AVY konusunda eğitime gereksinim duymaktadır. Disipline göre ihtiyaçların değiştiği düşünüldüğünde, makro düzeyde tüm araştırmacıları kapsayan genel bir eğitim, mikro düzeyde ise disipline özgü bir eğitim söz konusu olabilir. En az eğitim eksikliği hissedilen konu veri atıf stilleridir. Ancak dış kaynaklı verinin kullanımı arttığında bu konuda da eğitim ihtiyacı hissedilebileceği düşünülmektedir. Araştırma verilerinin yönetimi konusunda uzmanlaşmış veri kütüphanecileri araştırmacılara eğitim verebilir.

Genel anlamda AVY ile ilgili katkı sağlamak üzere fikrini sunan bir araştırmacı TÜBİTAK'ın kullanıcılara ulaşmada ne yazık ki başarılı olmadığını, TÜBİTAK'ın kullandığı VYP sistemini tesadüfen fark ettiğini ifade etmiştir. Hizmetin verilmesi kadar

duyurulması ve tanıtılması da önemlidir. Bir şeyi zorunlu kılmadan önce sağlayacağı faydaların anlatılması, işin özümseyerek yapılmasını sağlayacaktır. Akabinde nasıl yapılacağı da anlatılmalıdır.

Araştırma evreninin hem URAP sıralamasında ilk beşte yer alan üniversitelerin araştırma kültürünü almış hem de TÜBİTAK projelerinde yer almış araştırmacılardan oluştuğu düşünüldüğünde, AVY konusundaki eksikliklerin genelde daha fazla olabileceği yönünde yorum yapmak mümkündür. Bu sürecin geliştirilebilmesi için TÜBİTAK tarafından yapılması gerekenleri ve araştırmacılardan beklenenleri içeren öneriler aşağıda sunulmuştur.

TÜBİTAK: Daha önce de belirtildiği gibi, araştırma verilerinin yönetimi konusunda TÜBİTAK tarafında bazı güzel gelişmeler yaşanmıştır. Ayrıca, 2022 yılında Açık Bilim Politikasının revize edilmesi beklenmektedir (Satoğlu, 2021). Bu yasal düzenlemeyle fon sağlayıcı kuruluşun hibe gereksinimleri veri paylaşımını kapsayacak şekilde geliştirilebilir. Politika belgesine, VYP hazırlama ve uygulama, verilerin belirli bir süre saklanması, verilerin ilgili yayınla ve/veya proje raporuyla bağlantılı ve FAIR ilkeleriyle uyumlu bir şekilde yayımlanması ilkeleri eklenebilir. Aynı zamanda, verilerin gerekçe gösterilmeden yayımlanmaması durumunda yeni proje başvurusu yapamayacak şekilde bir yapılanmaya gidilebilir.

Araştırma verilerinin yönetimi ile ilgili araştırmacılara destek olacak bir veri yönetimi birimi kurulabilir. Öncelikle destek personeli konunun uzmanları tarafından eğitim almalı, yeterli donanım ve beceriye sahip olduktan sonra bu personel araştırmacılara yol göstermelidir. Projelerde araştırma verilerinin yönetilmesi için bir veri yöneticisi görevlendirilebilir.

TÜBİTAK tarafından büyük verinin sağlayacağı olanakların keşfedilmesi ve araştırmacıların büyük veri üzerinde çalışmaya yönlendirilmesi sağlanabilir, farklı disiplinlerdeki araştırmacılar bir araya getirilerek iş birliğine dayalı bir araştırma ortamı yaratılabilir. Verilerin paylaşılması için araştırmacı teşvik edilebilir. Açık veri, açık bilim çalışmalarından ayrı düşünülemez. Araştırmacılara, Açık Erişim Ödülü veya Araştırma Verileri Erişim Ödülü verilebilir.

Araştırmacıların bilgi ihtiyacını karşılamak için TÜBİTAK web sayfalarının (ULAKBİM, Aperta, ARBIS gibi) geliştirilmesine ihtiyaç vardır. Araştırma verilerinin yönetimi ile ilgili bir web sayfası oluşturulabilir. Web sayfasında rehberler, araçlar, eğitim, SSS, politika belgesi ve APERTA'ya ait alt başlıklar ve/veya menüler bulunabilir. Her bir alan altında şu bilgiler olabilir:

- *Rehberler:* VYP hazırlama, araştırma verileri, dosya türleri, dosya yapısı ve organizasyon, dokümantasyon ve üst veri, yasal ve etik konular, veri saklama, depolama, uzun süreli koruma, veri yayımlama ile ilgili eğitici dokümanları kapsamalıdır.
- *Araçlar:* Araştırma öncesinde plan hazırlama, araştırma sırasında verileri saklama ve yedekleme, araştırma sonrasında ise yayımlama ve arşivleme ile ilgili araçlar bu bölümde yer alabilir.
- *Eğitim:* Hem araştırma verilerinin yönetimi hem de teknik konularda eğitim verilebilir. Benzer özelliklere sahip gruplar (disiplin, yaş gibi) bir araya getirilerek eğitim verilebileceği gibi birebir eğitim vermek veya eğitim videolarından yararlanmak da mümkün olmalıdır. Araştırmacılar vakitleri olduğunda veya ihtiyaç anında bunları kullanabilir. Eğitime katılmak için web sayfasında bir talep formu yer alabilir. AVY konusunda eğitimler veri yaşam döngüsünün tüm aşamalarını kapsayabileceği gibi sadece belirli aşamalar için de eğitim verilebilir. Eğitim şekli yüz yüze veya çevrimiçi olabilir. Araştırmacılara uygulanacak olan mini bir test ile kişinin ihtiyacına uygun konuların yer aldığı eğitim materyallerine (rehberler, videolar gibi) yönlendirmeler yapılabilir.
- *Teknik konularda eğitim:* Araştırmacıların işini kolaylaştıracak birçok veri analizi, görselleştirme, düzenleme programı bulunmaktadır. Her araştırmacının yeni teknolojilere yatkın olmaması veya bu gelişmeleri takip etmemesi nedeniyle araştırmacılara kolaylık sağlayacak araçların kullanımı konusunda bilgilendirme ve/veya destek konusunda TÜBİTAK araştırmacılara yardımcı olabilir.
- *SSS (Sıkça Sorulan Sorular):* Araştırma verilerinin yönetimi sürecindeki tüm aşamalara göre bir sınıflandırma yapılarak her bir aşama altında ilgili sorulara ve yanıtlara yer verilebilir.
- *Dış kaynaklara bağlantı:* Araştırmacıya yardımcı olabileceği düşünülen MANTRA, Eğitim Portalı gibi dış kaynaklara yönlendirmeler yapılabilir.

- *İletişim:* Sosyal medya, canlı destek, yardım masası gibi çeşitli iletişim kanallarıyla kolayca ulaşılabilir olmak sağlanabilir. Araştırmacının ihtiyaç duyduğu anda erişilebilir olmak önem taşımaktadır. TÜBİTAK web sayfasında iletişim bilgileri yeterli değildir. Sıkıntı yaşanması durumunda kiminle hangi kanalla iletişim kurulacağı belirsizdir. Birçok kanaldan erişilebilir olmak araştırmacıya yardımcı olacaktır.

Araştırmacı: Araştırmacı, araştırma verilerini yönetirken her bir aşamayı ayrı ayrı düşünerek planlamalıdır. İlgili aşamalarda yapılması gerekenler aşağıda verilmiştir.

- *Planlama:* Araştırmacı, araştırma öncesinde veri yaşam döngüsündeki aşamaları düşünerek bir planlama yapmalı ve proje başvurusunda veri yönetim planını yazılı bir şekilde sunmalıdır. Planda yer alması gereken unsurlar [Bölüm 3.1](#)'de sunulmuştur.
- *Veri Toplama ve Düzenleme:* Veri edinme aşamasında kullanılacak verinin veri arşivlerinde olup olmadığı kontrol edilmelidir. Yeni veri toplanacaksa; veri elde etme aşamasında kullanılacak yöntemler, verinin türü, formatı miktarı, klasör düzeninin nasıl olacağı, dosya adlandırma kuralı, tanımlayıcı bilgiler, üst veri oluştururken kullanılacak standart gibi bilgiler planda yer almalı ve uygulanmalıdır. Kişilerden veri toplanacaksa, aydınlanmış onam formu alınmalıdır. Var olan veri kullanılacaksa, keşfetme alanında sunulan hususlara dikkat edilmelidir (bkz. [Bölüm 3.5](#)). Verinin düzenlenmesi, belgelenmesi ve/veya keşfedilmesi aşamasında kütüphaneden destek alınabilir.
- *Depolama ve Koruma:* Veri yönetim planında belirlenen sorumluluklara göre, olası riskler gözetilerek uygun depolama ortamı, saklama koşulları ve yedekleme stratejisiyle veri güvenliği sağlanmalıdır.
- *Yayımlama ve Paylaşma:* Üniversite tarafından finanse edilip edilmediğine bakılmaksızın, üniversite mensupları tarafından yapılan tüm araştırmalarda araştırmacı verilerini makinece okunabilir, anlaşılır, doğru, eksiksiz, güvenilir ve koşullar uygun olduğu sürece paylaşılabilir halde tutmakla yükümlü olmalıdır. Araştırmanın dayanağı olan araştırma verisinin hesap verilebilirlik, doğrulanabilirlik, şeffaflık için saklanması ve paylaşılması gerekir. Etik ve yasal koşullara dikkat ederek, gerekirse anonimleştirme yaparak veri/veri setleri

lisanslanmalıdır. Veri toplama aşamasında yapılan dokümantasyon işlemlerinden yararlanılarak veri arşivine yüklenmek üzere üst veriler ve verilerin bağlamını açıklayan beni oku dosyası oluşturulmalıdır. Bu dosyalar verilerle birlikte bir bütün halinde sisteme yüklenerek yayımlanma süreci başlatılmalıdır. Yayımlanma işlemi, veri arşivi sorumlusu ve/veya kütüphaneci kontrolü ve onay vermesi sonrası gerçekleşmelidir.

- *Keşfetme ve Yeniden Kullanma:* Yeniden kullanım için, araştırmacı gerekirse kütüphaneden destek alarak veri arşivlerini kullanmalı ve mevcut açık verileri keşfetmelidir. Bulunan veri/veri setinin lisans ve kullanım şartları kontrol edilerek yeniden kullanılırken atıf verilmelidir.

KAYNAKÇA

- Açık Veri (2022). UAVS2021 Hakkında. Erişim adresi: <https://acikveri.org/ulusal-arastirma-verileri-sempozyumu/>
- Akers, K. G. ve Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, 8(2), 5-26. doi:10.2218/ijdc.v8i2.26
- Al, U. ve Soydal, İ. (2015). Bilimsel iletişimin farklı bir yüzü: Geri çekilen makaleler. U. Al ve Z. Taşkın (Yay. haz.). *Prof. Dr. İrfan Çakın'a armağan* içinde (s. 22–37). Ankara: Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü. Erişim adresi: <http://yunus.hacettepe.edu.tr/~umutal/publications/retraction.pdf>
- Allard S. ve Aydınoğlu A.U. (2012). Environmental researchers' data practices: An exploratory study in Turkey, S. Kurbanoglu, U. Al, P.L Erdoğan, Y. Tonta, N. Uçak (ed.) *E-Science and Information Management. IMCW 2012* içinde (s. 13-24). Communications in Computer and Information Science, vol. 317. Springer, Berlin, Heidelberg.
- ANDS Guides. (2018). *Publishing and sharing sensitive data*. Erişim adresi: <https://www.ands.org.au/guides/sensitivedata>
- Araştırma Verileri Yönetimi Eğitim Portalı (t.y.). Erişim adresi: <https://acikveri.ulakbim.gov.tr/>
- Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., ...Wouters, P. (2004). Promoting access to public research data for scientific, economic, and social development. *Data Science Journal*, 3, 135-152.
- Aydınoğlu, A. U. (2019). Araştırma verileri eğitim portalı: veriyi tekrar kullanma. Erişim adresi: <https://acikveri.ulakbim.gov.tr/acik-veri-acik-bilim/bolum-7-veriyi-tekrar-kullanma/>
- Aydınoğlu A. U., Doğan, G. ve Taşkın, Z. (2017). Research data management in Turkey: Perceptions and practices, *Library Hi Tech*, 35 (2), 271-289. doi: 10.1108/LHT-11-2016-0134
- Berkeley Library. (t.y.). Data backup. Erişim adresi: <https://researchdata.berkeley.edu/services/data-backup>
- Berman, E. (2017). An exploratory sequential mixed methods approach to understanding researchers' data management practices at UVM: Integrated findings to develop research data services. *Journal of Esience Librarianship*, e1104. doi: 10.7191/jeslib.2017.1104
- Bote, J. J. ve Termens, M. (2019). Reusing data: Technical and ethical challenges. *DESIDOC Journal of Library & Information Technology*, 39(6), 329-337. doi: 10.14429/djlit.39.6.14807
- Briney, K. (2015). *Data management for researchers: Organize, maintain and share your data for research success*. Exeter: Pelagic Publishing, UK.
- Briney, K., Goben, A. ve Zilinski, L. (2015). Do you have an institutional data policy? A review of the current landscape of library data services and institutional data policies. *Journal of Librarianship and Scholarly Communication*, 3(2), eP1232. <http://dx.doi.org/10.7710/2162-3309.1232>
- Bulk Rename Utility. (2021). Erişim adresi: <https://www.bulkrenameutility.co.uk/>

- Burnham, A. (2012). University of Leicester research data – Definitions. Erişim Adresi: https://www2.le.ac.uk/services/research-data/old-2019-12-11/documents/UoL_ReserchDataDefinitions_20120904.pdf
- Buys, C. M. ve Shaw, P.L. (2015). Data management practices across an institution: Survey and report. *Journal of Librarianship and Scholarly Communication* 3(2). DOI: 10.7710/2162-3309.1225.
- Cahit Arf Bilgi Merkezi. (2017). TÜBİTAK ULAKBİM Ulusal Açık Bilim Komitesi. Erişim adresi: <http://cabim.ulakbim.gov.tr/ulakbim-acik-erisim-faaliyetleri/ulakbim-ulusal-acik-bilim-komitesi>
- Chawinga, W. D. ve Zinn, S. (2019). Global perspectives of research data sharing: A systematic literature review. *Library & Information Science Research*, 41(2), 109-122.
- Concordat on open research data*. (2015). Versiyon 10, 17 Temmuz 2015. Erişim adresi: <http://www.oulu.fi/yliopisto/sites/default/files/content/ConcordatOpenResearchData.pdf>
- Cornell University. (t.y.). Data citation. Erişim adresi: <https://data.research.cornell.edu/content/data-citation>
- Cornell University Research Data Management Service Group (t. y.). Metadata and describing data. Erişim adresi: <https://data.research.cornell.edu/content/writing-metadata>
- Cox, A., ve Verbaan, E. (2018). What are research data? *Exploring research data management* içinde (s. 19-32). Facet. doi:10.29085/9781783302802.003
- Cox, A. M., Kennan, M. A., Lyon, L. ve Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. *Journal of the Association for Information Science and Technology*, 68(9), 2182-2200. <https://doi.org/10.1002/asi.23781>
- Creamer, A., Morales, M. E., Crespo, J., Kafel, D. ve Martin, E. R. (2012). An assessment of needed competencies to promote the data curation and management librarianship of health sciences and science and technology librarians in New England. *Journal of eScience Librarianship*, 1(1), 18–26. <http://dx.doi.org/10.7191/jeslib.2012.1006>.
- Creative Commons Türkiye (2019a, 13 Mart). #CCTRwebinar: FAIR Veri Prensipleri ve Araştırma Verilerinin Yönetimi [Blog yazısı]. Erişim adresi: <http://creativecommons.org.tr/cctrwebinar-fair-veri-prensipleri-ve-arastirma-verilerinin-yonetimi/>
- Creative Commons Türkiye (2019b, 17 Nisan). #CCTRwebinar: Türkiye’de açık veri politikalarının geliştirilmesi [Blog yazısı]. Erişim adresi: <http://creativecommons.org.tr/cctrwebinar-turkiyede-acik-veri-politikalarinin-gelistirilmesi-2/>
- Curty, R. G. (2015). *Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences* (Yayın No. 3713677). [Doktora tezi, Syracuse University]. ProQuest Dissertations & Theses Global.
- The data harvest: how sharing research data can yield knowledge, jobs and growth. An RDA Europe report (2014). Erişim adresi: <https://rd-alliance.org/sites/default/files/attachment/The%20Data%20Harvest%20Final.pdf>
- DCC. (2017). Overview of funders' data policies. Erişim adresi: <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

- Digital Archiving Consultancy, Bioinformatics Research Centre, & National e-Science Centre. (2005). *Large-scale data sharing in the life sciences: Data standards, incentives, barriers and funding models*. Erişim adresi: <https://eprints.gla.ac.uk/14140/>
- DMPOnline. (2020). Erişim adresi: <https://dmptool.org/>
- DMPTool. (2020). Erişim adresi: <http://dmponline.ac.uk>
- Doğan, G. (2019). Araştırma verileri eğitim portalı: mevcut verilerin saptanması. Erişim adresi: <https://acikveri.ulakbim.gov.tr/acik-veri-acik-bilim/bolum-2-arastirma-verisi-hazirlama-sureci/2-4-mevcut-verilerin-saptanmasi/>
- Doğan, G., Taşkın, Z. ve Aydınoglu, A. U. (2021). Research data management in Turkey: A survey to build an effective national data repository. *IFLA journal*, 47(1), 51-64.
- Donnelly, M. (2012). Data management plans and planning. G. Pryor (Ed.) *Managing research data* içinde (s. 83-104). London: Facet.
- Dryad. (2020). Erişim adresi: <https://datadryad.org/>
- EDUCAUSE. (2020). *Guidelines for Data De-Identification or Anonymization*. Erişim adresi: <https://www.educause.edu/focus-areas-and-initiatives/policy-and-security/cybersecurity-program/resources/information-security-guide/toolkits/guidelines-for-data-deidentification-or-anonymization#>
- European Commission. (2012). *Online survey on scientific information in the digital age*. Erişim adresi: https://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf
- European Commission. (2016). *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020*. Erişim adresi: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Figshare. (2022). Erişim adresi: <https://figshare.com/>
- FORCE11. (2011-2021). *The FAIR Data Principles*. Erişim adresi: <https://force11.org/info/the-fair-data-principles/>
- The Graduate Institute. (2020). *Folder Structure*. Erişim adresi: <https://libguides.graduateinstitute.ch/rdm/folders>
- Grove, J. (2020, 12 Şubat). Scientists offered €1,000 to publish null results. Erişim adresi: <https://www.timeshighereducation.com/news/scientists-offered-eu1000-publish-null-results>
- Gürdal, G. (2020, 16 Ocak). Re: Yükseköğretim Araştırma Verisi ve Açık Veri Alt Çalışma Grubu [Elektronik mesajlaşma listesi mesajı]. Erişim adresi: <https://mailman.metu.edu.tr/mailman/private/kutup-1/2020-January/013526.html>
- Gürdal, G. ve Bitri, E. (2015). *Araştırma Verisi Yönetimi, Açık Veri ve Avrupa Birliği Bilimsel Veri Altyapısı: OpenAIRE2020, XVII. Akademik Bilişim Konferansı Eskişehir, 4-6 Şubat 2015*. Erişim adresi: <http://hdl.handle.net/11147/4265>

- Hacettepe Üniversitesi Biyoinformatik Anabilim Dalı (t.y.). *Açık Veri Yönetimi ve Paylaşımı Çalıştayı Programı*. Erişim adresi:
http://www.biyoinformatik.hacettepe.edu.tr/tr/acik_veri_yonetimi_ve_paylasim-20
- Hampton, S., Anderson, S., Bagby, S., Gries, C., Han, X., Hart, E.M., ...Zimmerman, N. (2015). The tao of open science for ecology. *Ecosphere*, 6(7). doi: 10.1890/ES14-00402.1.
- Harvard Medical School. (2022). *Biomedical Data Lifecycle*. Erişim adresi:
<https://datamanagement.hms.harvard.edu/about/what-research-data-management/biomedical-data-lifecycle>
- Hey, T. ve Trefethen, A. (2003). The data deluge: An e-science perspective. *Grid computing: Making the global infrastructure a reality*, 809-824.
- Hey, T. ve Trefethen, A. (2008). E-science, cyber-infrastructure, and scholarly communication. G. M. Olson, A. Zimmerman ve N. Bos (Ed.) *Scientific collaboration on the Internet* içinde (s. 15-31).
- Hswe, P. ve Holt, A. (2010). *Guide for Research Libraries: the NSF data sharing policy*, Association for Research Libraries. Erişim adresi: <http://old.arl.org/rtl/eresearch/escien/nsf/index.shtml>
- Holt, İ. (2018, 25 Ekim). Creative Commons lisanslarının açık bilimdeki yeri. Erişim adresi:
<https://zenodo.org/record/1471568#.X9s6X9gzbIU>
- IASSIST SIGDC. (2012). Quick guide to data citation. Erişim adres:
<https://www.icpsr.umich.edu/files/ICPSR/enewsletters/iassist2.html>
- IWGDD. (2009). *Harnessing the power of digital data for science and society*. Washington D.C., USA: Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council. Erişim adresi: http://www.nesc.ac.uk/technical_papers/UKeS-2006-02.pdf
- İYTE (2019). *Açık Bilim Yolunda Semineri: Program*. Erişim adresi:
<https://libguides.iyte.edu.tr/acikbilimiyolunda>
- James Cook University. (2020). *Research data management toolkit: Retention and preservation*. Erişim adresi: <https://libguides.jcu.edu.au/research-data-management-toolkit/retention-preservation>
- Jones, S. (2012). Research data policies: principles, requirements and trends. G. Pryor (Ed.) *Managing research data* içinde (s. 47-66). London: Facet.
- Kahn, M., Higgs, R., Davidson, J. ve Jones, S. (2014). Research data management in South Africa: how we shape up. *Australian Academic & Research Libraries*, 45 (4), s. 296-308.
- Kaiseer, J. (2008). NIH reports breach of patient records. *Science*, 319 (5871), s. 1746
doi: 10.1126 / science.319.5871.1746b
- Karaçay, B. (2018). *Mutlu Beyin*. Ankara: TÜBİTAK.
- Kennan, M. A. ve Markauskaite, L. (2015). Research data management practices: A snapshot in time. *International Journal of Digital Curation*, 10 (2), s. 69-95.
- Kroes, M. N. (2012). *A declaration of ALL European academies*. Erişim adresi:
https://www.alllea.org/wp-content/uploads/2015/09/OpenScience-Rome-Declaration-final_web.pdf

- Kroll, S. ve Forsman, R. (2010). *A Slice of Research Life: Information Support for Research in the United States*. Dublin, Ohio: OCLC Research. Erişim adresi: <http://www.oclc.org/research/publications/library/2010/2010-15.pdf>
- Liu, X. ve Ding, N. (2016). Research Data Management in Universities of Central China Practices at Wuhan University Library. *The Electronic Library*, 34 (5), s. 808 – 822.
- Lord, P. ve Macdonald, A. (2003). *E-Science curation report: Data curation for e-Science in the UK: An audit to establish requirements for future curation and provision*. Digital Archiving Consultancy Limited.
- Mantra Research Data Management Training. (2020a). Documentation, metadata, citation. Erişim adresi: https://mantra.edina.ac.uk/documentation_metadata_citation/
- Mantra Research Data Management Training. (2020b). Organising data: what's in a name? Erişim adresi: <https://mantra.edina.ac.uk/organisingdata/>
- Mantra Research Data Management Training. (2020c). Protecting sensitive data. Erişim adresi: <https://mantra.edina.ac.uk/protectionrightsandaccess/>
- Mantra Research Data Management Training. (2020d). Research data in context: what are research data? Erişim adresi: <https://mantra.edina.ac.uk/researchdataexplained/>
- Mantra Research Data Management Training. (2020e). Why manage your data? Erişim adresi: <https://mantra.ed.ac.uk/datamanagementplanning/>
- Michigan State University Libraries. (2021). *LibGuides: How to cite data: general info*. Erişim adresi: <https://libguides.lib.msu.edu/citedata>.
- Molloy J. C. (2011). The Open Knowledge Foundation: Open Data Means Better Science. *PLoS Biology*. 9(12): e1001195. doi:10.1371/journal.pbio.1001195
- NASA (2019). *Mars Climate Orbiter*. Erişim adresi: <https://solarsystem.nasa.gov/missions/mars-climate-orbiter/in-depth/>
- National Institutes of Health (NIH). (2003). *NIH Data Sharing Policy and Implementation Guidance*. Erişim adresi: https://grants.nih.gov/grants/policy/data_sharing/
- National Science Foundation. (2010). Scientists Seeking NSF funding will soon be required to submit data management plans. *News Release 10-077*. Erişim adresi: https://www.nsf.gov/news/news_summ.jsp?cntn_id=116928
- National Science Foundation. (2011). *Digital research data sharing and management*. Erişim adresi: www.nsf.gov/nsb/publications/2011/nsb1124.pdf
- Nicol, A., Caruso, J. ve Archambault, É. (2013). Open data access policies and strategies in the European research area and beyond. *info@ science*, 1, 495-6505.
- OpenAIRE Research data management briefing paper (2016). Erişim adresi: https://dspace.vsb.cz/bitstream/handle/10084/111436/OpenAIRE_RDM_Briefing_Paper_Master.pdf?sequence=4&isAllowed=y
- OpenAIRE Ufuk 2020. (t.y.). *Ufuk 2020 Açık araştırma veri pilot: OpenAIRE nasıl yardımcı olabilir?* Erişim adresi: http://acikerisim.ankos.org.tr/belgeler/FactSheet_OR_Data_Pilot_Turkish.pdf

- Organisation for Economic Co-operation and Development. (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Erişim adresi: <http://www.oecd.org/science/sci-tech/38500813.pdf>
- Patel, D. (2016). Research data management: a conceptual framework. *Library Review*, 65 (4/5), s. 226-241.
- Perrier L., Blondal E., Ayala A. P., Dearborn D., Kenny T., Lightfoot D., ... MacDonald, H. (2017) Research data management in academic institutions: A scoping review. *PLoS ONE* 12(5): e0178261. <https://doi.org/10.1371/journal.pone.0178261>
- Peters, C. (2017, Ocak 26). *Research data management : basics & best practices*. Erişim adresi: https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1000&context=rdsc_workshops
- Piwowar, H., Day, R. ve Fridsma D. (2007). Sharing detailed research data is associated with increased citation rate, *PLoS ONE*, 2 (3), e308. Erişim adresi: www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0000308
- Piwowar, H. A., Vision, T. J. ve Whitlock, M. C. (2011). Data archiving is a good investment. *Nature*, 473(7347), 285–285.
- PLoS One. *Data availability* (2019). Erişim adresi: <https://journals.plos.org/plosone/s/data-availability>
- Polat, C. (2006). Bilimsel bilgiye açık erişim ve kurumsal açık erişim arşivleri. *Atatürk Üniversitesi Fen Edebiyat Fakültesi Sosyal Bilimler Dergisi*, 6(37), 53-80. Erişim adresi: <http://eprints.rclis.org/9029/>
- Princeton University Library. (2020). Research Data Management at Princeton: Sharing and Preservation. Erişim adresi: <https://libguides.princeton.edu/c.php?g=102546&p=665868>
- Pryor, G. (2012). Why manage research data? G. Pryor (Ed.) *Managing research data* içinde (s. 1-16). London: Facet.
- Ray, J. M. (2014). *Research Data Management Practical Strategies for Information Professionals*. M. J. Ray (ed.) Introduction to Research Data Management, s. 12-29. Indiana: Purdue University.
- re3data.org - Registry of Research Data Repositories (t.y.). Erişim adresi: <https://doi.org/10.17616/R3D>
- Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 : establishing Horizon 2020 – the Framework Programme for Research and Innovation (2014–2020) and repealing Decision No 1982/2006/EC. Erişim adresi: https://ec.europa.eu/home-affairs/system/files/2020-09/h2020-eu-establact_en.pdf
- ReNamer (t.y.). Erişim adresi: <https://www.den4b.com/products/renamer>
- Renwick, S., Winter, M. ve Gill, M. (2017). Managing research data at an academic library in a developing country. *IFLA Journal*, 43(1), 51-64.
- Riley, J. (2017). *Understanding metadata: what is metadata and what is it for*. Baltimore: National Information Standards Organization. Erişim adresi: <https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf>

- Satoğlu, M. M. (2021, Mayıs). *TÜBİTAK veri yönetimi çalışmaları ve APERTA* [sunum]. Ulusal Araştırma Verileri Sempozyumu (UAVS), 25-26 Mayıs 2021. doi: 10.5281/zenodo.4896586
- Savage C. J. ve Vickers, A. J. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS ONE* 4(9): e7078. Erişim tarihi: 06.04.2020 doi: 10.1371/journal.pone.0007078
- Scaramozzino, J.M., Ramírez, M.L. ve McGaughey, K.J. (2012). A study of faculty data curation behaviors and attitudes at a teaching-centered university. *College & Research Libraries*, 73 (4), 349-365.
- Schneider, R. (2013). Research data literacy. S. Kurbanoglu ve diğerleri (Ed.), *Worldwide Commonalities and Challenges in Information Literacy Research and Practice: European Conference, ECIL 2013, Istanbul, Turkey, October 22-25, 2013* içinde s. 134-140. Springer International, Cham,
- Si, L., Xing, W., Zhuang, X., Hua, X. ve Zhou, L. (2015). Investigation and of research data services in university libraries. *The Electronic Library*, 33 (3), s. 417-449.
- Smale, N., Unsworth, K., Denyer, G. ve Barr, D. (2018). The history, advocacy and efficacy of data management plans. *bioRxiv*, 443499
- Spichtinger, D. ve Siren, J. (2017). The development of research data management policies in Horizon 2020. *Research data management - A European perspective* içinde s. 11-23. doi:10.1515/9783110365634-002
- Stanford Libraries. (t.y.). Best practices for file naming. Erişim adresi: <https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-naming>
- Stanford Libraries. (2022). Create metadata for your research project - guides at Stanford University. Erişim adresi: <https://guides.library.stanford.edu/research-metadata/metadata-tools>
- Steen, R. G. (2011). Retractions in the scientific literature: is the incidence of research fraud increasing? *Journal of Medical Ethics*, 37(4), 249–253. <https://doi.org/10.1136/jme.2010.040923>
- Stewart, J. ve Crossley, J. (2013) Library readiness for research data management. *ALISS Quarterly*, 8 (4). ss. 13-16. ISSN 17479258. Erişim adresi: <http://eprints.uwe.ac.uk/30294>
- Space.com (01 Nisan, 2009). Hidden planet discovered in old hubble data. Erişim adresi: <https://www.space.com/6527-hidden-planet-discovered-hubble-data.html>
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., ve Frame, M. (2011). Data sharing by scientists: practices and perceptions. *PloSone*, 6(6),e21101.
- Tenopir, C., Birch, B., ve Allard, S. (2012). *Academic libraries and research data services: current practices and plans for the future*. Chicago, IL: Association of College and Research Libraries.
- Tenopir, C., Sandusky, R. J., Allard, S ve Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research* 36, 84–90.
- Tonta, Y. (2012). *Açık erişim, kurumsal arşivler ve MedOANet projesi*. Ulusal Açık Erişim Çalıştayı, 8-9 Kasım 2012, Ankara. Erişim adresi: http://ae2012.acikerisim.org/wp-content/uploads/sites/6/2016/11/yasar_tonta.pdf

- Tonta, Y. (2013a). *Açık erişimin geleceği ve araştırma verilerine açık erişim*. Bilkent'te Kütüphanecilik Seminerleri. 17 Aralık 2013. Erişim adresi: library.bilkent.edu.tr/activities/librarianship-seminars/presentations/yasar-tonta.pptx
- Tonta, Y. (2013b). *Araştırma verilerinin yönetimi ve açık veri*. 2. Ulusal Açık Erişim Çalıştayı, 21-22 Ekim 2013, İzmir. Erişim adresi: <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-izmir-21-22ekim2013-open-access-research-data-management.pdf>
- Tonta, Y. (2015). *Açık erişim, PASTEUR4OA projesi ve konsorsiyal anlaşmalar*. ANKOSLink 2015, 9-11 Nisan 2015, Antalya. Erişim adresi: tarihinde <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-ankos2015-turkiyede-acik-erisim-calismalari-11-Nisan-2015-v2.pdf>
- Tonta, Y. (2016). *Açık bilim, açık erişim, açık veri* [sunum]. Açık Veri ve Açık Devlet Paneli. Kamu Politikaları Enstitüsü, Bursa, 4 Haziran 2016. Erişim adresi: <https://yunus.hacettepe.edu.tr/~tonta/yayinlar/tonta-acik-bilim-acik-erisim-acik-veri-kpe-4-haz-2016-bursa-v2.pdf>
- Tonta, Y. ve Akbulut, M. (2019). Türkiye'de lisansüstü tezlere açık erişim. *Türk Kütüphaneciliği*, 33(4), 219–248. Erişim adresi: http://yunus.hacettepe.edu.tr/~tonta/Yayinlar/tonta_akbulut_lisansustu_tezlere_acik_erisim.pdf
- Tonta, Y. ve Al, U. (2015). Araştırma verilerinin yönetimi. *Türk Kütüphaneciliği* 29 (1), s. 36-45. Erişim adresi: <http://www.tk.org.tr/index.php/TK/article/view/2520/2524>
- UK Data Archive. (2011). Managing and sharing data: best practice for researchers. <https://ukdataservice.ac.uk/media/622417/managingsharing.pdf>
- UK Data Service. (2020a). Data security. Erişim adresi: <https://www.ukdataservice.ac.uk/manage-data/store/security.aspx>
- UK Data Service. (2020b). Organising data. Erişim adresi: <https://www.ukdataservice.ac.uk/manage-data/format/organising.aspx>
- UK Data Service. (2020c). Recommended formats. Erişim adresi: <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>
- UKRI (2011). Common principles on data policy. Erişim adresi: <https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>
- The University of British Columbia. (2018). Data pseudonymization. Erişim adresi: <https://researchdata.library.ubc.ca/share/anonymize-and-de-identify/data-pseudonymization/>
- University of California Curation Center (2020). DMPTool. Erişim adresi: https://dmptool.org/general_guidance#types-of-data
- University of Cambridge. (2020a). Looking after and sharing your data. Erişim adresi: <https://www.data.cam.ac.uk/data-management-guide/looking-after-and-sharing-your-data>
- University of Cambridge. (2020b). Organising your data. Erişim adresi: <https://www.data.cam.ac.uk/data-management-guide/organising-your-data>

- University of Cape Town (2022). Research data management roles and responsibilities. Erişim adresi: <http://www.researchsupport.uct.ac.za/research-data-management-roles-and-responsibilities>
- University of Leicester (2019). Research data. Erişim adresi: <https://www2.le.ac.uk/services/research-data/old-2019-12-11/rdm/what-is-rdm/research-data>
- University of Minnesota. (2020). Creating a de-identified data set. Erişim adresi: <https://www.healthprivacy.umn.edu/policies-procedures/creating-de-identified-data-set>
- University of Minnesota Libraries. (2020a). File naming. Erişim adresi: <https://www.lib.umn.edu/datamanagement/file-naming>
- University of Minnesota Libraries. (2020b). Sensitive data. Erişim adresi: <https://www.lib.umn.edu/datamanagement/sensitive>
- URAP. (2018). URAP Türkiye özel bölümü. Erişim adresi: <http://tr.urapcenter.org/2018/>
- USGS. (t. y.). Why share your data. Erişim adresi: <https://www.usgs.gov/data-management/why-share-your-data>
- UWA Library. (2022). Guides: research data management toolkit: data collection. Erişim adresi: <https://guides.library.uwa.edu.au/RDMtoolkit/data-collection>
- Ünal, Y., Chowdhury, G., Kurbanoglu, S., Boustany, J. ve Walton, G. (2019). Research data management and data sharing behaviour of university researchers'. *ISIC: The Information Behaviour Conference* içinde, 09 October 2018 - 11 October 2019, Krakow, Poland.
- Ünal, Y. ve Kurbanoglu, S. (2018). Araştırma verilerinin yönetimi: Türk araştırmacılar verilerini açmaya hazır mı? *Türk Kütüphaneciliği*, 32(4), 287-311. doi: 10.24146/tkd.2018.42
- Van Loon, J.E., Akers, K.G., Hudson, C. ve Sarkozy, A. (2017). Quality evaluation of data management plans at a research university. *IFLA Journal*, 43 (1), s. 98-104.
- Van Uffelen, C. (2019). Another laptop stolen at TPM. Erişim adresi: <https://www.delta.tudelft.nl/article/another-laptop-stolen-tpm>
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L. ve Horton, L. (2011). Managing and sharing data: a best practice guide for researchers. Erişim adresi: <https://dam.ukdataservice.ac.uk/media/622417/managingsharing.pdf>
- Vilar, P. ve Zabukovec, V. (2019). Research data management and research data literacy in Slovenian science. *Journal of Documentation*, 75 (1), 24-43. doi: 10.1108/JD-03-2018-0042
- Vines, T.H., Albert, A.Y.K., Andrew, R.L., Débarre, F., Bock, D.G., Franklin, M.T...Rennison, D.J. (2014). The availability of research data declines rapidly with article age. *Current Biology*, 24(1), 94-97.
- Wakeling, S.G., Hannay, P. ve Baig, Z. (2017). A review of data breaches and losses that occurred from laptops that were stolen or otherwise misplaced in 2015 and 2016. Valli, C. (Ed.). *The Proceedings of 15th Australian Information Security Management Conference* içinde, 5-6 December, 2017, Edith Cowan University, Perth, Western Australia. (pp.97-107). Erişim adresi: <https://ro.ecu.edu.au/ism/212/>

- Wellcome Trust (2008). Q&A: Wellcome Trust policy on data management and sharing. Erişim adresi: <http://web.archive.org/web/20080810123841/http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-management-and-sharing/WTX035045.htm>
- Whitmire, A. L., Boock, M. ve Sutton, S. C. (2015). Variability in academic research data management practices implications for data services development from a faculty survey. *Program: electronic library and information systems*, 49 (4), s. 382 – 407.
- Whyte, A. ve Tedds, J. (2011). *Making the case for research data management*. Edinburgh: Digital Curation Centre. Erişim adresi: www.dcc.ac.uk/resources/briefing-papers/making-case-rdm
- Wiley. (2014). Researcher data sharing insights. <https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ...Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 1-9.
- Willetts, D., Livanov, D., Schütte, G., Harayama, Y., Carrozza, M.C., Goodyear, G., ...Geoghegan-Quinn, M. (2013). G8 Science Ministers Statement. Erişim adresi: <https://www.gov.uk/government/news/g8-science-ministers-statement>
- Zencir, M. B. (2019). *Ankara Üniversitesi akademisyenlerinin araştırma verilerinin yönetimi ile ilgili tutumları ve bir model önerisi* (Doktora tezi, Ankara Üniversitesi). Erişim adresi: <https://tez.yok.gov.tr/UlusalTezMerkezi/>
- Zenodo. (2020). Erişim adresi: <https://zenodo.org/>
- Zhang, M. (2016, 12 Nisan). *Writing data management plans* [Video]. Erişim adresi: <https://www.youtube.com/watch?v=ZJFtYmOlksc>
- Zhu, Y. (2020). Open-access policy and data-sharing practice in UK academia. *Journal of Information Science*, 46(1), 41-52.

EK 1. Anket Soruları

ARAŞTIRMA VERİLERİNİN YÖNETİMİ

Sayın Katılımcı;

Bu anket, HÜ Sosyal Bilimler Enstitüsü Bilgi ve Belge Yönetimi Bölümü'nde Prof. Dr. Serap Kurbanoglu'nun danışmanlığında yürütülen doktora tezinde kullanılmak üzere araştırma verilerinin yönetimi hakkında bilgi toplamak için hazırlanmıştır. Bu çalışmayla, TÜBİTAK tarafından 2014-2018 yılları arasında desteklenen araştırma projelerinde yer alan araştırmacıların, araştırma verilerinin yönetimi konusuna yönelik farkındalık düzeyleri, davranışları ve tutumlarını belirlemek amaçlanmaktadır. Anketin uygulanması için Hacettepe Üniversitesi Etik Komisyonundan gerekli izinler alınmıştır. Gerekli olduğu durumlarda diğer kurumlardan da izin alınacaktır.

Çalışmaya katılım gönüllülük esasına dayanmaktadır. Ankette kimlik belirleyici hiçbir bilgi istenmemekte ve araştırma kapsamı dışında hiçbir soru sorulmamaktadır. Katılım sırasında herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini yarıda bırakıp araştırmadan çekilme hakkına sahipsiniz. Bu durum size hiçbir sorumluluk getirmeyecektir.

Anketin yanıtlanması yaklaşık 10 dakika sürecektir. Araştırma kapsamında toplanan veriler gizli kalacaktır ve sadece bilimsel amaçlar doğrultusunda kullanılacaktır.

Bir araştırmacı olarak haklarımı anladım ve bu araştırmaya gönüllü olarak katılıyorum.

- Evet (Katılmayı kabul ediyorum)
- Hayır (Katılmayı kabul etmiyorum)

DEĞERLİ KATKILARINIZ İÇİN ÇOK TEŞEKKÜR EDERİM.

Araştırmacı

Canan TAVLUOĞLU

HÜ BBY Bölümü Doktora Öğrencisi

1. BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ KONUSUNDA FARKINDALIK VE UYGULAMA EĞİLİMLERİ

Araştırma verisinin tanımı: Araştırma çıktılarının temelini oluşturmada yardımcı olan toplanmış, üretilmiş, gözlenmiş veya yaratılmış her türlü basılı, fiziksel veya elektronik veridir. Örneğin,

- Belgeler (text, Word), elektronik tablolar, slaytlar
- Laboratuvar defterleri, saha defterleri, karalamalar, günlükler
- Anketler, transkripsiyonlar, kod kitapları
- Ses kayıtları, videolar
- Fotoğraflar, filmler
- Protein veya genetik diziler
- Test yanıtları
- Araştırma sürecinde sağlanan, toplanan dijital nesne koleksiyonu
- Veritabanı içerikleri (video, ses, metin, resimler)
- Modeller, algoritmalar, komut dosyaları
- Bir uygulamanın içeriği (input, output, logfiles)
- Metodolojiler ve iş akışları
- Standart işletim prosedürleri ve protokolleri
- İçerik analizi
- Deneysel veri
- Odak grup kayıtları; görüşme notları

1. Araştırmalarınızda elektronik ortamda toplanmış ve/veya üretilmiş dijital (sayısal) veri veya sonradan dijitalleştirilmiş veri kullanıyor musunuz? Araştırma verisi için yukarıdaki tanımı baz alınız.

- Evet
- Hayır (Son sayfaya yönlendirme)

2. Araştırmalarınızda kullandığınız verilerin dosya türlerini belirtiniz. (İlgili tüm seçenekleri işaretleyiniz)

- Standart ofis dokümanları (metin (.doc, .txt gibi), excel (.xls gibi), sunum (.ppt gibi))
- Yapılandırılmış bilimsel ve istatistiksel veri (SPSS (.sav, .sdq, .spv gibi), GIS (.gpx, .kml gibi), EViews (.wf1))
- Kodlanmış metin (.xml, .sgml)
- İnternet ve web-tabanlı veri (web sayfaları, e-postalar, bloglar, sosyal ağ verisi)
- Veri tabanları (Access, Oracle, MySQL)

- Görseller (.jpeg, .gif, .tiff, .png gibi)
- Ses dosyaları (.wav, .mp3 gibi)
- Videolar (.mov, .mts, .mp4 gibi)
- Yapılandırılmış grafikler (.cad, .cam, .vml gibi)
- Ham (makine tarafından üretilmiş) veri
- Arşivlenmiş veri (.zip, .rar gibi)
- Yazılım uygulamaları (modelleme araçları, editörler, derleyiciler)
- Kaynak kodu (komut, Java, C, C++ gibi)
- Yapılandırma verisi (parametre düzenlemeleri, loglar, kütüphane dosyaları gibi)
- Diğer (lütfen açıklayınız):

3. Araştırmalarınız için veriyi genellikle nasıl elde edersiniz? (İlgili tüm seçenekleri işaretleyiniz)

- Yeni veri toplarım/yaratırım
- Üniversitedeki araştırma ekibimden edinirim
- Kendi araştırma ağımdan edinirim
- Bilinen bir kaynaktan edinirim
- Bilinen birden fazla kaynaktan edinirim
- Diğer (lütfen açıklayınız):

4. Diğer kaynaklardan sağladığınız veriyi genellikle nasıl kullanıyorsunuz? (İlgili tüm seçenekleri işaretleyiniz)

- Diğer kaynaklardan veri kullanmıyorum
- Olduğu gibi sorunsuz bir şekilde
- Temizleme ve/veya değişiklikler için harcanan biraz çabayla
- Kullanılabilir hale getirmek için çok çaba ve zaman harcadıktan sonra
- Diğer (lütfen açıklayınız):

5. Aşağıdakilerden hangisi araştırmalarınızda kullandığınız verinin miktarını daha iyi tanımlar?

- Bilmiyorum
- 1 GB (gigabyte) veya daha az
- 1 GB - 100 GB
- >100 GB – 1 TB (terabyte)
- >1 TB – 100 TB
- >100 TB – 1 PB (petabyte)
- 1 PB'tan fazla
- Diğer (lütfen açıklayınız):

6. Araştırma verilerinizi araştırma sırasında nerede tutuyorsunuz? (İlgili tüm seçenekleri işaretleyiniz)

- Kişisel bilgisayar(lar)
- Bulut hizmetleri (Dropbox, Google Drive, iCloud gibi)
- Harici disk/Flash bellek

- Bölüm/Kurum sunucusu
- E-posta hesabımda
- Diğer (lütfen açıklayınız):

7. Verilerinizi hangi sıklıkta yedekliyorsunuz?

- Yedeklemiyorum
- Değişiklik yaptığım her an
- Günlük
- Haftalık
- Aylık
- Yılda birkaç kez
- Yıllık
- Birkaç yılda bir
- Diğer (lütfen açıklayınız):

8. Yaptığımız araştırma/proje bittikten sonra verilerinizi ne kadar süre saklıyorsunuz?

- Saklamıyorum
- 1 yıldan az
- 1-2 yıl
- 3-5 yıl
- 6-10 yıl
- Daima

9. Uzun süreli koruma, veri bütünlüğünün bozulmadan saklanması, depolanması ve veriye erişim sağlanması sürecidir. Uzun süreli koruma için verilerinizi nerede saklıyorsunuz? (İlgili tüm seçenekleri işaretleyiniz)

- Verilerimi uzun süreli korumuyorum
- Konusal/kurum dışı arşivler
- Kurumsal arşivler
- Kendi cihazlarımız (bilgisayar, tablet, harici disk gibi)
- Bulut hizmetleri (Dropbox, Google Drive, iCloud gibi)
- Açık erişimli bir veri havuzu
- Ticarî bir veri havuzu
- Kamuya açık bir web sitesi
- Basılı olarak
- Diğer (lütfen açıklayınız):.....

10. Uzun süreli koruma sizce kimin sorumluluğunda olmalıdır? (İlgili tüm seçenekleri işaretleyiniz)

- Kendi sorumluluğumda
- Proje ekibinin
- Kurumun/Üniversitenin
- Bölümün
- Fon sağlayan kuruluşun
- Kütüphanenin

- Yayıncının
- Hiç kimsenin
- Diğer (lütfen açıklayınız):

11. Verilerin kim tarafından, nasıl, ne zaman, nerede toplandığı gibi veri hakkında bilgileri içeren üst veri (metadata) bilgilerini kaydediyor musunuz?

- Evet, çalıştığım dosya üzerine kaydediyorum.
- Evet, üst veri bilgilerini ayrı bir dosyada tutuyorum.
- Hayır, çünkü gerekli bulmuyorum. (Lütfen 13. soruya geçiniz.)
- Hayır, çünkü zaman bulamıyorum. (Lütfen 13. soruya geçiniz.)
- Hayır, çünkü teknik desteğe ihtiyaç duyuyorum. (Lütfen 13. soruya geçiniz.)
- Diğer (lütfen açıklayınız)

12. Üst veri oluştururken herhangi bir bilimsel üst veri standardı kullanıyor musunuz?(İlgili tüm seçenekleri işaretleyiniz)

- Hayır, kullanmıyorum
- Evet, ABCD (Access to Biological Collections Data Schema) kullanıyorum.
- Evet, AVM (Astronomy Visualization Metadata Standard) kullanıyorum.
- Evet, CIF (Crystallographic Information Framework) kullanıyorum.
- Evet, CF (Climate and Forecast) kullanıyorum.
- Evet, CSDGM (Content Standard for Digital Geospatial Metadata) kullanıyorum.
- Evet, DC (Dublin Core) kullanıyorum.
- Evet, DwC (Darwin Core) kullanıyorum.
- Evet, DIF (Directory Interchange Format) kullanıyorum.
- Evet, EML (Ecological Metadata Language) kullanıyorum.
- Evet, FGDC Metadata Standard (Federal Geographic Data Committee) kullanıyorum.
- Evet, FITS (Flexible Image Transport System) kullanıyorum.
- Evet, ISO (International Standards Organization) kullanıyorum.
- Evet, NetCDF (Network Common Data Form) kullanıyorum.
- Evet, NISO MIX kullanıyorum.
- Evet, OGIS (Open GIS) kullanıyorum.
- Evet, OME-XML (Open Microscopy Environment-XML) kullanıyorum.
- Evet, laboratuvarımda standartlaştırılmış üst veri kullanıyorum.
- Evet, Diğer (lütfen belirtiniz):

13. Araştırmalarınıza ait verileri anlamlandırmak ve/veya aradığınız dosyayı kolayca bulmak için yapılan dokümantasyon (dosya adlandırma, veriyi tanımlama vb.) işlemlerine yönelik olarak aşağıdaki seçeneklerden lütfen size uygun olanlarını işaretleyiniz.

- Herhangi bir dokümantasyon işlemi yapmam.
- Beni oku dosyası kullanırım.
- Veri sözlükleri kullanırım.

- Çalışmama ait klasörler belirli bir düzen içerisinde (projeye göre, tarihe göre, veri türüne göre vb)
- Aradığım dosyayı kolay bulmak için etiketleme yaparım.
- Dosya adlandırırken belirli bir kurala uyarak tutarlı adlandırma yaparım.
- Dosya adlandırmada sürüm (versiyon) numarası kullanırım
- Diğer (lütfen belirtiniz):

14. Veri Yönetim Planı, araştırma projesi yürütülürken veya tamamlandığında, elde edilen araştırma verilerinin nasıl yönetileceğine (veri nasıl toplanacak, nasıl düzenlenecek, nerede saklanacak, paylaşılacak mı) dair bilgiler içeren bir belgedir. Araştırmalarınız için hiç “Veri Yönetim Planı” hazırladınız mı?

- Evet
- Hayır (Lütfen 17. soruya geçiniz.)

15. Veri Yönetim Planı hazırlama nedeniniz neydi? (İlgili tüm seçenekleri işaretleyiniz)

- Kişisel tercihim
- Araştırma ekibimin tercihi
- Fon sağlayıcı kurumun zorunlu tutması
- Fon sağlayıcı kurumun tavsiye etmesi
- Çalıştığım kurumun zorunlu tutması
- Çalıştığım kurumun tavsiye etmesi
- Diğer (lütfen açıklayınız):

16. Oluşturduğunuz Veri Yönetim Planının içeriğinde hangi alanlar bulunmaktadır? (İlgili tüm seçenekleri işaretleyiniz)

- Veri toplama (veri türü, formatı, büyüklüğü, dosya yapısı, isimlendirme gibi)
- Dokümantasyon ve üst veri (yöntem, değişkenler, veri hakkında bilgi)
- Etik ve yasal konular (anonimleştirme, veri sahipliği, veri paylaşımı ve lisanslama)
- Depolama ve yedekleme (verinin nerede, kim tarafından yedekleneceği, veriye erişim, veri güvenliği)
- Seçim ve koruma (uzun süreli koruma planlaması)
- Veri paylaşımı (paylaşım koşulları)
- Sorumluluklar ve kaynaklar (sorumlu(lar) kim, yazılım, donanım, teknik uzmanlık gibi kaynak ihtiyaçlarının planlanması vb.)
- Diğer (lütfen açıklayınız) :

17. DMPTool, DMPonline gibi veri yönetimi planlama araçlarını biliyor musunuz?

- Biliyorum ve kullanıyorum
- Biliyorum ama kullanmıyorum
- Bilmiyorum

18. Araştırma verilerinin yönetilmesinde Veri Yönetim Planı'nın araştırmacılara yardımcı olduğunu düşünüyor musunuz?

- Evet
- Hayır
- Fikrim yok

19. Veri yüklemek ve indirmek için kullandığınız veri arşivleri (Zenodo, DRYAD, Figshare gibi) varsa lütfen listeleiniz.

.....

20. Araştırma verilerinin yönetimi, araştırma verisinin toplanması, düzenlenmesi, saklanması, korunması, yayımlanması ve paylaşılması konularını kapsayan bir süreçtir. Araştırma verilerinin yönetimi konusunda eğitim gereksinimi duyuyor musunuz?

- Evet
- Hayır (Lütfen 22. soruya geçiniz.)

21. Aşağıda belirtilen hangi konularda eğitim almak istersiniz? (İlgili tüm seçenekleri işaretleyiniz)

- Veri yönetim planı
- Dokümantasyon (Yöntem, format, üst veri oluşturma, standartlar gibi)
- Organizasyon (Dosya düzenleme, isimlendirme, versiyon belirleme gibi)
- Veri atıf stilleri
- Veri depolama ve koruma
- Veri paylaşımı
- Mevcut veri/veri setleri bulma
- Yasal ve etik konular
- Diğer (lütfen açıklayınız):.....

22. TÜBİTAK ULAKBİM tarafından hazırlanan ve araştırma verilerinin yönetimi konusunda temel bir rehber olan Araştırma Verileri Yönetimi Eğitim Portalı'nı biliyor musunuz?

- Biliyorum ve kullanıyorum
- Biliyorum ama kullanmıyorum
- Bilmiyorum

23. Sizce araştırma verileri yönetimindeki temel problemler nelerdir? (İlgili tüm seçenekleri işaretleyiniz)

- Fikrim yok
- Depolama sorunları
- Verilerin kaybolması
- Güvenlik sorunları
- Koruma ile ilgili sorunlar
- Entelektüel mülkiyet sorunları

- Kalite sorunları
- Verilerin paylaşılmaması
- Diğer (lütfen açıklayınız) :

.....

2. BÖLÜM: VERİ PAYLAŞIMI YAKLAŞIMLARI

24. Araştırma verilerinizi başkalarıyla paylaşır mısınız?

- Verilerimi paylaşıyorum
- Verilerimi kısmen paylaşıyorum (sadece bazı bölümlerini)
- Verilerimi kesinlikle paylaşmam (Lütfen 27. soruya geçiniz)

25. Araştırma verilerinin paylaşımı ile ilgili olarak size uygun olan seçeneği/seçenekleri işaretleyiniz.

- Verilerimi etik ve yasal koşullar elverdiğince paylaşıyorum.
- Verilerim herkese açıktır.
- Verilerimi istek olması halinde paylaşıyorum.
- Verilerimi sadece araştırma grubumdaki kişilerle paylaşıyorum.
- Verilerimi çalıştığım kurumdaki diğer araştırmacılarla paylaşıyorum.
- Verilerimi kendi disiplinimdeki diğer araştırmacılarla paylaşıyorum.
- Verilerimi bilimsel camia ile paylaşıyorum.
- Verilerimi sadece tanıdığım kişilerle/araştırmacılarla paylaşıyorum
- Diğer (lütfen açıklayınız):

26. Verilerinizi başkalarıyla nasıl paylaşıyorsunuz? (İlgili tüm seçenekleri işaretleyiniz) (Lütfen 28. soruya geçiniz)

- E-posta
- Üniversite web sitesi
- Dergi makalesine ek bağlantı
- Kişisel web sayfası
- Veri havuzları
- Diğer (lütfen açıklayınız):

27. Verilerinizi paylaşmama nedenleriniz nelerdir? (İlgili tüm seçenekleri işaretleyiniz)

- Verilerimin açıklama olmadan anlaşılacak durumda olmaması
- Verilerimin temizlenmemiş olması
- Fon yetersizliği
- Üst veri standardı yetersizliği
- Başkalarının benim verilerime ihtiyacı olmadığını düşünüyorum
- Zaman yetersizliği
- Verilerimi sunacak yer olmaması
- Bu konuda zorlayıcı olmaması (fon veren kuruluş, çalıştığım kuruluş gibi)
- Verilerimi kamuya açma hakkımın olmaması

- Verilerimi kamuya açacak teknik bilgi yetersizliği
- Bilimsel gücü kaybetme korkusu
- Yasal ve etik sorunlar
- Verinin hatalı kullanılabilceği endişesi
- Verinin yanlış yorumlanabileceği endişesi
- Diğer (lütfen açıklayınız):

28. 14.03.2019 tarihinde yürürlüğe giren TÜBİTAK Açık Bilim Politikası'nın farkında mısınız?

- Farkındayım ve okudum
- Farkındayım ama okumadım
- Farkında değilim

29. Araştırma verileri yönetimi ile ilgili katkı sağlamak istediğiniz hususlar varsa lütfen belirtiniz.

.....

3. BÖLÜM: DEMOGRAFİK BİLGİLER

30. Yaşınız?

.....

31. Bölümünüz?

.....

32. Kaç yıldır araştırma faaliyetinde buluyorsunuz? (İlk yayın (tez, proje çıktısı, makale gibi) üretme ve/veya ürün geliştirme sürecinden itibaren)

.....

33. Unvanınız?

- Profesör
- Doçent
- Dr. Öğretim Üyesi
- Dr. Öğretim Görevlisi
- Öğretim Görevlisi
- Dr. Araştırma Görevlisi
- Araştırma Görevlisi
- Diğer (lütfen açıklayınız):

34. Kurumunuz?

.....

DEĞERLİ KATKILARINIZ İÇİN TEŞEKKÜR EDERİZ.

EK 2. Etik Kurul İzni



T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük



Sayı : 35853172-300
Konu : Canan TAVLUOĞLU (Etik Komisyon İzni)

SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 14.10.2019 tarihli ve 12908312-300/00000809140 sayılı yazınız.

Enstitünüz Bilgi ve Belge Yönetimi Anabilim Dalı Doktora programı öğrencilerinden **Canan TAVLUOĞLU**'nun Prof. Dr. S. Serap KURBANOĞLU danışmanlığında hazırladığı "Araştırma Verilerinin Yönetimi: Hacettepe Üniversitesi Örneği" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun 05 Kasım 2019 tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-İmzalıdır
Prof. Dr. Rahime Meral NOHUTCU
Rektör Yardımcısı

Evrakın elektronik imzalı suretine <https://belgedogrulama.hacettepe.edu.tr> adresinden 64834D98-3027-46b69290-bd5847150c8f3 beldi ile erişebilirsiniz.
Bu belge 5070 sayılı Elektronik İmza Kanunu'na uygun olarak Ölümlü Elektronik İmza ile imzalanmıştır.

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon:0 (312) 305 3001-3002 Faks:0 (312) 311 9992 E-posta:yazim@hacettepe.edu.tr İnternet
Adres: www.hacettepe.edu.tr

Serda TOPRAK





HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
BİLGİ VE BELGE YÖNETİMİ ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 01/07/2022

Tez Başlığı : *Araştırma Verilerinin Yönetimi: TÜBİTAK Projelerinde Yer Alan Araştırmacılar Üzerine Bir Değerlendirme*

Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 133 sayfalık kısmına ilişkin, 01/07/2022 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda işaretlenmiş filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 7'dir.

Uygulanan filtrelemeler:

- Kabul/Onay ve Bildirim sayfaları hariç
- Kaynakça hariç
- Alıntılar hariç
- Alıntılar dâhil
- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Tarih ve İmza

Adı Soyadı: Canan Tavluoğlu

Öğrenci No: N13240504

Anabilim Dalı: Bilgi ve Belge Yönetimi

Programı: Doktora

Statüsü: Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. S. Serap Kurbanoglu