



Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Bilgi ve Belge Yönetimi Anabilim Dalı

**SAĞLIK BİLİMLERİNDE ARAŞTIRMA VERİLERİNİN YÖNETİMİ:  
HACETTEPE ÜNİVERSİTESİ LİSANSÜSTÜ ÖĞRENCİLERİNİN  
FARKINDALIKLARININ DEĞERLENDİRİLMESİ**

Eda ÖZEL

Yüksek Lisans Tezi

Ankara, 2023



SAĞLIK BİLİMLERİNDE ARAŞTIRMA VERİLERİNİN YÖNETİMİ: HACETTEPE  
ÜNİVERSİTESİ LİSANSÜSTÜ ÖĞRENCİLERİNİN FARKINDALIKLARININ  
DEĞERLENDİRİLMESİ

Eda ÖZEL

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Bilgi ve Belge Yönetimi Anabilim Dalı

Yüksek Lisans Tezi

Ankara, 2023

## KABUL VE ONAY

Eda Özel tarafından hazırlanan ‘‘Saęlık Bilimlerinde Arařtırma Verilerinin Yönetimi: Hacettepe Üniversitesi Lisansüstü Öğrencilerinin Farkındalıklarının Deęerlendirilmesi’’ başlıklı bu çalıřma, 19/01/2023 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiřtir.

---

Prof. Dr. Melih ELÇİN (Bařkan)

---

Prof. Dr. Fahrettin ÖZDEMİRCİ (Üye)

---

Prof. Dr. Özgür KÜLCÜ (Danıřman)

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Uęur ÖMÜRGÖNÜLŐEN  
Enstitü Müdürü

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

...../...../.....

**Eda ÖZEL**

<sup>1</sup>“*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*”

- (1) *Madde 6. 1. Lisansüstü tezle ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının 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 internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının 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, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. \* Kurum ve kuruluşlarla yapılan işbirliği 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 danışmanının 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. zgr KLC** danıřmanlıđında tarafımdan retildiđini ve Hacettepe niversitesi Sosyal Bilimler Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.

**Eda ZEL**

## TEŞEKKÜR

Hem mesleki olarak hem de lisansüstü eğitimim süresince bu zor yolda her umutsuzluğa kapıldığımda desteğiyle toparlanıp devam etmemi sağlayan, pes etmeme asla izin vermeyen, tez çalışmamın her aşamasında zamanını, desteğini ve ilgisini esirgemeyen sayın hocam Prof. Dr. Özgür KÜLCÜ'ye bana ve tezime sunduğu katkı için teşekkür ederim.

Tez çalışmamla yakından ilgilenip her soruma sabırla cevap veren sevgili Dr. Öğr. Üyesi Betül AKBUĞA ÖZEL'e ve jürime katılmayı kabul ederek tez çalışmama sundukları katkılar için değerli hocalarım Prof. Dr. Melih ELÇİN ve Prof. Dr. Fahrettin ÖZDEMİRCI'ye teşekkür ederim.

Hem lisans hem yüksek lisans eğitimimde bilgi birikiminden yararlandığım tüm Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü hocalarıma teşekkürü borç bilirim. Ayrıca tez sürecim boyunca yardımlarını esirgemeyen sevgili yöneticim H.Ü. Öğrenci İşleri Daire Başkan Vekili Özkan AY'a, anlayışı ve desteği için mesai arkadaşım Uğur ÖZTÜRK'e teşekkür ederim. Çalışma kapsamında anketimi uygulamama olanak sağlayan H.Ü. Sağlık Bilimleri Enstitüsü ile ankete katılarak araştırmamı tamamlamamı sağlayan lisansüstü öğrencilerine kalpten teşekkür ederim.

Son olarak, başta can dostum, çocukluk arkadaşım Deniz ÇAVUŞ olmak üzere, özellikle son üç yıldır fark etmenin ve öğrenmenin nefes aldığımız sürece devam ettiğini deneyimlediğim kendimi bulma yolculuğumda, benim bugünkü ben olmama katkı sağlayan herkese en derin teşekkürlerimi sunarım.

*Varlığıyla hayatıma ışık saçan, birlikte yeniden doğup büyüdüğüm,  
Canım kızım, Defnem'e...*



## ÖZET

ÖZEL, Eda. *Sağlık Bilimlerinde Araştırma Verilerinin Yönetimi: Hacettepe Üniversitesi Lisansüstü Öğrencilerinin Farkındalıklarının Değerlendirilmesi*, Yüksek Lisans Tezi, Ankara, 2023.

Araştırmalarda bütçeden en büyük payı veri toplama çalışmalarının alması nedeniyle fon sağlayan kuruluşların verinin yeniden kullanımı beklentisi artmakla birlikte, araştırma verilerinin yönetimi uygulamaları giderek önem kazanmaktadır. Sağlık araştırmalarında elde edilen verilerin paylaşımı ve yeniden kullanımı uluslararası düzeyde ekonomik ve bilimsel olarak önemli bir güç haline gelmiştir. Ancak özellikle sağlık verilerinin yaşam döngüleri boyunca iyi yönetilmesi, veri kaynaklarının sürdürülebilir olması ve ihlallere karşı koruma sağlanması için başarılı veri yönetim uygulamaları oldukça önemlidir. Bu amaçla araştırma verilerinin yönetimi ve verinin yeniden kullanımı ile ilgili kapsamlı politikalar belirlemek kadar önemli olan bir diğer konu araştırmacılara veri okuryazarlığı becerilerini geliştirebilmeleri için eğitim imkanlarının sunulabilmesidir.

Bu çalışmada Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü'ne kayıtlı öğrencilerin veri yönetim alışkanlıkları, eğitim ihtiyaçları ve veri paylaşımı hakkındaki düşünceleri ile ilgili mevcut durum ortaya koyulmuştur. Ayrıca ders programlarında veri okuryazarlığı becerilerini geliştirebilecek içerikler olup olmadığı incelenmiştir. Paralel karma desende tasarlanan araştırmanın nicel kısmında katılımcılara 31 soruluk bir anket uygulanmıştır. Ankete 140 kişi katılmış olup ankette elde edilen bulgulara göre eğitim ihtiyaçlarının özellikle verinin yeniden kullanımı ve telif hakları ile veriye atf konularında yüksek olduğunu belirtmişlerdir. Verilerini paylaşma konusunda istekli olmalarının yanı sıra paylaşımın önündeki en büyük engelin etik ihlal olabileceği düşüncesi olduğunu belirtmişlerdir. Araştırmanın nitel kısmında ise enstitüye bağlı programların ders programları incelenerek bilimsel bir araştırmayı planlama, yürütme ve raporlama becerilerini geliştirmeyi hedefleyen derslerin içerikleri tematik analiz yöntemiyle incelenmiştir. İçerikler veri yönetimi uygulamaları bakımından incelendiğinde yeterli olmadığı görülerek eğitim programlarına veri okuryazarlığı içerikleri eklenmesi önerisinde bulunulmuştur.

### **Anahtar Sözcükler**

Sağlık verisi, araştırma verisi, araştırma verilerinin yönetimi, veri yönetim planı, veri okuryazarlığı

## ABSTRACT

ÖZEL, Eda. *Research Data Management in Health Sciences: Assessment of Hacettepe University Graduate Students' Awareness*. Master Thesis. Ankara, 2023.

Since data collection activities take the largest share of the budget in research, the expectations of funding organizations for data reuse are increasing, and research data management practices are becoming increasingly important. Sharing and reuse of data gathered in health research has become an important economic and scientific force at the international level. However, successful data management practices are especially important to manage health data well throughout their lifecycle, to ensure the sustainability of data resources and to protect against data breaches. For this purpose, as important as setting comprehensive policies on research data management and data reuse is the provision of training opportunities for researchers to develop data literacy skills.

In this study, the current situation of the data management practices, training needs and views about data sharing of students in Hacettepe University Institute of Health Sciences was revealed. In addition, it was examined whether there are contents that can improve data literacy skills in the curriculum. In the quantitative part of the parallel mixed designed research, a questionnaire consisting of 31 questions was conducted to the participants. 140 respondents participated in this study, and according to the findings obtained from the survey, their training needed improvement especially in data reuse, copyright and data citation. Although they are willing to share their data, they think that the biggest barrier to sharing may be ethical violations. In the qualitative part of the study, the curricula of the programs affiliated to the institute were examined and the contents of the courses aiming to develop the skills of planning, conducting and reporting a scientific research were examined by thematic analysis method. When the contents were examined in terms of data management practices, most were found to be not sufficient, and it was suggested that data literacy should be added to the curricula.

### Keywords

Health data, research data, research data management, data management plan, data literacy

## İÇİNDEKİLER

<b>KABUL VE ONAY</b> .....	i
<b>YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI</b> .....	ii
<b>ETİK BEYAN</b> .....	iii
<b>TEŞEKKÜR</b> .....	iv
<b>ÖZET</b> .....	vi
<b>ABSTRACT</b> .....	vii
<b>İÇİNDEKİLER</b> .....	viii
<b>KISALTMALAR DİZİNİ</b> .....	x
<b>TABLolar DİZİNİ</b> .....	xi
<b>ŞEKİLLER DİZİNİ</b> .....	xii
<b>1. BÖLÜM: GİRİŞ</b> .....	1
<b>1.1. KONUNUN ÖNEMİ</b> .....	1
<b>1.2. ARAŞTIRMANIN AMACI, PROBLEMİ VE HİPOTEZİ</b> .....	3
<b>1.3. ARAŞTIRMANIN KAPSAMI</b> .....	5
<b>1.4. ARAŞTIRMANIN YÖNTEMİ</b> .....	6
<b>1.5. ARAŞTIRMANIN DÜZENİ</b> .....	9
<b>1.6. KAYNAKLAR</b> .....	9
<b>2. BÖLÜM: ARAŞTIRMA VERİLERİNİN YÖNETİMİ</b> .....	11
<b>2.1. ARAŞTIRMA VERİSİ ve PAYLAŞIMI</b> .....	11
<b>2.2. VERİ YAŞAM DÖNGÜSÜ</b> .....	16
<b>2.3. VERİ YÖNETİM PLANI</b> .....	19
<b>2.4. ARAŞTIRMA VERİLERİNİN YÖNETİMİ UYGULAMALARI..</b> .....	21
2.4.1 Planlama .....	22
2.4.2. Veri Toplama ve Düzenleme .....	23
2.4.3. Depolama ve Koruma .....	26
2.4.4. Yayınlama ve Paylaşım .....	27
2.4.5. Yeniden Kullanım .....	28

<b>3. BÖLÜM: SAĞLIK BİLİMLERİNDE VERİ, YENİDEN KULLANIMI VE VERİ OKURYAZARLIĞI .....</b>	<b>31</b>
<b>3.1. SAĞLIK VERİSİ .....</b>	<b>31</b>
<b>3.2. SAĞLIK VERİLERİNİN YENİDEN KULLANIMI .....</b>	<b>36</b>
3.2.1. Mahremiyet .....	41
3.2.2. Anonimleştirme .....	43
<b>3.3. SAĞLIKTA VERİ OKURYAZARLIĞI .....</b>	<b>45</b>
3.3.1. Tıp Eğitimi ve Veri Okuryazarlığı .....	45
3.3.2. Veri Okuryazarlığı Eğitimi .....	46
<b>4. BÖLÜM: BULGULAR .....</b>	<b>50</b>
<b>4.1. ARAŞTIRMA VERİLERİNİN YÖNETİMİ FARKINDALIK     ANKETİ .....</b>	<b>50</b>
4.1.1. Bilgi Arama Davranışları .....	53
4.1.2. Araştırma Verilerinin Yönetimi .....	62
4.1.3. Araştırma Verilerinin Paylaşımı .....	69
4.1.4. Katılımcıların Eğitim İhtiyacı .....	74
4.1.5. Katılımcıların Görüşleri .....	76
<b>4.2. DERS İÇERİKLERİ ANALİZİ .....</b>	<b>77</b>
<b>5. BÖLÜM: SONUÇ VE ÖNERİLER .....</b>	<b>83</b>
<b>5.1. SONUÇ .....</b>	<b>83</b>
<b>5.2. ÖNERİLER .....</b>	<b>89</b>
<b>KAYNAKÇA .....</b>	<b>91</b>
<b>EK 1. ORJİNALLİK RAPORU.....</b>	<b>104</b>
<b>EK 2. ETİK KURUL İZİNİ .....</b>	<b>106</b>

## KISALTMALAR DİZİNİ

ARBİS	Arařtırmacı Bilgi Sistemi
AVY	Arařtırma Verilerinin Yönetimi
BAP	Bilimsel Arařtırma Projeleri
BMJ	British Medical Journal
CC	Creative Commons
CNRS	Centre National de la Recherche Scientifique
EMBL	European Molecular Biology Laboratory
FAIR	Findability, Accessibility, Interoperability, Reusability
FTR	Fizik Tedavi ve Rehabilitasyon
GDPR	General Data Protection Regulation
HHMI	Howard Hughes Medical Institute
IGWWD	Interagency Working Group On Digital Data
MAXQDA	Max Qualitative Data Analysis
MMV	The Medicines for Malaria Venture
NCATS	The National Center for Advancing Translational Sciences
NIH	National Institute of Health
NISO	Ulusal Bilgi Standartları Organizasyonu
NSF	National Science Foundation
OECD	Organisation for Economic Co-operation and Development
OSS	Open Source
PMC	PubMed Central
QID	Quasi Identifier
SPSS	Statistical Package for the Social Sciences
TSH	Thyroid Stimulating Hormone
TÜBİTAK	Türkiye Bilimsel ve Teknolojik Arařtırma Kurumu
URAP	University Ranking by Academic Performance
WEF	World Economic Forum

## TABLOLAR DİZİNİ

<b>Tablo 1.</b> Ankete katılan öğrencilerin fakültelere göre dağılımı .....	7
<b>Tablo 2.</b> Programlarda verilen ve araştırmaya konu olan dersler ile ilgili genel bilgiler.	8
<b>Tablo 3.</b> Üst Veri Türleri.....	25
<b>Tablo 4.</b> Lisanslar ve Koşulları.....	28
<b>Tablo 5.</b> Fakülte ve bölümlere göre öğrenci dağılımı.....	51
<b>Tablo 6.</b> Eğitim düzeyleri ve dönemlerine göre öğrenci dağılımı.....	52
<b>Tablo 7.</b> Öğrenci sayısının yaşa göre dağılımı.....	52
<b>Tablo 8.</b> Katılımcılar arasında çalışanların unvanları.....	53
<b>Tablo 9.</b> Bilgi arama davranışları.....	54
<b>Tablo 10.</b> Bilgiyi internette ararken tercih edilen yaklaşımlar.....	56
<b>Tablo 11.</b> Katılımcıların bilgiyi başka kaynaktan teyit alışkanlıkları.....	58
<b>Tablo 12.</b> Veri elde etme biçimleri.....	59
<b>Tablo 13.</b> Fakültelere göre bilgi arama davranışlarını öğrenme şekli.....	61
<b>Tablo 14.</b> Fakültelere göre kullanılan veri türleri.....	62
<b>Tablo 15.</b> Fakültelere göre kullanılan veri formatları.....	63
<b>Tablo 16.</b> Fakültelere göre verilerini saklama süreleri.....	64
<b>Tablo 17.</b> Fakültelere göre üst veri oluşturma alışkanlıkları.....	65
<b>Tablo 18.</b> Üst veri oluşturmada kullanılan yöntemler.....	65
<b>Tablo 19.</b> Araştırma devam ederken veriyi saklama ortamları.....	66
<b>Tablo 20.</b> Araştırma tamamlandıktan sonra veriyi saklama ortamları.....	67
<b>Tablo 21.</b> Araştırma devam ederken veriyi yedekleme ortamları.....	67
<b>Tablo 22.</b> Araştırma tamamlandıktan sonra veriyi yedekleme ortamları.....	68
<b>Tablo 23.</b> Araştırma verilerinin paylaşımı.....	70
<b>Tablo 24.</b> Görüşlerdeki ifadelerin sıklığı.....	70
<b>Tablo 25.</b> Veri Paylaşımı önündeki engeller.....	71
<b>Tablo 26.</b> Veri Paylaşımı hakkındaki görüşler.....	73
<b>Tablo 27.</b> Verilerin yeniden kullanımını ile ilgili görüşler.....	74
<b>Tablo 28.</b> Katılımcıların eğitim ihtiyaçları.....	75
<b>Tablo 29.</b> Katılımcıların eğitim ihtiyaçları 2 .....	75
<b>Tablo 30.</b> Ortak verilen dersler.....	78
<b>Tablo 31.</b> Nitel veri seti ile ilgili genel bilgiler .....	78
<b>Tablo 32.</b> Kod kitabı .....	79

## ŞEKİLLER DİZİNİ

<b>Şekil 1.</b> Araştırmanın akış diyagramı.....	6
<b>Şekil 2.</b> Biyomedikal veri yaşam döngüsü .....	17
<b>Şekil 3.</b> Sağlık hizmetlerinde veri ekosistemi.....	33
<b>Şekil 4.</b> Tüm dersler Tek vaka modeli kod haritası.....	81
<b>Şekil 5.</b> SBE601 kodlu dersin kod haritası.....	82

# 1. BÖLÜM

## GİRİŞ

### 1.1. KONUNUN ÖNEMİ

Araştırma verileri, bilimsel arařtırmaların ana varlığı, arařtırmanın temeli ve aynı zamanda arařtırmanın ıktısıdır. Verilerin yeniden kullanımı, bu alana yapılan yatırımların getirisini artırır, açık bilimsel sorgulamayı güçlendirir, alıřmaların ve görüşlerin çeřitliliğini arttırır, yeni arařtırma alanlarını destekler ve ilk arařtırmacılar tarafından öngörölmeıen konuların keřfedilmesini saęlar. Arařtırma verilerinin paylařımı, arařtırmada řeffaflığı ve tekrarlanabilirlięi teřvik etme, eęitim ve öęretime katkıda bulunma ve yeniliki arařtırmaları teřvik etme potansiyeline sahiptir. Arařtırmanın tekrarlanabilir olması, bulguların doęruluęu ve sonucunun güvenilirlięini ortaya koymada büyük öneme sahiptir. Saęlık verileri bir bireyin saęlık durumuyla ilgili tıbbi muayene raporları, laboratuvar testleri, tıbbi görüntölleme sonuçları vb kiřisel bilgileridir ve bu verilerin ikincil kullanımı, bireylerin saęlık hizmetleri deneyimlerini geliřtirir, hastalıklar ve tedavileri hakkında yeni bilgilerin keřfedilmesini, saęlık sistemlerinin etkinlięi ve verimlilięinin anlařılmasını saęlayarak, halk saęlıęı politikalarını destekler. Günümüzde küresel veri hacminin yaklaşık %30'u saęlık sektörü tarafından üretilmektedir ancak bu verilerin tamamı kullanılmamaktadır. Örneęin, hastaneler tarafından üretilen tüm verilerin %97'si kullanılmamaktadır (World Economic Forum, 2019).

Hey, Tansley ve Tolle'ye (2009, s. 17) göre yayınlanmış arařtırmalarda analiz sonuçları paylařılan veriler buzdaęının görünen kısmı olup, toplanan ancak sistematik bir řekilde düzenlenmeyen veya yayınlanmayan ok daha büyük miktarda veri bulunmaktadır. Özellikle tıp alanında yapılan arařtırmaların ok maliyetli ve insan saęlıęını doęrudan etkileyen arařtırmalar olması, bu arařtırmalar sonucunda yeni ilaların ve tedavi yöntemlerinin geliřtiriliyor olması arařtırmaların güvenilirlięini daha önemli hale getirmektedir. Prinz ve arkadaşları (2011) oęunluęu kanser alanında olan ve 23 arařtırmacı tarafından yayınlanan 67 arařtırmanın verilerini yeniden deęerlendirmiş ve



sadece %21'inin doğrulanabilir olduğunu ortaya koymuşlardır. Projelerin neredeyse üçte ikisinin yeniden yatırım yapılacak araştırmalar olmadığını belirtmişlerdir. Benzer şekilde Begley ve Ellis (2012) kanser alanında yapılmış 53 araştırma üzerinde yaptıkları çalışmada bu oranı %11 olarak ortaya koymuşlardır. Psikoloji alanında yapılan bir çalışmada ise araştırmacılar 3 psikoloji dergisinde yayınlanan 100 makale üzerinde yaptıkları çalışmada sadece %36'sının verilerinin doğrulanabilir olduğunu ortaya koymuşlardır (Open Science Collaboration, 2015). Sağlık bilimleri alanında, veri paylaşımının faydaları sadece bilimin gelişmesi için değil, aynı zamanda insan sağlığı ve yaşamı için de özellikle önemlidir. Daha hızlı ve engellerden arındırılmış veri akışı, sonuçların doğrulanması ve olası ihlallerin ortaya çıkarılması bilimin hiçbir alanında sağlık bilimlerinde olduğu kadar önemli değildir (Milewska ve diğerleri, 2021).

Özellikle sağlık araştırmacıları, büyük miktarlarda hassas veri ile çalışmalarından dolayı araştırma verilerinin yönetiminde önemli bir yere sahiptir (Akers ve Doty, 2013). Ancak sağlık araştırmacıları nadiren veri yönetimi eğitimi almakta dolayısıyla veri yönetimi konusunda eğitim eksiklikleri bulunmaktadır (Federer ve diğerleri, 2016). Araştırma verilerinin yönetimi ile ilgili yapılan ve farklı disiplinlerden çeşitli araştırmacı gruplarının farkındalıklarını değerlendiren benzer çalışmalar görmek mümkündür (Cox ve diğerleri, 2019; Kim, 2015; Scaramozzino, Ramirez ve McGaughey, 2012; Tenopir ve diğerleri, 2011). Ülkemizdeki çalışmalara baktığımızda ise, ilk olarak Aydınoglu, Doğan ve Taşkın'ın (2017), 25 farklı üniversiteden 532 akademisyen üzerinde gerçekleştirdiği araştırma görülmektedir. Ünal ve Kurbanoglu (2018), Türkiye'deki 44 üniversiteden toplam 540 akademisyen üzerinde farkındalıkları üzerine araştırma gerçekleştirmiştir. Zencir (2019) doktora tezinde Ankara Üniversitesi'nde Bilimsel Araştırma Projeleri Birimi (BAP) desteğiyle proje yürüten farklı disiplinlerden 194 araştırmacı üzerinde bir araştırma gerçekleştirmiştir. Doğan, Taşkın ve Aydınoglu (2021), ARBİS'e kayıtlı farklı disiplinlerden 1577 araştırmacının farkındalıklarını değerlendirmiştir. Tavluoglu (2022) ise doktora tezinde URAP sıralamasında en başarılı üniversiteler arasında ilk beşte olup 2014 ile 2018 arasında TÜBİTAK desteğiyle araştırma yürüten 408 akademisyeni değerlendirmiştir. Sağlık bilimleri alanındaki araştırmacıların araştırma verilerinin yönetimi konusunda farkındalıklarının ve yeterliliklerinin değerlendirildiği çalışmalar uluslararası literatürde mevcuttur (Akers ve Doty, 2013; Bardyn, Resnick ve Camina,

2012; Borghi ve Gulick, 2018; Federer, Lou ve Joubert, 2016; Tenopir ve diğeri, 2011), ancak ülkemizde belli bir disiplin özelinde araştırma bulunmamaktadır.

Ülkemizde yapılan arařtırmalar akademisyenlerle sınırlı olarak ve farkındalıklarını deęerlendirmek üzere yürütülmüřtür. Uluslararası literatürde ise lisans ve lisansüstü öğrenciler üzerinde yapılan (Adamick, Reznik-Zellen ve Sheridan, 2012; Carlson ve diğeri, 2011; Pasek ve Mayer, 2019; Piorun ve diğeri, 2012) ve belli bir disipline özgü arařtırmacılar üzerinde gerçekleştirilen çalışmalar da mevcuttur (Chen ve Wu, 2017; Hyndman ve Goody, 2020; Strasser ve Hampton, 2012). Her ne kadar politikalar arařtırma verilerinin paylaşımı ve veri yönetim planı hazırlamada arařtırmacılara farkındalık sağlıyor olsa da geçmiş toplantılarda alınan kararlara, belirlenen hedeflere ve bugünkü uygulamalara baktığımızda planlanandan yavaş ilerledięi görülmektedir. Ülkemizde yapılan arařtırmalar, arařtırma verilerinin yönetimi konusunda politika eksiklięi olduęunu ve arařtırmacıların bilgi ve farkındalık düzeylerinin düşük olduęunu göstermektedir.

Bu çalışmayı daha önce yapılan çalışmalardan ayıran önemli farklardan biri sağlık bilimleri alanındaki arařtırmacılara odaklanmış olmasıdır. Bu kapsamda yine ülkemizde yapılan dięer çalışmalardan farklı olarak akademisyenler üzerinde deęil, arařtırma becerilerinin geliřtięi lisansüstü eğitime devam eden öğrenciler üzerinde gerçekleştirilmiştir.

## **1.2. ARAŐTIRMANIN AMACI, PROBLEM VE HİPOTEZ**

Arařtırmacıların çoęu arařtırma verilerinin yönetimi konusunda eğitimi tecrübeli arařtırmacılarından genç arařtırmacılara aktarılması yoluyla almaktadırlar. Literatürdeki çalışmaların ortaya koyduęu sorunlar benzer olmakla birlikte eğitim eksiklięi tüm arařtırmalarda öne çıkmakta ve disiplinler arasında fark göstermemektedir.

Veri yönetimi sorumluluęu öncelikle arařtırmacılara aittir, ancak kurum ve kuruluşlar eğitim konusunda destekleyici bir çerçeve sağlayabilir. Arařtırma verilerinin yönetimi ve veri yönetim planı eğitimi, kütüphaneciler ve bilgi bilimciler ile eş zamanlı olarak, son sınıf lisans öğrencileri veya birinci sınıf lisansüstü öğrenciler, yani arařtırma verileriyle ilgilenme olasılıęı daha yüksek olan öğrenciler için müfredat içinde genel bir kurs olarak

sağlanabilir. Bu eğitimler farklı disiplinlerin farklı veri alışkanlıkları olduğundan, disipline özgü olabilirler (Doğan, Taşkın ve Aydınoglu, 2020, s. 10). Araştırmacılara araştırma verilerinin paylaşımı ve veri okuryazarlığı üzerine eğitim verilmesi yalnızca araştırma verilerinin paylaşılmasını desteklemeye değil, araştırma kültürünün gelişmesine ve veri odaklı ilerlemesine yarar sağlayacaktır. Bu anlamda eğitim programları planlanırken mevcut durumun belirlenmesi, farklı disiplinlere yönelik farklı içeriklerin hazırlanması noktasında araştırmacıların ihtiyaçlarının belirlenmesi kritik önem taşımaktadır.

Bu araştırmacının amacı Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü öğrencilerinin araştırma verilerinin paylaşımı hakkındaki görüşleri ile veri yönetim alışkanlıkları ve eğitim içeriklerinin bu konuları destekleyip desteklemediği noktasında mevcut durumu ortaya koymaktır.

Araştırmanın problemi “Sağlık bilimleri lisansüstü öğrencilerinin araştırma verilerinin yönetimi konusunda bilgi eksikliği vardır” şeklinde oluşturulmuştur. Araştırmanın nicel kısmında yanıt aranan temel araştırma soruları ve alt soruları şunlardır:

1. Sağlık bilimleri lisansüstü öğrencilerinin bilgi arama ve veri yönetim alışkanlıkları nelerdir? Öğrencilerin;

- Bilgiye erişimde hangi yaklaşımları tercih etmektedirler?
- Veri elde etme biçimleri nelerdir?
- Veri saklama ve yedekleme alışkanlıkları nelerdir?

2. Sağlık bilimleri lisansüstü öğrencilerinin veri paylaşımı ile ilgili farkındalıkları ne düzeydedir? Öğrencilerin;

- Araştırma verilerinin paylaşımı ile ilgili düşünceleri nelerdir?
- Daha önce veri yönetim planı hazırlamışlar mıdır?
- Daha önce başka araştırmacıların verilerini kullanmışlar mıdır?
- Daha önce verilerini başka araştırmacılarla paylaşmışlar mıdır?

3. Sağlık bilimleri lisansüstü öğrencilerinin araştırma verilerinin yönetimi ve veri yönetim uygulamaları hakkında eğitime ihtiyacı var mıdır? Öğrencilerin;

- Eğitime en çok ihtiyaç duydukları konular nelerdir?
- Eğitim ihtiyaçları fakültelere göre değişmekte midir?

Araştırmanın nitel kısmında ise “Eğitim programlarında, araştırma verilerinin yönetimi ve veri yönetim planı uygulamaları konusunda bilgilendirmeye yönelik içerik mevcut mudur?” sorusuna cevap aranmıştır. Araştırmanın hipotezleri şu şekilde belirlenmiştir;

- Bilgi arama yaklaşımları fakültelere göre farklılık göstermektedir.
- Veri elde etme biçimleri fakültelere göre farklılık göstermektedir.
- Veri saklama ve yedekleme alışkanlıkları fakültelere göre farklılık göstermektedir.
- Paylaşılan veri kullanma istekliliği fakültelere göre farklılık göstermektedir.
- Verilerini paylaşma istekliliği fakültelere göre farklılık göstermektedir.
- Veri yönetim planı hazırlama ile fon destekli araştırma yürütüp yürütmeme durumu arasında ilişki vardır.
- Eğitim ihtiyaçları fakültelere göre farklılık göstermektedir.
- Eğitim programlarında araştırma verilerinin yönetimi ve veri yönetim planı uygulamaları ile ilgili yeterli içerik yoktur.

### **1.3. ARAŞTIRMANIN KAPSAMI**

Bu araştırma, araştırma verilerinin yönetimi uygulamaları ile sağlık veri okuryazarlığı eğitim programlarını ve sağlık bilimleri araştırmacılarının bu konulardaki uygulamalarını ve görüşlerini kapsamaktadır. Bu kapsamda araştırma Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü’ne bağlı programlarda 2022-2023 Güz Dönemi itibariyle kayıtlı bulunan öğrencilerin araştırma verilerinin paylaşımına bakışı, veri yönetim alışkanlıkları ile eğitim ihtiyaçlarının ortaya koyulmasının hedeflendiği “Araştırma Verilerinin Paylaşımı Farkındalık Anketi”ne verdikleri cevaplar ile ve bu programların “[akts.hacettepe.edu.tr](https://akts.hacettepe.edu.tr)” adresinde belirtilen ders içerikleri ile sınırlandırılmıştır. Hacettepe Üniversitesi URAP verilerine göre 2021-2022 eğitim öğretim yılında Türkiye’de hem tüm üniversitelerin genel sıralamasında hem de üniversite niteliklerine göre sıralamada “Tıp Fakültesi Olan Üniversiteler” arasında birinci sırada olmasının yanı sıra, sağlık bilimleri alanında önde gelen bir araştırma üniversitesidir (URAP, 2022).

## 1.4. ARAŞTIRMANIN YÖNTEMİ

Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü öğrencilerinin araştırma verilerinin yönetimi konusunda farkındalık düzeyini ve veri yönetim alışkanlıklarını ortaya koymak amacıyla betimleme yöntemi kullanılmıştır. Betimleme yöntemi incelenen olayın, durumun, koşulların veya bunların özelliklerinin ne olduğunu açıklayan, ortaya koyan incelemelerdir (Kaptan, 2000, s. 59).

Araştırma paralel karma desende tasarlanmıştır (bkz. Şekil 1). Bu desende nicel ve nitel yöntemler kullanılarak iki farklı veri türü eş zamanlı olarak toplanıp, bağımsız olarak analiz edilir. Araştırılan sorunun tam resmini ortaya koymak için iki analiz sonucunda elde edilen bulgular karşılaştırılır (Creswell and Plano Clark, 2017). Eş zamanlı paralel karma yöntemin amacı aynı konuda farklı ancak tamamlayıcı veriler elde etmektir (Morse, 1991, s. 122).

**Şekil 1.** Araştırmanın akış diyagramı



Araştırmanın nicel kısmında öğrencilerin farkındalık düzeylerini ortaya koymak amacıyla daha önce yapılmış çalışmalar (Doğan, Taşkın ve Aydınoglu, 2021; Tavluoglu, 2022; Zencir, 2019) referans alınarak “Araştırma Verilerinin Paylaşımı Farkındalık Anketi”

oluşturulmuştur. Pilot uygulama ve uzman görüşü alınması sonrasında son şekli verilen ankette, öğrencilerin demografik bilgileri, ana bilim dalı, bilgi arama davranışları, veri saklama, paylaşma ve kullanma alışkanlıkları ile araştırma verilerinin paylaşımı ve yeniden kullanımı ile ilgili düşüncelerini ortaya koymayı amaçlayan 31 soru bulunmaktadır. Anket aracılığıyla toplanan verilerin SPSS 22 paket programı ile analizi sonucu ortalama ve standart sapma hesaplanarak betimleyici tablolar ortaya koyulmuş, Ki-kare testleri yapılmıştır.

Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü'ne bağlı 6 fakültede bulunan 81 ana bilim dalı ve bu ana bilim dallarına bağlı 177 program bulunmaktadır. Anket formu bu programlara kayıtlı toplam 1856 öğrenciye (Google Forms aracılığıyla) 2022-2023 Eğitim-Öğretim Yılı Güz dönemi program kayıt haftasından hemen sonra 4 Ekim 2022 günü e-posta yoluyla gönderilmiştir. 18 Ekim 2022 tarihi itibarıyla 140 kişiden geri dönüş alınmış ve bu tarihten sonra anket formu dönmemiş olup teknik kısıtlılıklar nedeniyle hatırlatma e-postası gönderilememiştir. Ankete cevap verme motivasyonlarının düşük oluşu, geçersiz/kullanılmayan e-posta adresleri ve spam kutusuna düşen e-postaların görülmemiş olması bu araştırmanın sınırlılıkları dahilinde değerlendirilebilir. Anket sonuçları evreni temsil etmese de paralel karma desen kapsamında nicel verilerin toplandığı araştırma grubunu temsil etmektedir. Ankete katılan öğrencilerin eğitim gördükleri lisansüstü programlar fakülterlere göre gruplanmış ve dağılımı Tablo 1'de gösterilmiştir.

**Tablo 1.** Ankete katılan öğrencilerin fakülterlere göre dağılımı

Fakülteler	S	%
Eczacılık Fakültesi	15	10,7
Fizik Tedavi ve Rehabilitasyon	33	23,5
Hemşirelik Fakültesi	12	8,6
Sağlık Bilimleri Fakültesi	33	23,5
Spor Bilimleri Fakültesi	8	5,7
Tıp Fakültesi	39	28
Toplam	140	100

Araştırmanın nitel kısmında ise eğitim eksikliğinin olası nedenlerini araştırmak amacıyla müfredatlarında veri yönetim uygulamaları ile ilgili bilgi sahibi olabilecekleri içerikler olup olmadığını ortaya koymak amacıyla içerik analizi yöntemi kullanılmıştır. İçerik analizinde temelde yapılan işlem, birbirine benzeyen verileri belirli kavramlar ve temalar çerçevesinde bir araya getirmek ve bunları okuyucunun anlayabileceği bir biçimde düzenleyerek yorumlamaktır (Yıldırım ve Şimşek, 2008, s. 227; Neuman, 2012, s. 663). Bu bölümde Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü'ne bağlı ana bilim dallarında verilen ve öğrencilerin bilimsel bir araştırmayı planlama, yürütme ve raporlama becerilerini geliştirmeyi hedefleyen dersler ile veriyle ilgili çalışmaların yapıldığı dersler üzerine odaklanılmıştır. Bu kapsamda “akts.hacettepe.edu.tr” bağlantısı üzerinden içeriği paylaşılan tüm dersler incelenmiş ve araştırmanın odağıyla ilgili olanlar listelenmiştir. “600 Özel Konular” ve “700 Özel Konular” dersleri tez aşamasında geçirilen dönemlere karşılık olarak yüklenilen dersler olduğu için kapsama dahil edilmemiştir. 46 bölüm (%26) için ise belirtilen içeriğe rastlanmamıştır. 131 bölüm (%74) için elde edilen toplam 180 ders yüksek lisans-doktora ve seçmeli-zorunlu olarak Tablo 2’de gösterilmiştir. Yüksek lisans ve doktora programlarında farklı bölümlerde ortak olarak verilen aynı kodlu dersler veri setinde tekrar olmaması için bir ders sayıldıktan sonra analiz için toplam 42 ders sayısına ulaşılmıştır.

**Tablo 2.** Programlarda verilen ve araştırmaya konu olan dersler ile ilgili genel bilgiler

	Yüksek Lisans	Doktora	Toplam
Zorunlu	54	19	73
Seçmeli	89	18	107
Toplam	143	37	180

Elde edilen toplam 42 derse ait dokümanı araştırma projesi süreçleri ve veri yönetim uygulamaları bakımından değerlendirmek üzere temalar belirlenmiştir. Kodlama sırasında dokümanlarda geçen ifadelerin özgün haline bağlı kalınarak verilerin analizinde betimsel yaklaşım benimsenmiştir. MAXQDA 22 paket programı ile analiz edilerek ilişki haritalar çıkarılmıştır.

## 1.5. ARAŞTIRMANIN DÜZENİ

Araştırma beş bölümden oluşmaktadır. Birinci bölümünde konunun önemi, araştırmanın problemi ve amacı, araştırma soruları ve hipotezler, araştırmanın kapsamı, araştırmanın yöntemi, araştırmanın düzeni ve kaynaklar bulunmaktadır.

İkinci bölümde araştırma verisi, açık bilim, araştırma verilerinin yönetimi ve veri paylaşımını teşvik etmek için gerçekleştirilen uluslararası toplantılar, literatür bilgisi olarak sunulmuştur. Veri yaşam döngüsü, araştırmanın temeli olan sağlık verisi üzerinden açıklanmıştır. Ayrıca veri yönetim planı ve araştırma verilerinin yönetimi uygulamaları hakkında bilgi verilmektedir.

Üçüncü bölümde sağlık verisinin tanımı, önemi, özellikleri, analizi, diğer veri türlerinden farkları anlatılmıştır. Sağlık verisi ekosistemi, veri kaynakları ve kullanım alanlarından bahsedilmiştir. Sağlık verilerinin yeniden kullanımının önemi ve paylaşımında en önemli husus olan mahremiyet ve anonimleştirme ile ilgili bilgi verilmiştir. Ayrıca veri okuryazarlığı ve veri okuryazarlığı eğitim programları ile ilgili bilgi verilmektedir.

Dördüncü bölümde nicel ve nitel yöntemlerle toplanan verilerin analizi edilmesiyle ortaya çıkan bulgular tablolar ve haritalar ile açıklanmıştır.

Beşinci bölüm olan sonuç ve öneriler bölümünde karma yöntem ile elde edilen nicel ve nitel analizlerin sonuçları birlikte yorumlanarak önerilerde bulunulmuştur.

## 1.6. KAYNAKLAR

Literatür taraması için Hacettepe Üniversitesi Kütüphaneleri basılı ve elektronik kaynakları ile aşağıdaki elektronik veri tabanlarından faydalanılmıştır:

- Google Akademik,
- Bilgi Dünyası,
- EBSCOHost,
- ScienceDirect,
- PubMed,



- National Library of Medicine,
- ULAKBİM TR-Dizin,
- Web of Science,
- YÖK Tez Veri Tabanı,
- Scopus Veritabanı.

Bu tez çalışmasının yazımında Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü tarafından hazırlanan Tez Yazım Yönergesi (2019) ve kaynak gösterme kuralı olarak American Psychological Association (APA) Kaynak Gösterme Kuralları 6. Basımı'ndan yararlanılmıştır.

## 2. BÖLÜM

### ARAŞTIRMA VERİLERİNİN YÖNETİMİ

#### 2.1. ARAŞTIRMA VERİSİ VE PAYLAŞIMI

21. yyda hesaplamalı bilim olarak adlandırılan üçüncü paradigma ile başlayan çok miktarda veri üretimi “Veri-Yoğun Bilimsel Keşif” olarak adlandırılan dördüncü paradigmaya evrilmiştir. Verilerin sensörler tarafından toplanarak veya simülasyonlar aracılığıyla üretilerek elde edilmesi ve bilgisayarlarda depolanması, veri yoğun bilime yönelik teknik ve teknolojilerin farklılığı, bilimsel keşif için yeni dördüncü bir paradigma olarak veri yoğun bilimi, hesaplamalı bilimden ayırma gerekliliği doğurmuştur. "E-Bilim" olarak da adlandırılan bu dördüncü paradigma teori, deney ve hesaplamayı bir araya getirmekte ve bilimin finanse edilme, iletme ve yayınlanma biçimlerinde değişikliklere yol açmıştır (Bell, Hey ve Szalay, 2009, s. 1297).

Açık bilim hareketi 2000'li yılların başında başlamış ancak hız kazanması 10 yılı bulmuştur. OECD (2015, s. 9) açık bilimi, araştırmacılar, hükümetler, araştırma fonlama kurumları, araştırmacıların kendisi veya kamu desteğiyle gerçekleştirilen araştırmaların sonuçlarının birincil çıktılarını araştırmayı hızlandırmak için bir araç olarak hiçbir kısıtlama olmaksızın veya minimum kısıtlama ile dijital formatta kamuya açık hale getirme çabaları olarak tanımlamaktadır. Munafò ve arkadaşlarına (2017, s. 5) göre açık bilim, içerik üretme ve bir iddiayı kanıtlama sürecini şeffaf ve başkalarının erişimine açık hale getirmektir. Nielsen'a (2011) göre ise farklı alanlardan gelen bilginin araştırma sürecinde mümkün olduğu kadar erken bir aşamada açık bir şekilde paylaşılması gerektiği fikridir. Açık Bilim, geniş anlamda kaynakların ve fikirlerin paylaşılması olarak tanımlanmakta ve bunların gelecekte kullanılmak üzere kamuya açık ve ücretsiz olarak erişilebilir hale getirilmesine vurgu yapmaktadır (Laakso ve diğerleri, 2011).

Alkan (1989) araştırmayı, “verilerin belirli bir amaç için toplanması, analiz edilmesi ve yorumlanarak, araştırmanın amacı doğrultusunda güvenilir çıktılar elde edilmesi süreci”

olarak tanımlanmaktadır. Veri ise, araştırma kapsamında analiz ederek, güvenilir ve özgün sonuçlar elde etmek için toplanan, gözlemlenen veya oluşturulan materyali ifade eder (University of Edinburgh Information Service, 2014). Green, MacDonald ve Rice'a (2009) göre araştırma verisi, orijinal araştırma sonuçları üretmek için analiz etmek amacıyla toplanan, gözlemlenen veya oluşturulan veridir. OECD ise araştırma verisini, bilimsel araştırmalar için kullanılan ve araştırma bulgularını sağlamak için gerekli olan kayıtlar (sayısal skorlar, metinsel kayıtlar, görüntüler ve sesler) olarak tanımlanmaktadır. Bir veri seti, araştırma veya bilim amacıyla, bir konuda gözlemlerin, nesnelere veya diğer varlıkların sistematik ve kısmi bir temsili oluşturur (OECD, 2007, s. 13). Borgman (2015a, s. 29) araştırma verilerini "araştırma amacıyla olguların kanıtı olarak kullanılan varlıklar" olarak ele alır. Veriler, bilimin temel hammaddesidir ve yeniden kullanım için büyük bir potansiyele sahip ulusal ve uluslararası ölçekte değerli bir varlıktır. Araştırma verilerinin yönetimi, herhangi bir araştırma projesinde üretilen verilerin planlanması, toplanması, saklanması, kullanılması, paylaşılması ve korunmasına ilişkin uygulamaları ifade eder (UK Edinburgh, 2020). Corti ve arkadaşları (2019, s. 2) araştırma verilerinin yönetimini veri üreten bir deneyin planlanması, birincil verilerin toplanması, işlenmesi ve analiz edilmesi, verilerin yayınlanması ve paylaşılması, korunması ve yeniden kullanımını içeren bir yaşam döngüsü görevi olarak tanımlanmaktadır. Borgman (2015b, s. 209) ise veri yönetimini, araştırma süreci boyunca ve sonrasında bütünlüğün, erişimin ve kullanılabilirliğin sağlanmasının genel bir çerçevesi olarak görmektedir.

Ortaya çıkışı 1960'lara dayanan araştırma verilerinin paylaşımı, verinin erişilebilirliği ve araştırmanın doğrulanabilirliği noktasında oldukça önemlidir. Bu nedenle hem şeffaflık hem de yeni araştırmalarda tekrar kullanma potansiyelinin, araştırma verilerini paylaşmanın öneminin, araştırmanın kilit paydaşları olan araştırma kuruluşları, fon sağlayan kuruluşlar ve araştırmacılar tarafından tanınması önemlidir (Sturges ve diğerleri, 2015; s. 2446). Minnesota Üniversitesinde 780 araştırmacı üzerinde yapılan bir çalışmada dörtte birinden fazlasının yedekleme eksikliği nedeniyle önemli verilerini kaybettiğini ve %72'sinin verileri yedeklemek için CD, DVD ve flash sürücüler gibi güvenli olmayan ortamlar kullandığını ortaya koymuştur. Araştırmaya katılanların %92'si verilerini öncelikle kendi grupları içinde veya kampüs içinde paylaştıklarını belirtirken üniversitenin bir yılda aldığı fon desteği 769 milyon dolar olarak belirtilmiştir (University of Minnesota President's Emerging Leaders Program, 2009).

Her ne kadar Avrupa’da 2000’lerin başında, Türkiye’de ise son birkaç yılda önemli bir konu haline gelmiş olsa da araştırma verilerinin depolandığı veri bankalarının geçmişte özellikle sağlık alanında çok daha öncesine dayanmaktadır. 1965 yılında kurulan küçük molekülü organik ve metal organik kristallerin yapılarına erişilebilen Cambridge Kristalografik Veri Merkezi, 1971 yılında kurulan, proteinlerin ve nükleik asitlerin üç boyutlu yapılarını sunan Protein Databank’ın ilk versiyonu olan SEARCH (Meyer, 1997, s. 1594), 1974’te kurulan Avrupa Moleküler Biyoloji Laboratuvarı (EMBL), 1987 yılında kurulan Japonya DNA Veri Bankası, bu alanda kullanıma sunulan ilk veri tabanlarındandır.

2000’li yıllara gelindiğinde, küresel internet trafiğindeki hızlı artış ile yıllık veri üretiminin iki katına çıktığı tahmin edilmektedir. Bu dijital devrim, çoğu bilimsel süreçlerin dijitale kaymasının yanında büyük miktarda verinin üretildiği araştırmalarda köklü değişikliklere neden olmuştur. Bilgisayar tabanlı veri toplama araçlarının artan kullanımı, birçok disiplinlerdeki araştırma verilerinin ölçeğini genişletmiştir. Küresel ekonomik krizle birlikte de veri ekonomik bir değere dönüşmüştür. Bilimin yürütülme ve iletme biçiminde makalelerden elde edilen sonuçların çoğunun yeniden üretilmemesiyle ortaya çıkan güvenilirlik ve tekrarlanabilirlik sorununun farkına varılması, açık bilim hareketini başlatmıştır (Martone ve Nakamura, 2022, s. 2).

Budapeşte Açık Erişim Girişimi “açık erişim”in tanımını ilk kez 2002 yılında “hakem değerlendirmesinden geçmiş bilimsel literatürün, internet aracılığıyla; finansal, yasal ve teknik engeller olmaksızın, serbestçe erişilebilir, okunabilir, indirilebilir, kopyalanabilir, dağıtılabilir, basılabilir, taranabilir, tam metinlere bağlantı verilebilir, dizinlenebilir, yazılıma veri olarak aktarılabilir ve her türlü yasal amaç için kullanılabilir olması” olarak açıklamıştır. Açık erişimin teşviki için tamamlayıcı stratejiler önermiş, tüm disiplinlere ve ülkelere açık erişim için genel bir çağrı yaparak önemli bir fon desteği sağlamıştır (Budapeşte Açık Erişim Girişimi, 2012).

Budapeşte toplantısından kısa bir süre sonra, açık erişim hakkında iki kamuoyu bildirisi daha yayınlandı. Bethesda bildirisi 2003 yılında Howard Hughes Tıp Enstitüsü (ABD), BMJ Yayın Grubu (İngiltere), Harvard Tıp Fakültesi (ABD), Amerikan Hücre Biyolojisi Derneği (ABD), Rockefeller Vakfı (ABD) ve Wellcome Trust tarafından onaylanmıştır. Berlin Bildirgesi ise aynı sene aralarında Fransa’dan Centre National de la Recherche

Scientifique (CNRS) ve Almanya'dan Max Planck Society (MPS) gibi büyük ulusal araştırma kurumlarının da bulunduğu 60'tan fazla kuruluş tarafından imzalanmıştır. 2005 yılında ise Ulusal Sağlık Enstitüsü (NIH) hakemli makalelerin kabul edildikten sonra "mümkün olan en kısa sürede", ancak en geç 12 ay içinde açık erişim deposu PubMed Central'a (PMC) gönderilmesini talep etmiştir (Bosch, 2008, s. 404). 2013 yılından bu yana ise, Avrupa Komisyonu, ABD Bilim ve Teknoloji Politikası Ofisi ve Küresel Araştırma Konseyi de dahil olmak üzere küresel bilimsel kuruluşlar, araştırmaya kamu erişiminin artırılmasını destekleyen politikalar benimsemişlerdir (Gewin, 2016, s. 117).

Ülkemizde araştırma verilerinin paylaşımı ve açık bilim konusunda dahil olunan ilk çalışmalar Anadolu Üniversite Kütüphaneleri Konsorsiyumu'nun (ANKOS) Açık Erişim ve Kurumsal Arşivler (AEKA) Çalışma Grubu'nun kurulması, OpenAIRE2020, MedOAnet, PASTEUR4OA ve Ufuk2020 projeleri olarak sıralanabilir. 2014 yılında Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü ve Goethe-Institute Ankara, "Uluslararası Araştırma Verilerinin Yönetimi Çalıştayı"nı gerçekleştirmişlerdir. 2017 yılında TÜBİTAK, "Araştırma Verilerinin Yönetimi Eğitim Programı"nı sunmuştur. Aynı sene "Açık Veri Portalı Çalıştayı" gerçekleştirilmiştir. 2018 yılında ise Boğaziçi Üniversitesi'nde "Araştırma Verilerinin Yönetimi Çalıştayı" gerçekleştirilmiştir. 2019 yılında bu kez İzmir Yüksek Teknoloji Enstitüsü "Açık Bilim Yolunda: Bilimi Özgür Bırak" isimli seminer düzenlemiştir. Yine 2019 yılında TÜBİTAK "Açık Bilim Politikası"nı yayınlamış ve YÖK üniversitelere "Açık Bilim Politikası" hazırlama zorunluluğu getirmiştir. 2021 yılında "Ulusal Araştırma Verileri Sempozyumu" ile "Açık veri Yönetimi ve Paylaşımı Çalıştayı" gerçekleştirilmiştir.

Veri paylaşımı, bir yandan politikalar aracılığıyla hız kazanırken, diğer yandan araştırmacılar için hala birçok soru işareti barındırmaktadır. Açık erişim anlayışının faydaları ve açık erişimi destekleme çabaları, kuruluşları fon sağladıkları araştırmaların açık erişim havuzlarında kamuya açık olmasını zorunlu kılan politikalar geliştirmeye teşvik etmiştir. Bilimsel araştırmaların tüm çıktıklarına erişim imkanı, veri paylaşımı ve verilerin yeniden kullanımı ile daha fazla araştırma üretilmesini sağlayarak, araştırma sistemini daha etkin ve verimli hale getirme potansiyeline sahiptir (Piwowar ve Vision, 2013). Corti ve arkadaşları (2019, s. 3) araştırma verilerinin yeniden kullanımının faydalarını paydaşlar açısından değerlendirmiş ve şöyle listelemişlerdir:

- Bilimsel sorgulama ve tartışmayı teşvik eder,
- Yenilikçiliği ve potansiyel yeni veri kullanımlarını teşvik eder,
- Veri kullanıcıları ve veri yaratıcıları arasında yeni iş birliklerine yol açar,
- Şeffaflığı ve hesap verebilirliği en üst düzeye çıkarır,
- Araştırma bulgularının incelenmesini sağlar,
- Araştırma yöntemlerinin iyileştirilmesini ve doğrulanmasını teşvik eder,
- Veri toplama işleminin tekrarlanmasının maliyetini azaltır,
- Araştırmanın etkisini ve görünürlüğünü artırır,
- Verileri ve sonuçlarını oluşturan araştırmayı teşvik eder,
- Kendi başına bir araştırma çıktısı olarak araştırmacıya doğrudan bir kredi sağlayabilir,
- Eğitim ve öğretim için önemli kaynaklar sağlar.

Açık Bilgi Vakfı (Open Science Foundation) 2010 yılında yayınladığı Panton İlkeleri ile verinin açıklığı konusuna ışık tutmuştur. Bilimsel araştırma veri setlerinin yasal statüsünün açıkça belirtilmesi gerektiğini ve lisanslamanın doğru şekilde yapılması gerektiğini vurgulamıştır. Daha sonra ilkeler Açık Bilgi Vakfı - Açık Veri Çalışma Grubu tarafından geliştirilmiştir (Gray, 2010).

Wilkinson ve arkadaşları (2016), iyi bir veri yönetiminin kendisinin bir hedef olmaktan ziyade verinin paylaşılma sürecinden sonra araştırmacılar tarafından yeniden kullanımını sağlayan temel kanal olduğunu ve mevcut dijital ekosistemin araştırma verilerinin paylaşımına, dolayısıyla araştırma yatırımlarından maksimum fayda sağlamaya engel olduğunu belirtmişlerdir. Bu sebeple araştırma verilerinin yeniden kullanılmasını destekleyen bir altyapı geliştirmek için akademi, sanayi, fon kuruluşları ve yayıncıları temsil eden paydaşlarla FAIR (Findability, Accessibility, Interoperability, Reusability) olarak adlandırdıkları aşağıdaki ilkeleri yayınlamışlardır.

*Bulunabilirlik (Findability)*: Bu ilkeye göre veri standart ve tekil bir tanımlayıcı ile kalıcı olarak tanımlanarak, Google gibi arama motorlarında tarama yapıldığında açık olarak erişilebilir olmalıdır. Ayrıca veri, hakkında detaylı bilgi verebilecek zengin üst veri ile tanımlanmalı ve bu üst veri makine tarafından okunabilir olmalıdır (Lamprecht ve diğerleri, 2020, s. 41)

*Erişilebilirlik (Accessibility):* Erişilebilirlik ilkesi orijinal kaynağa yani veriye artık erişilemediğinde bile standartlaştırılmış bir iletişim protokolü ve üst verilere erişebilmeyi ifade eder. Protokol açık, serbest ve evrensel olarak uygulanabilir olmalı ve üst verileri doğrudan veri deposundan veya web sitesinden indirilebilmelidir (Lamprecht ve diğerleri, 2020, s. 45).

*Birlikte Çalışabilirlik (Interoperability):* İki veya daha fazla sistem veya bileşenin bilgi alışverişinde bulunabilme ve bu bilgi alışverişini kullanabilme yeteneği olarak tanımlanmaktadır. Üst veri, bir kaynağın bulunabilirliğini, yeniden kullanılabilirliğini, yorumlanmasını ve değerlendirilmesini sağlama amacına hizmet edebilecek herhangi bir kaynak tanımıdır. Bu ilke üst verilerin makine tarafından kolayca bulunabilir ve okunabilir olmasını ifade eder (Lamprecht, 2020, s. 46).

*Yeniden Kullanım (Reusability):* Yeniden kullanılabilirlik, bir kişinin yazılımı tekrar üretilebilir şekilde yeniden kullanabilmesini amaçlar. Yeniden kullanım ilkesi üst veri öğelerinin makine tarafından işlenebilir ve doğrulanabilir olmasını, kullanımının kolay olmasını ve veri paylaşımını sürdürmek için verilerin atıf yapılabiliyor olmasını ifade eder (Lamprecht, 2020, s. 49).

## 2.2. VERİ YAŞAM DÖNGÜSÜ

Veriler de bir tür yaşam döngüsü boyunca ilerlemektedir ve bu kavram genellikle araştırmacıların veri yönetiminin kapsamını ve anlamını anlamalarına yardımcı olmak için kullanılır. Araştırma verileri, toplanmasından, yaşam döngüsünün tamamlanmasına kadar formatında, uygulamasında, kullanımında ve hatta belki de amacında birçok değişikliğe uğramaktadır. Verilerin yaşam döngüsü içinde geçireceği dönüşümlerin iyi tanımlanması ve adlandırılması kuruluşların, ihtiyaçlara yönelik olarak daha iyi hedefler belirleyebilmesini sağlar (Carlson, 2014, s. 63). Veri yönetimi faaliyetlerini bilimsel araştırma süreçleriyle ilişkilendirmek ve yaşam döngüsü modelini araştırma faaliyetlerine uygulamak, verilerin iyi tanımlanmasını, korunmasını, erişilebilir olmasını ve yeniden kullanıma uygun olmasını sağlar. Ayrıca bilimsel araştırma verilerini yönetmeye yönelik politika ve uygulamaların değerlendirilmesine, geliştirilmesine ve değişikliğe ihtiyaç

duyulan alanların belirlmesine yardımcı olacak bir yapı olarak hizmet eder (Faundeen ve diğeri, 2013).

Literatürde çeşitli veri yaşam döngüsü modelleri bulunmakta olup bu çalışmada veri yaşam döngüsü araştırmaya temel olan sağlık verisi üzerinden açıklanmıştır. Harvard Üniversitesi Longwood Tıp Fakültesi'ne bağlı olarak araştırmacıların mevcut ve gelecekteki araştırma veri yönetimi ihtiyaçlarını karşılamak için rehberlik, kaynak ve çözümleri paylaşmak ve en iyi uygulamaları geliştirmek amacıyla Araştırma Veri Yönetimi Çalışma Grubu kurulmuştur. Grup yüksek hacimli tarama ve görüntüleme verilerinin yönetimi, bilgi teknolojisi, eğitim programlama ve kütüphane bilimleri gibi çeşitli alanlarda araştırmacılara destek sunmaktadır. Yayınladıkları veri yaşam döngüsü modeli Şekil 2'de görülmektedir (Harvard, 2022).



**Şekil 2.** Biyomedikal veri yaşam döngüsü (Harvard, 2022)

Biyomedikal veri yaşam döngüsünün süreçleri şu şekilde tanımlanmıştır (Harvard, 2022);

Planlama: Veri yaşam döngüsünün ilk aşaması planlamada geçerli veri politikalarının, bu politikalara uyumluluğun ve klinik veri yönetimine ilişkin tüm hususların gözden



geçirilmesi gerekmektedir. Bu aşamada planlamaya referans olması için bir kontrol listesi oluşturulması önerilmektedir. Planlanması önerilen konular şunlardır;

- Geçerli veri politikalarını, bu politikalara uyumluluğu ve klinik veri yönetimine ilişkin tüm hususların gözden geçirilmesi,
- Standartlaştırılmış dizin yapıları ve dosya adlandırma kuralları geliştirilmesi,
- Veri yönetimi ve paylaşımının sağlanması için roller ve sorumluluklar atanması.

*Veri Toplama:* Bu aşama veri kümelerinin ve toplama süreçlerinin organizasyonunu ve entegrasyonunu içerir. Ne tür veriler üretileceği, veri dokümantasyonu ve üst veriler için hangi standartların kullanılacağı belirlenir. Veri toplama sürecinin başında yapıyı tanımlamak, tüm veri ve belge dosyaları için organizasyonel bir temel sağlar.

*Analiz:* Verinin analize hazır hale getirilmesi süreçlerini ve dokümantasyonunun yapılmasını içeren aşamadır. Toplanan verinin, sayısallaştırması, kontrol edilmesi, doğrulanması, temizlenmesi ve anonimleştirilmesi, verilerin tanımlanması ve analizini kapsar. Bu işlemler veri kümeleri üzerinde iyi bir şekilde gerçekleştirildiğinde verilerin analizi mümkün olduğunca net, tutarlı ve hatasız sonuçlar verir.

*Değerlendirme ve Arşivleme:* Verilerin korunması ve uzun vadeli erişimi için nasıl arşivleneceği, verinin türü, fon sağlayıcısı ve/veya üniversite için saklama gereksinimlerinin neler olduğunun belirlenmesi ve arşivlenmesi süreçlerini içerir.

*Paylaşım:* Veri paylaşımında hangi politikaların göz önünde bulundurulacağı verinin paylaşılacağı veri havuzunun değerlendirilmesi ve seçimi aşamalarını kapsar.

*Yeniden Kullanım:* Veriye nasıl erişileceği, nasıl yeniden kullanılacağı, entelektüel mülkiyet hakları, veri kullanım anlaşmalarının belirlenmesi aşamalarını kapsar.

*Depolama ve Yönetim:* Biyomedikal Veri Yaşam Döngüsü'nün her aşaması veri depolama ve yönetimi etrafında şekillenir. Verilerin güvende kalmasını ve önerilen güvenlik protokollerine uyulmasını sağlamak için yaşam döngüsü boyunca uygun depolama bakımı zorunludur. Araştırma verilerinin depolanması ve iletilmesi için giderek daha karmaşık düzenlemeler gerekmektedir. Araştırma öznelinin mahremiyetini

korumak ve hassas, kişisel olarak tanımlanan bilgilerin güvenliğini sağlamak için güçlü bir veri gizliliği ve güvenlik planlaması gereklidir.

Yaşam döngüsünün ilk üç aşaması araştırma tamamlanana kadar olan veri süreçlerini içerirken diğer aşamalar, araştırma tamamlandıktan sonra verilerin korunmasını, paylaşılmasını ve son olarak yeni çalışmalar yapmak veya orijinal sonuçların tekrarlanabilirliğini test etmek için verilerin yeniden kullanılmasını içerir. Veri yönetim planı, veri yaşam döngüsünün aşamaları ile benzer süreçleri içermektedir ve bu süreçlerin bir politikaya dayanarak dokümente edilmesi olarak tanımlanabilir.

### 2.3. VERİ YÖNETİM PLANI

Veri yönetim planı bir araştırma projesinin planlama aşamasında geliştirilen ve araştırma verilerinin proje sırasında nasıl toplanacağı, üretileceği ve işleneceğinin yanı sıra daha sonra nasıl, nerede saklanacağı ve paylaşılacağını detaylandıran resmi bir belgedir (Harvard, 2022). Veri yönetim planları, araştırmacıların verilerini yönetmelerine yardımcı olan ve araştırma sona erdikten sonra verilerin yüksek kalitede, erişilebilir ve yeniden kullanılabilir olmasını sağlayan araçlardır. Ayrıca araştırmacının veri ile ilgili süreçlerinin düzenli, anlaşılır ve şeffaf olmasını sağlar. Smale ve arkadaşları (2018), veri yönetim planlarının ilk olarak 1966 yılında, karmaşık havacılık ve mühendislik projelerinde kullanıldığını, 1970'lerin sonlarından 1980'lere kadar, veri yönetim planı kullanımının çeşitli disiplinlerde yaygınlaştığını belirtmişlerdir. Veri yönetim planları, karmaşık projelerin veri toplama ve analiz aşamaları sırasında veri yönetimi gereksinimlerini karşılmasına yardımcı olmak için proje yönetimi araçları olarak kullanılmaya devam etti.

2003 yılında Amerikan Sağlık Enstitüsü (NIH) araştırma sonuçlarının insan sağlığını iyileştirecek bilgi, ürün ve prosedürlere dönüştürülmesini hızlandırma motivasyonu ile başvuru sahiplerinden veri yönetim planı talep etmeye başlamıştır (Thoegersen, 2015, s.2).

2004 Ocak ayında OECD ülkelerinin bilim ve teknoloji bakanları Paris'te bir araya gelerek araştırma verilerine erişim konusunda uluslararası kılavuz ilkelere duyulan ihtiyacı tartışmışlardır. Bu toplantıda 30 OECD ülkesi ile Çin, İsrail, Rusya ve Güney Afrika hükümetleri "OECD Kamu Destekli Araştırma Verilerine Erişim Bildirgesi"ni

kabul etmişlerdir. Bu bildirge ile açık erişim verilerinin olumlu etkisini kabul ederek imzacı ülkelerde kamu tarafından finanse edilen araştırmaların verileri için açık erişim düzenlemeleri üzerinde çalışma taahhüdü vermiştir. Ayrıca OECD'yi "ortak kabul gören ilkelere dayalı bir dizi OECD yönergesi geliştirmeye" davet etmişlerdir.

2005 yılında, İngiltere’de altı büyük araştırma fonu kuruluşu (Biotechnology and Biological Sciences Research Council, Department of Trade and Industry, Joint Information Systems Committee for Support for Research, Medical Research Council, Natural Environment Research Council, Wellcome Trust) değişen veri ortamını daha fazla sorgulamak için bir danışmanlık finanse etmiştir. Bu danışmanlık raporu, İngiltere’de fon kaynaklarının veri yönetim planlarının fon başvurusu ile sunulmasının zorunlu tutulmasını öneren ilk rapordur (Smale ve diğerleri, 2018, s. 6).

2009 yılında Amerika’da Dijital Veriler Üzerine Kurumlar Arası Çalışma Grubu (IGWWD) verinin korunmasını ve paylaşılmasını en üst düzeye çıkarmak amacıyla bir çerçeve veri yönetim planı hazırlayarak fon sağlayıcıların proje sahiplerini veri yönetimini planlamaya teşvik etmelerini tavsiye eden bir rapor yayınlamıştır (Thoegersen, 2015, s. 2)

2011 yılında Amerika’da ilk defa veri yönetim planı hazırlama zorunluluğu National Science Foundation tarafından sağlanmıştır (National Science Foundation, 2011). Aynı sene UK Research and Innovation ise halkın araştırma, teknoloji ve inovasyon ile ilgili karar alma süreçlerine katılımını sağlamak için 7 ilke yayınlamıştır.

2013 yılında ise Avrupa Konseyi Ufuk 2020 programı ile veri yönetim planı hazırlama zorunluluğu getirmiştir. Yine 2015 yılında Birleşik Krallık'ta, toplanan ve üretilen araştırma verilerinin, ilgili yasal, etik, düzenleyici çerçeveler içerisinde tutarlı bir şekilde ve maliyetler göz önünde bulundurularak mümkün olan her yerde başkalarının kullanımına açık hale getirilmesini sağlamak amacıyla çok paydaşlı bir grup tarafından on ilkeye dayanan Açık Araştırma Verileri Konkordatosu yayınlanmıştır (Concordat on Open Research Data, 2016).

Çeşitli kuruluşlar bir araya gelerek araştırmacıların ve diğer veri paydaşlarının büyük araştırma fon kuruluşlarının gereksinimlerine göre veri yönetim planları geliştirebilmeleri

için DMP Online ve DMPTools'u kullanıma sunmuştur (DMPonline, 2022; DMPTool, 2022). Veri yönetim planları fon sağlayıcı kuruluşların beklentilerine, ihtiyaçlarına veya araştırmacı ve araştırmaya göre değişebileceği gibi, tüm veri yönetim planlarının içermesinin beklendiği ortak alanlar mevcuttur. Veri yönetim planının içermesi beklenen ortak alanlar şunlardır (Borgman, 2015a, s. 35):

- Araştırmanın amacı doğrultusunda toplanacak verilerin neler olduğu,
- Verilerin düzenlenmesi (üst veriler, standartlar vb.),
- Verilerin paylaşımı ile ilgili detaylar,
- Hukuki ve etik konular, veri paylaşımına ilişkin kısıtlamalar,
- Telif hakları,
- Verilerin depolanma ve yedeklenmesi ile ilgili önlemleri,
- Veri yönetiminden sorumlu kişiler, kurumlar,
- Maliyet ve kaynak ihtiyacı.

Kurumsal düzeydeki politikalar, araştırmacılar arasında veri paylaşımı kültürünün zorunlu kılınmasına yardımcı olur. Politika, veri paylaşımı bağlamında şunları açıkça belirtmelidir: amaç, kapsam, uygulanabilirlik ve veri gönderimi, lisanslama, üst veri girişi, veri sınıflandırması, telif hakkı anlaşmaları ve varsa veri geri çekme taleplerinin değerlendirileceği koşullar, verilerin kullanım şartları ve koşulları, hassas verilerin gizliliğinin korunması, verilerin güvenlik ihlallerine karşı korunması ile ilgili olarak veriye katkıda bulunanlara yönelik kılavuz ilkeler, fikri mülkiyet endişeleri vb. ile ilgili olarak veriye katkıda bulunanlara yönelik yönergeler (Patel, 2016, s. 228).

#### **2.4. ARAŞTIRMA VERİLERİNİN YÖNETİMİ UYGULAMALARI**

Araştırma verilerinin yönetimi bir proje boyunca aktif olarak gerçekleşen bir dizi süreç ve proje yönetimi stratejisidir. Bu süreçlerin birçoğu bu teknikler üzerine inşa edilir. Bunun nedeni, bilimsel araştırma gruplarının etkili bir şekilde işbirliği yapabilmeleri için, veri yönetimi prosedürleri de dahil olmak üzere bir projedeki görevleri ve ekibin nasıl çalışacağı konusunda ortak bir anlayış oluşturmaları gerektiğidir. Bu nedenle, iyi veri yönetimi uygulamaları açıkça tanımlanmış rol ve sorumlulukları da içerir (Briney, Coates ve Goben, 2020). Veri yoğun bilimin potansiyelinin gerçekleştirilmesi için, araştırma

ürünlerinin keşfedilmesi ve kullanılması için uygun sistemler, hizmetler, araçlar, içerik, politikalar, uygulamalar ve insan kaynakları gereklidir (Borgman ve diğerleri, 2015b, 208).

Araştırma verilerinin yönetimi uygulamalarını, veri yaşam döngüsü ve veri yönetim planından bağımsız olarak düşünmek yanlış olur. Yaşam döngüsü boyunca verinin geçireceği işlemler planlanarak veri yönetim planı oluşturulur ve bu planda belirtilen işlemler veri yönetim uygulamaları ile gerçekleştirilirler. Başarılı AVY uygulamaları, araştırmanın bütünlüğünü ve şeffaflığını sağlayarak araştırma verilerinin etkisini ve erişimini en üst düzeye çıkarır (Neylon, 2017). Uygulamaların anlaşılması, kurumların hizmetleri ve sağladıkları destek için strateji geliştirmelerinde yardımcı olacak ve ileriki araştırmalar için veri yönetimi uygulamalarında temel alanların ana hatlarının çizilmesine katkı sağlayacaktır (Krahe ve diğerleri, 2020, s. 110). Veri paylaşımı gerektiren ve giderek daha şeffaf hale gelen bir klinik araştırma altyapısı, kanıta dayalı tıbbın temeli olan kanıt sentezinin geleceği için büyük umut vaat etmektedir (Ross, 2016). Veri yönetim uygulamaları planlama, veri toplama ve düzenleme, depolama ve koruma, yayınlama ve paylaşım ile yeniden kullanım olmak üzere beş süreçten oluşur.

#### **2.4.1. Planlama**

Uzun vadeli, verimli ve uygun maliyetli, veri keşfine ve verilerin yeniden kullanımına olanak tanıyan bir araştırma ve veri ortamı oluşturmak, uzun vadeli planlama yaklaşımlarının benimsenmesini gerektirir (National Academies of Sciences, Engineering and Medicine, 2020, 118). Veri yönetim planı bir kurumun içinde araştırma projesinin tarafları arasında gayri resmi olabileceği gibi fon sağlayıcı kuruluşlar veya projeler resmi bir veri yönetim planı sunulmasını gerektirebilir (Harvard, 2022). Plan yapmak zaman, emek ve ekonomik açıdan tasarruf edilmesine yardımcı olur ve araştırma sürecini kolaylaştırır. Planlama ayrıca daha geniş bağlamı ve farklı seçeneklerin sonuçlarını göz önünde bulundurarak sağlıklı kararlar alınmasını sağlar (DCC, 2019). Hassas veriler kullanıldığında, plan ayrıca verileri güvenli ve yönetmeliklere uygun hale getirmek için hangi adımların atıldığını da açıklamalıdır. Proje sırasında ve bittikten sonra verilerin paylaşım ve yeniden kullanım için nasıl erişilebilir olacağı ve belgeleneceği ayrıntılı olarak açıklanır (Briney, Coates ve Goben, 2020). Harvard Üniversitesi

Araştırma Veri Yönetim Merkezi veri yönetim planı hazırlarken cevaplanması gereken soruları şöyle listelemiştir (Harvard, 2022):

- *Veri türleri:* Verilerinizin kaynağı nedir? Verileriniz hangi formatlarda? Verileriniz sabit mi olacak yoksa zaman içinde değişecek mi? Projeniz ne kadar veri üretecek?
- *Üst veriler:* Verilerinizi nasıl belgeleyecek ve açıklayacaksınız?
- *Depolama, yedekleme ve güvenlik:* Verilerinizi nasıl ve nerede depolayacak ve güvence altına alacaksınız?
- *Koruma/gizlilik için hükümler:* Hangi gizlilik ve mahremiyet konularını ele almalısınız?
- *Erişim ve paylaşım:* Verilerinize diğer araştırmacıların erişimini nasıl sağlayacaksınız? Başkaları verilerinizi nasıl keşfedecek?
- *Arşivleme ve erişim sağlama:* Verileri korumak ve uzun vadeli erişim sağlamak için planlarınız nelerdir?
- *Roller ve plan gözetimi:* Proje boyunca veri yönetiminin unsurlarından kim sorumlu olacak ve uygulama için hangi kaynaklar gerekli?

#### 2.4.2. Veri Toplama ve Düzenleme

Hem örneklem seçimi hem de veri toplama sürecinin kendisine ilişkin kararlar, veri kalitesini ve nihayetinde varılabilecek sonuçları etkiler (Wilson ve Miller, 2014, s. 33). Klinik verilerin yeniden kullanımını birçok potansiyel avantaj sunmasına rağmen, veri kalitesi sorunları, veri entegrasyonu ve birlikte çalışabilirlik sınırlamaları gibi birçok zorlukla da karşı karşıyadır (Meystre ve diğerleri, 2017).

Verilere erişimi en üst düzeye çıkarmak, dosyaların düzenli kalmasını sağlamak ve dosyaların tanımlanmasını kolaylaştırmak için bir adlandırma kuralı oluşturulması gerekir. Bu kural, dosyaların ne içerdiğini ve diğer dosyalarla nasıl ilişkili olduğunu açıklayan bir çerçevedir. Dosya adlandırma kuralları aşağıdaki formatlarda belirlenebilir (Harvard, 2022):

- Deney koşulları

- Veri türü
- Araştırmacı adı/unvanı
- Laboratuvar adı/konumu
- Proje veya deney adı veya kısaltması
- Deneyin tarihi veya tarih aralığı
- Tarih atamaları için YYYYAAAGG iyi bir formattır. Bu biçim, tüm dosyaların kronolojik sırada kalmasını sağlar. Dosya adına bir zaman damgası eklemek için YYYYAAAGGZssdd biçimi kullanılabilir.
- Deney numarası veya numune kimliği
- Sıralı bir numaralandırma sistemi kullanırken, netlik için baştaki sıfırları kullanın ve dosyaların sıralı olarak sıralandığından emin olun. Örneğin, "1, 2, ...10, 11 ... 100, 101 ..." yerine "001, 002, ...010, 011 ... 100, 101 ..." kullanın.
- Bir dosyanın en güncel sürümünü belirtmek için sürüm isimleri kullanılması
- Dosya adının sonuna sürüm bilgisi ekleyerek bir dosyanın sürümleri belirtilir, örn. dosya\_adi\_v2.xxx
- Bir sürüm numarası kullanılabilir (örn. "v01" veya "v02")
- Sürüm tarihi kullanılabilir (ISO 8601 formatını kullanın: YYYYAAAGG veya YYYY-AA-GG)
- Tire (-), alt çizgi (\_) kullanın veya her kelimenin ilk harfini büyük yazın
- Kısa çizgiler: dosya-adi.xxx
- Küçük harfler: dosya\_adi.xxx
- Ayrılma yok: dosyaadi.xxx
- Baş harfi (metnin her bölümünün ilk harfi büyük yazılır): DosyaAdı.xxx
- Özel karakterlerden kaçınılmalı: ~ ! @ # \$ % ^ & \* ( ) ` ; : < > ? . , [ ] { } ' " |
- Her farklı veri kümesi için bir klasördeki diğer dosyalar hakkında bilgi içeren düz metin veya BENİOKU dosyaları oluşturma

Verilerin tutarlı bir şekilde tanımlanması için çeşitli üst veri standartları vardır. Üst veriler verilerin başkaları tarafından anlaşılmasını ve keşfedilmesini sağlar ve bu geniş dokümantasyonun verileri ayrıntılı olarak tanımlayan bir alt kümesidir (DCC, 2022). Üst veriler, bir bilgi kaynağını tanımlayan, açıklayan, konumlandıran veya başka şekilde erişimi, kullanmayı veya yönetmeyi kolaylaştıran yapılandırılmış bilgilerdir. Üst veri,

genellikle veri hakkındaki veri olarak tanımlanır. Verilerin nasıl ve neden oluşturulduğuna ilişkin bağlamsal ayrıntıları yakalamak esastır. Üst veriler, verilerinizin nasıl oluşturulduğuna, analiz edildiğine ve saklandığına ilişkin bağlamın açık, ayrıntılı ve dolayısıyla tekrarlanabilir olmasını sağlar. Amerika Birleşik Devletleri Ulusal Bilgi Standartları Organizasyonu (NISO), üst verileri dört gruba ayırmaktadır (Riley, 2017, s.6);

**Tablo 3.** Üst Veri Türleri

Türü	Amacı
Tanımlayıcı üst veriler	Bir kaynağı bulmak veya anlamak için
İdari üst veriler	
Teknik üst veriler	Dosyaların kodunu çözmek ve işlemek için
Koruma üst verileri	Dosyaların uzun vadeli yönetimi
Haklar üst verileri	İçeriğe bağlı fikri mülkiyet hakları
Yapısal üst veriler	Kaynakların parçalarının birbirleriyle ilişkileri
İşaretleme dilleri	İçerikteki diğer yapısal veya anlamsal özellikler için meta verileri ve işaretleri entegre eder

İyi tanımlanan üst veriler, verileri şimdi ve gelecekte anlamayı, kullanmayı ve paylaşmayı sağlayarak diğer araştırmacıların verileri uzun vadede keşfetmesine, erişmesine, yeniden kullanmasına ve alıntılmasına yardımcı olur. Ayrıca verilerin uzun vadeli arşivlenmesini de kolaylaştırır. Biyomedikal verilere kolayca erişebilmek için kullanılan üst veri türleri ise şunlardır (Harvard, 2022b):

*Reaktif Üst Veriler:* Verileri oluşturmak için kullanılan klinik numuneler, biyolojik reaktifler (örn. hücre hatları, antikolar, RNA'lar), kimyasal reaktifler (örn. ilaçlar) vb. hakkında bilgiler.

*Teknik Üst Veriler:* Araştırma cihazları ve ilgili yazılımlar tarafından otomatik olarak oluşturulan bilgiler.

*DeneySEL Üst Veriler:* DeneySEL koşullar (örn. tahlil türü, zaman noktaları), deneySEL protokol ve verileri oluşturmak için kullanılan ekipman hakkında bilgiler.



*Analitik Üst Veriler:* Yazılım adı ve sürümü, kalite kontrol parametreleri ve çıktı dosyası türü ayrıntıları dahil olmak üzere veri analizi yöntemleri hakkında bilgiler.

*Veri Kümesi Seviyesi Meta Verileri:* Araştırma projesinin hedefleri, katılımcı araştırmacılar, ilgili yayınlar ve fon kaynakları hakkında bilgiler.

Üst veri standartları arasında en çok kullanılan Dublin Core Metadata Initiative (DC) ve Data Documentation Initiative (DDI) olsa da pek çok alana özgü araştırmacıların yararlanabileceği üst veri standartları mevcuttur (Metadata Standards Directory, 2022). Sağlık bilimleri alanında Darwin Core, Genome Metadata, ISA-Tab, MIBBI, NeXus, OME-XML, PDBx/mmCIF, FHIR standartları örnek olarak gösterilebilir.

### **2.4.3. Depolama ve Koruma**

Veri paylaşımı ve yeniden kullanımı, veriyi depolamayı gerektirir ve açık veri politikalarının süreklilik varsayımları, kurumların uzun vadeli desteğini içermelidir (Bishop ve diğerleri, 2020, s. 4). Disipline özgü veri depolarının gücü, veri türleri ve araştırma bağlamları arasındaki tutarlılıkta yatmaktadır. Dolayısıyla, verilerin yeniden kullanımını destekleyen ve böylece bilgi keşfi ve analiz hizmetlerine yardımcı olan prosedürler geliştirmek mümkündür. Veri depoları, belirli bir veri seti hakkında kullanıcı davranış verilerini ve yorumlarını (veri setinin arama sonuçlarını alıntılara göre sıralamak, indirme sayısı ve veriyi 'popüler' olarak etiketlemek gibi) toplayarak referans olarak sunar (Duan, Wang ve Song, 2022).

Verileri yedekleme ve depolamada yaygın kabul gören yöntem 3-2-1 kuralıdır. Bu kurala göre verilerin en az üç kopyası olmalı ve iki kopyası en az iki farklı tür saklama ortamında saklanmalıdır. Verilerin bir kopyası ise dış ortamda yedeklenmelidir. Bu yöntem ile hem çalınmaya karşı hem de yedekleme ortamının bozulması veya teknolojisinin eskimesine karşı önlem alınmış olur. Araştırma tamamlandıktan sonra ise araştırma veri çıktılarını yayınlamak için bir arşiv veri deposu seçilebilir. Veri havuzları, verilerin saklanması ve gelecekte veriyi üreten araştırmacı veya diğer araştırmacılar tarafından kullanılmak üzere erişilebilmesini sağlar. Verilerin ve diğer çıktıların uygun şekilde arşivlenmesiyle, araştırmaya atıfta bulunulması, yeniden kullanılması ve arama motorlarında keşfedilmesi daha olasıdır (Berkley Library, 2022).

#### 2.4.4. Yayımlama ve Paylaşım

Araştırma materyallerinin projenin aktif aşamasındayken paylaşılacak üzere hazırlanması, yıllar öncesine ait çalışmaların yeniden yapılandırılmasından çok daha kolaydır. Bu uygulama, daha sonra yeniden kullanımı kolaylaştırma konusunda kişinin kendisine sağladığı doğrudan faydaların yanı sıra, yayınlanan bulguların yeniden üretilebilirliğini ve diğer araştırmacıların bu çalışmayı kullanma, genişletme ve atıfta bulunma kolaylığını artırır (Gorgolewski ve Poldrack, 2016).

Verinin yeniden kullanımı için açık telif hakkı beyanlarına ve veri lisanslarına sahip olmak çok önemlidir. Her araştırma projesinin verilerin sahipliğiyle ilgili hüküm ve koşulları açıkça tanımlaması gerekir. Buna alternatif olarak, veriler bir veri havuzunda saklandığında, veri havuzlarının kullanıcılarla karmaşık lisans anlaşmaları yapması zor olabileceğinden, bir ön koşul olarak her veri setiyle birlikte telif hakkı beyanının açık bir şekilde tanımlanması ve izin verilen kullanımların belirtilmesi gerekebilir. Herhangi bir belirsizliği ortadan kaldırmak için araştırmacı, veri üretimine katılan herkesin elde edebilecekleri her türlü telif hakkını bireysel araştırmacıya devretmesini isteyebilir (Patel, 2016, s. 230).

Verilerin kullanımını etkin bir şekilde teşvik etmek amacıyla, veri havuzları verilerin ticari olmayan kullanımına izin veren bir mekanizmaya sahip olmalıdır. Bunlar Open Data Commons (ODC) (Open Knowledge Foundation, t.y.) modeline ya da Creative Commons (CC) (Creative Commons, t.y.) modeline dayanabilir. Bu iki lisanslama modeli ve açıklamaları aşağıdaki Tablo 4'te gösterilmektedir.

ODC modeli altında, türev çalışmaların da orijinal çalışma ile aynı lisans altında yayınlanması gerekir. CC modeli, türev çalışmaların orijinal çalışmadan farklı bir lisans altında yayınlanması için esneklik sağlar. Örneğin, türev yok (ND) seçeneği araştırmacıların türev çalışmaları kısıtlamasına ve ticari değil (NC) seçeneği lisans sahiplerinin verilerinin ticari kullanımını kısıtlamasına izin verir. İdeal olarak ND'nin çok kısıtlayıcı bir seçenek olmasından kaçınılmalıdır çünkü açık erişim ve paylaşımın ruhuna ve ilkelerine aykırıdır (Patel, 2016, s. 235).

**Tablo 4.** Lisanslar ve koşulları (Tavluoğlu, 2020)

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- ND		Orijinali değiştiremez veya ticari olarak kullanılamaz ve atıf verilmeli

#### 2.4.5. Yeniden Kullanım

Veri depoları, araştırmacıların orijinal araştırmalardan elde ettikleri veri kümelerini depoladıkları (ve keşfettikleri) yerlerdir. Belirli disiplinler ve belirli veri türleri etrafında tasarlanmış veri havuzları vardır. İki yaygın veri deposu türü kurumsal depolar ve disiplinler depolardır. Araştırmacıların verilerini depolayabileceği ve veri taraması yapabileceği en çok kullanılan genel veri depolarına, Figshare, ICPSR, Açık Erişim Rehberi, Açık Bilgi Uluslararası Projeleri örnek gösterilebilir. Sağlık alanında kullanılan veri depoları olarak ise Patients Like Me, Science: Data Deposition, NIH: Data Sharing Repositories, Nature: Recommended Data Repositories, Biosharing, DataONE, Dryad sıralanabilir (Colorado State University, 2022). Dijital kütüphaneler genel sistemler olma

eğiliminde olmayıp, belirli topluluklar ve içerik türleri için tasarlandıklarında en etkili halde olurlar (Borgman ve diğerleri, 2015b, s. 208).

İlk veri depolarının 1960'ların sonunda kurulmasından sonra verilerin kataloglanması için bibliyografik standartların geliştirilmesi on yılı bulmuştur. 1970 yılında Amerikan Kütüphane Derneği (ALA), "Makinede Okunabilir Veri Dosyalarının Kataloglanması için Kurallar (Rules for Cataloging Machine Readable Data Files)" konulu bir alt komite kurarak diğer şeylerin yanı sıra makinede okunabilir veri dosyalarına bibliyografik kontrol getirmekle görevlendirdi. Akademik atıf uygulamalarının arşivleme uygulamalarını yakalamaya başlaması için uzun bir süre geçmesi gerekmiştir (Altman ve Crosas, 2013, 63). Yayınlanan verilerin yeniden kullanımında veri sahibine atıfta bulunulması, hem veri sahibi için bir ödül olarak hem de araştırma bulgularının kaynağını belgelemek için önemlidir. Önceki çalışmalar, kamuya açık veri setlerine sahip makalelerin, mevcut verileri olmayan benzer çalışmalara göre daha fazla sayıda atıf aldığını ortaya koymuştur. 2014 yılında yayınlanan Veri Atıf İlkeleri Ortak Deklarasyonu (Joint Declaration of Data Citation Principles) veri alıntılama önerileri üzerine sekiz temel ilke sunmaktadır (FORCE11, 2014):

1. Önem: Veriler araştırmanın meşru, alıntılanabilir ürünleri olarak görülmelidir ve veri atıflarına diğer araştırma nesnelere yapılan atıflarla aynı önem verilmelidir.
2. Kredi ve Atıf: Veri atıfları, tek bir atıf tarzı veya mekanizmasının tüm veriler için geçerli olmayabileceğini kabul ederek, verilere katkıda bulunan herkese bilimsel kredi ve yasal atıf verilmesini kolaylaştırmalıdır.
3. Kanıt: Bilimsel literatürde, bir iddianın verilere dayandığı her yerde ve her zaman, ilgili verilere atıfta bulunulmalıdır.
4. Benzersiz Tanımlama: Bir veri atıfı, makine tarafından uygulanabilir, küresel olarak benzersiz ve bir topluluk tarafından yaygın olarak kullanılan kalıcı bir tanımlama yöntemi içermelidir.
5. Erişim: Veri atıfları hem insanların hem de makinelerin atıfta bulunulan verileri bilinçli bir şekilde kullanması için gerekli olan verinin kendisine ve ilişkili üst verilere, belgelere, kodlara ve diğer materyallere erişimi kolaylaştırmalıdır.

6. Kalıcılık: Benzersiz tanımlayıcılar, veriler ve üst veriler, tanımladıkları verilerin ömrünün ötesinde bile kalıcı olmalıdır.
7. Belirlilik ve Doğrulanabilirlik: Veri atıfları, bir iddiayı destekleyen belirli verilerin tanımlanmasını, bunlara erişimi ve bunların doğrulanmasını kolaylaştırmalıdır.
8. Birlikte çalışabilirlik ve esneklik: Veri atıf yöntemleri, topluluklar arasındaki değişken uygulamalara uyum sağlamak için yeterince esnek olmalı, ancak topluluklar arasında veri atıf uygulamalarının birlikte çalışabilirliğini tehlikeye atacak kadar farklılık göstermemelidir.

### 3. BÖLÜM

## SAĞLIK BİLİMLERİNDE VERİ, YENİDEN KULLANIMI VE VERİ OKURYAZARLIĞI

### 3.1. SAĞLIK VERİSİ

1960'lardan itibaren hızla gelişen bilgisayar teknolojisiyle temeli atılan elektronik sağlık kayıtlarının kullanımı, yalnızca hastaların tıbbi bilgilerinin kaydedilip okunmasını ve dünyanın hemen her yerinden erişilebilir olmasını sağlamakla kalmayıp, aynı zamanda sağlık kayıtlarının biçimini ve dolayısıyla sağlık hizmetini değiştirmiştir. Kayıtların epidemiyolojik araştırmalar için kullanılmasıyla, düşük kalitede verilerin hasta zararına yol açabileceği keşfedilmiştir (Evans, 2016, s. 48-50). Sağlık verileri, bir kişinin sağlık durumuyla ilgili kişisel bilgileri ifade eder. Bu hem tıbbi verileri (reçeteler, tıbbi muayene raporları, laboratuvar testleri, tıbbi görüntülemeler, vb.) hem de sağlıkla ilgili idari ve mali bilgileri (randevuların planlanması, sağlık hizmetleri için faturalar ve hastalık izni yönetimi için tıbbi raporlar, vb.) kapsar. Sağlık sektörü, birçok farklı kaynaktan yoğun veri üretiminin ve akışının olduğu bir sektördür. Sağlık verileri benzersiz ve tekildir. İnsanların ve sağlık geçmişlerinin birbirlerinden farklı oluşu nedeniyle hiçbir hasta-veri kombinasyonu aynı değildir. Veriler eklenebilirler ve daha fazla veri ile daha güçlü istatistiksel analizler yapılabilir. Ayrıca sağlık verileri tükenmediği gibi hızlı bir şekilde kopyalanabilir, dağıtılabılır ve çoğaltılabilirler (Telenti ve Jiang, 2020, s. 1006).

Farklı biçimlerde ve birçok kaynaktan üretilen sağlık verileri: Hastalara ait demografik veriler, tıbbi geçmişleri, aktif tıbbi problemleri, aşular, alerjiler, ilaçlar, vital bulgular, laboratuvar ve radyoloji testlerinden elde edilen sonuçlar, patoloji raporları, sağlık hizmeti sunucuları tarafından hazırlanan ilerleme notları, sağlık talepleri ve fatura kayıtları gibi idari ve mali belgeler, elektronik tıbbi kayıtlar, hasta kayıtları, uzaktan algılayıcılar, giyilebilir cihazlar, akıllı sayaçlar ve sensörlerden gelen veriler, yaşamsal bulgu cihazlarından elde edilen veriler, kişilerin parmak izleri, retina taramaları, Facebook, Twitter ve diğer sosyal medya platformları ile kullanıcıların konumlarını,

sağlık davranışlarını, duygularını ve sosyal etkileşimlerini içeren veriler, bireylerin günlük hayatlarındaki egzersiz düzeylerini gösteren adım sayıları, kilo değişimleri, beslenme şekilleri, tükettikleri gıdalar, yararlı/zararlı alışkanlıkları vs. gibi yapılandırılmış ve yapılandırılmamış verilerdir (Altındış, 2019). Sağlık veri kaynakları Priyanka ve arkadaşları tarafından şöyle gruplandırılmaktadır (2014, s. 5866):

- Klinik veriler
- Yayınlar
- Klinik referanslar
- Genomik veriler
- Akışlı veriler (Streamed data)
- Web ve sosyal ağ verileri
- İşletme, organizasyonel ve harici veriler.

Ehrenstein ve arkadaşlarına (2017, 245, s. 5) göre ise veri kaynakları;

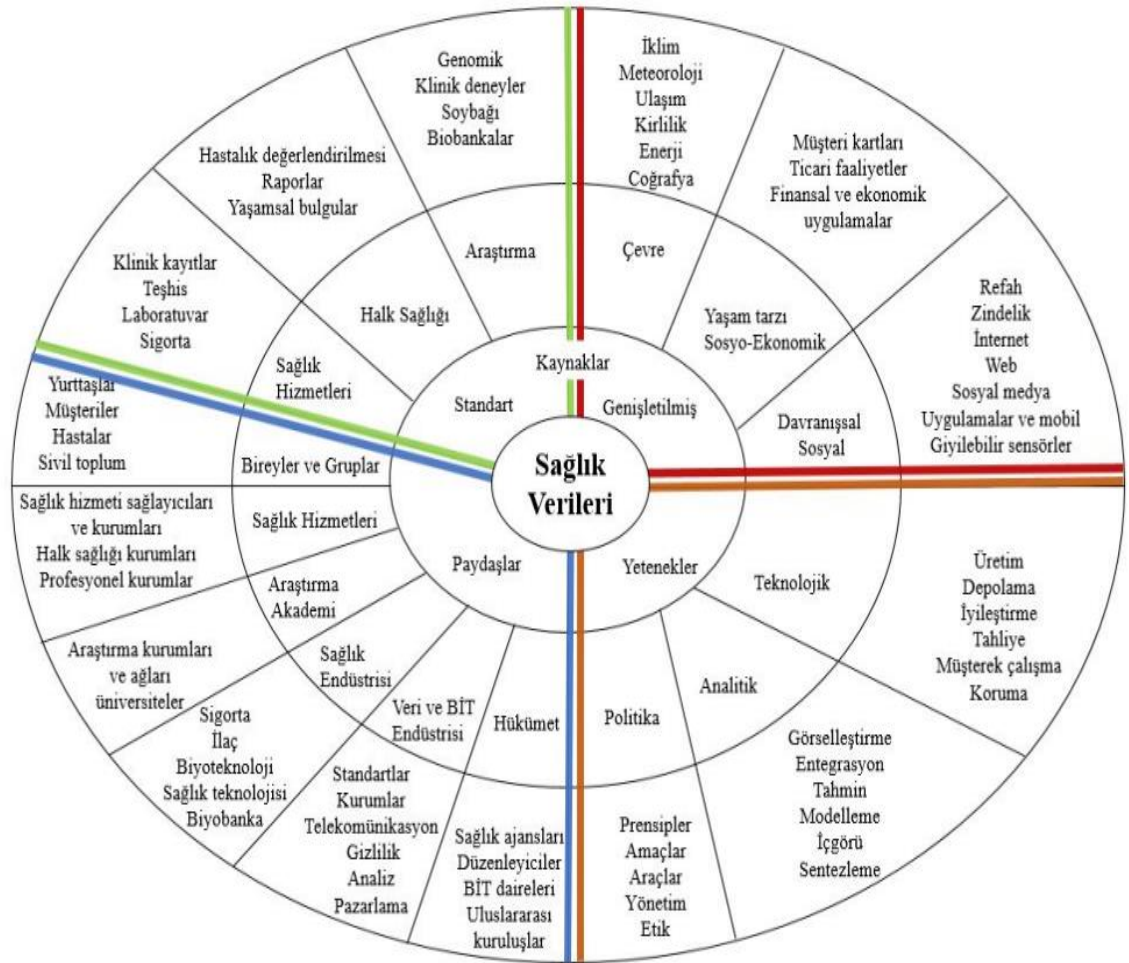
- Makine tarafından oluşturulan veriler,
- Biyometrik veriler,
- İnsan tarafından üretilen veriler,
- İşlem verileri,
- Davranış verileri,
- Epidemiyolojik veriler,
- Yayınlanmış veriler,
- Günlük yaşamdan sağlıkla ilişkilendirilebilecek veriler, şeklinde gruplandırılmaktadır.

Altındış ve Morkoç (2018) çok çeşitli kaynaklardan ve çok miktarda üretilen sağlık verilerinin kullanım alanlarını şöyle listelemişlerdir;

- Hasta bakımında ve araştırmada
- Hastalıkların sürveyansı ve halk sağlığı yönetiminde
- Tıp eğitimi ve öğretiminde
- Kurumsal öğrenmede
- Araştırma ve geliştirmede
- Kişiselleştirilmiş tıpta

- Yönetimde
- Genomik ve biyoinformatikte

Sağlık verisi ekosistemi çok sayıda paydaş ve sistemden meydana gelmektedir. Yeni analitik modeller, veri kaynakları ve paydaşlar giderek daha dinamik ilişkiler kurdukça, sağlık verisi ekosisteminin gelişen bir yapı olduğunu söyleyebiliriz. Sağlık verileri için veri ekosistemi Şekil 3'te gösterilmiştir (Vayena ve diğerleri., 2018, 6 s. 7; akt. İnce ve diğerleri, 2018, s. 850).



Şekil 3. Sağlık hizmetlerinde veri ekosistemi

Klinik araştırmalarda veriler üç yaklaşım kullanılarak toplanır. Klasik yaklaşımda veriler, araştırma protokolü tarafından ele alınan soruya göre bir form ile manuel ve ileriye dönük olarak toplanır. İkinci yaklaşım olan klinik kayıta ise veriler ileriye dönük olarak



toplanır, ancak önceden tanımlanmış bir araştırma sorusu yoktur. Verilerin yeniden kullanımı olarak tanımlanan son yaklaşım ise, elektronik sağlık kayıtlarının giderek daha fazla uygulanmasıyla ortaya çıkmıştır ve otomatik olarak kaydedilen verilerin, veri kaydı sonrasında tanımlanan bir araştırma sorusunu ele almak üzere yeniden kullanılmasını içermektedir (Lamer ve diğerleri, 2022).

Geçtiğimiz yüzyılda sağlık veri yönetiminin tanımı, teknolojik ilerlemelerin yanı sıra daha iyi ve gelişmiş hasta bakımı ihtiyacını karşılamak için çok sayıda dönüşüm geçirmiştir. Kişisel ve demografik kimlik, tıbbi durumun geçmişi, devam eden tedavi, laboratuvar testleri ve radyoloji sonuçlarına dayalı olarak bir hastanın kimliğini ve sağlığını tanımlayan tıbbi veriler, bir sağlık veri yönetim sisteminin ortak gereksinimidir. Tıbbi kayıtlar ister basılı belgeler ister dijital kayıtlar şeklinde olsun, sağlık veri yönetim sistemlerinin evrimi boyunca birincil bileşen olmuştur (Ismail ve diğerleri, 2020). Klinik verilerin erişime açılması, heterojen ve çok boyutlu klinik veri kümelerinin, veri ambarlarında veya ağlarda birleştirilmesi, entegrasyonu ve paylaşılması anlamına gelir (Meystrea ve diğerleri, 2017, s. 40). Klinik bilgi sistemleri geçici sorguları destekleyecek şekilde yapılandırılmamıştır, daha ziyade tek bir hastanın bilgilerinin alınmasını destekler.

Günümüzde ise bilişim teknolojilerindeki gelişmelerle birlikte büyük miktarda verinin analiziyle anlamlı bilgiler elde etmenin mümkün hale gelmesi ve kapsamlı veri tabanları sayesinde araştırmacıların hastalıkların altında yatan sebepleri daha kolay ve hızlı keşfetmesi mümkün olmuştur. Araştırmacıların güncel gelişmeleri takip edebilmesi, hastalıkların altında yatan sebepleri keşfetmesi, tanı ve teşhis süreçlerini yönetmeleri daha hızlı ve kolay olmuştur. Araştırmacılar yüksek kalitede klinik verileri daha hızlı ve daha verimli bir şekilde üretebilmekte ve veri madenciliği yöntemleriyle analiz edebilmektedir. Sağlık bilimlerinde veri madenciliği kullanım alanları şöyle sıralanabilir (Bains, 2016);

- Büyük boyutlarda verinin analizinin mümkün olması
- Kanıta dayalı tıp ve hastane hatalarının önlenmesi
- Halk sağlığında politika oluşturmaya yardımcı olması
- Veri toplama maliyetlerinden tasarruf sağlaması
- Hastalıkların erken teşhisi ve/veya önlenmesi

- Pandemik hastalıkların erken tespiti ve yönetimi için politika oluşturulması
- Non-invaziv teşhis ve karar desteği
- Olumsuz ilaç olayları

Bilgi teknolojilerindeki eğilim büyük miktarda veriden, veri madenciliği yöntemleri uygulayarak anlamlı sonuçlar elde etmektir. Bu bilgiler teşhis ve tedavi çalışmalarına, hekimlerin doğru karar vermelerine ve sağlık hizmetlerinin kalitesini arttırmaya katkıda bulunmaktadır. Literatürde sağlık verilerinin veri madenciliği yöntemleri ile analiz edilerek ilaçların bilinmeyen yeni yan etkilerinin keşfedildiği (Ayvaz ve diğerleri, 2016; Yıldırım, 2015), tümörlerin araştırıldığı ve türlerinin belirlendiği (Hughes ve Morin, 2018; Nath, Pal, Mukhopadhyay ve Mondal, 2020), %90 üzeri başarıyla hastalıkların teşhis edilebildiği (Bardou, Zhang ve Ahmad, 2018; Kayalp, Dogan, Ceylan, Senes ve Yucel, 2013; Toğaçar, Ergen ve Sertkaya, 2019), tıbbi konularda yeni bilgilerin keşfedildiği (Baydilli, 2020; Yarkoni, Poldrack ve Nichols, 2011) çok sayıda çalışma vardır. Ancak Canlas'a (2009, s. 8) göre, tıp alanında veri madenciliği uygulamak, tıp mesleğinin kendine has özellikleri nedeniyle çok zor bir iştir. Geleneksel veri madenciliği veri kümelerindeki kalıplar ve eğilimlerle ilgilenirken, tıpta veri madenciliği daha çok kalıplara ve eğilimlere uymayan azınlık ile ilgilenmektedir. Yaklaşımındaki bu farklılığı artıran şey, çoğu standart veri madenciliğinin çoğunlukla bu kalıpları ve eğilimleri açıklamakla değil, tanımlamakla ilgilenmesidir. Ancak, tıpta bu açıklamalara ihtiyaç vardır çünkü küçük bir fark, yaşam ve ölüm arasındaki dengeyi değiştirebilmektedir. Krzysztof ve Moore (2002) tıbbi verilerin benzersiz olduğu ana noktaları dört genel başlık altında toplamaktadırlar:

**Tıbbi verilerin heterojenliği:** Ham tıbbi veriler hacimli ve heterojendir. Tıbbi veriler, çeşitli görüntülerden, hastayla yapılan görüşmelerden, laboratuvar verilerinden ve hekimin gözlem ve yorumlarından toplanabilir. Tüm bu bileşenler hastanın tanı, prognoz ve tedavisini etkileyebilir ve göz ardı edilemez. Depolanan bilgiler, kolayca anlaşılır bir biçimde mevcut değilse daha az kullanışlı hale gelir. Görselleştirme teknikleri bu ortamda giderek artan bir rol oynayacaktır, çünkü görüntüler insanlar için en kolay kavranabilen tekniklerdir ve sonuçların tek bir anlık görüntüsünde çok fazla bilgi sağlayabilirler.

**Etik, yasal ve sosyal konular:** Tıbbi veriler insan denekler üzerinde toplandığından, hastaların zarar görmesini ve verilerinin kötüye kullanılmasını önlemek için tasarlanmış etik ve yasal gelenek vardır. Hastalara ait sağlık verilerinin kime ait olduğu konusu belirsizdir. Verilerinin gizliliği ve güvenliği ihlal edildiğinde sorumlunun kim olacağı da (doktor, sağlık kurumu veya veri madencisi vb.) belirsizdir. Tıpta etik, yasal ve sosyal konuların ana noktaları beş başlık altında toplanabilir: veri sahipliği, insan verilerinin gizliliği ve güvenliği, beklenen faydalar ile idari sorunlar.

**İstatistik felsefesi:** Veri madenciliği yöntemlerinin, özellikle istatistiklerin ve bu yöntemlerin altında yatan temel varsayımların tıbbi veriler için temelde farklı olabileceği düşünülmektedir. Araştırma sırasında araştırmacı bakış açısını değiştirirse verilerde değişiklik olmasa bile verilerin yorumunun değişeceğini düşünmektedirler.

**Tıbbın özel durumu:** Tıbbın gerek bilim ve felsefede gerekse de günlük yaşamda özel bir yeri vardır. Tıbbi müdahalelerin sonuçları insan hayatıyla ilgili bir zorunluluktur ve geri döndürülemez sonuçları olabilir. Hekimler tüm meslekler arasında en uzun çıraklık süresine sahiplerdir ve hastaların hekimlerden beklentisi oldukça yüksektir.

### 3.2. SAĞLIK VERİLERİNİN YENİDEN KULLANIMI

Verinin yeniden kullanımı, benzer ya da farklı bir uygulama için verilerin geri kazanılması ya da yeniden işlenmesi anlamına gelmektedir (Curty, 2015, s. 35). Faniel ve Jacobensen'e (2010) göre verinin yeniden kullanımını tanımlayan araştırmaların vardığı ortak nokta, verinin başlangıçta amaçlanandan farklı bir amaç için olan ikincil kullanımınıdır. Safran ve arkadaşları (2007, s. 1) sağlık verisinin ikincil kullanımını, kişisel sağlık verilerinin tedavi amacıyla doğrudan kullanım amacının dışında araştırma, kalite/güvenlik değerlendirmesi, halk sağlığı politikaları, ödeme, akreditasyon ve pazarlama gibi amaçlarla kullanımı olarak ifade eder. Heaton (2004) ise verinin ikincil kullanımının üç ana işlevine işaret etmektedir: verilerin yeniden kullanımı, veri replikasyonu ve sentez. İlk işlev olan verilerin yeniden kullanımı, verinin orijinal toplayıcıları tarafından ortaya koyulandan farklı, yeni veya ek araştırma sorularının araştırılmasını ifade eder. İkinci işlev olan veri kopyalama, orijinal çalışmayı doğrulamak, çürütmek veya iyileştirmek amacıyla kullanılan yeniden kullanım türünü

ifade eder. Üçüncü işlev olan sentez ise bilgiyi sentezlemek için kullanılacak birleştirme ve bütünleştirme gibi farklı teknikleri içerir.

Klinik verilerin yeniden kullanımı, sağlık hizmetlerinin kalitesinin yükseltilmesinde, sağlık yönetimini iyileştirmede, sağlık hizmetleri maliyetlerini azaltmada, halk sağlığı yönetiminde ve klinik araştırmalarda kullanılması için önemlidir. Tıptaki hızlı ilerleme ile tedavi yöntemleri, tıbbi müdahaleler, tanı ve test imkanları gelişirken, eski uygulamalar ile ilgili de hekimlerin bilgilerini sürekli güncellemesi gerekmektedir. Sağlıkta araştırma verilerinin paylaşımı ve yeniden kullanımı ile araştırmacılar yüksek kalitede klinik kanıtları daha hızlı ve daha verimli bir şekilde üretebilmektedir. Böylelikle klinik araştırmacılar ve hekimler daha fazla araştırmaya katılabilmektedir (Coorevits ve diğerleri, 2013, s. 548).

Temel ve klinik öncesi araştırmalar özellikle önemlidir çünkü gelecekteki çalışmaların üzerine inşa edileceği temeli oluştururlar. Sonunda klinik çalışmalara ve insanlığa fayda sağlayan yeni ilaçlara dönüşecek olan heyecan verici, yeni fikirleri sağlayan klinik öncesi araştırmalardır (Hackam ve Redelmeier, 2006). Veri paylaşımı, araştırma sonuçlarının insan sağlığını iyileştirmeye yönelik bilgi, ürün ve prosedürlere hızlı bir şekilde dönüştürülmesi için gereklidir. Tıbbi veriler, insan sağlığını iyileştirmeye yönelik araştırmalarda akademi ve tıp kurumları tarafından kullanılmaktadır. Ancak klinik veri havuzları genellikle gerçek sonuç verilerinden ziyade süreçle ilgili verilere sahiptir. Veriler toplandığında, verilerin nasıl ve neden toplandığına ilişkin bağlam aktarılamayabilir. Bu doğal önyargılar ve sınırlamalar nedeniyle, araştırmacılar rutin olarak toplanan klinik verilerin analizlerini yaparken dikkatli olmalıdır (Safran, 2014, s. 53). Verilerin yeniden kullanılabilmesi için "üretildikleri bağlamlardan ve ilişkilerden ayrı, serbestçe dolaşan metalar" olmaktan daha fazlası olarak görülmeleri gerekir (Mauthner, 2012, s. 3).

Yayımlanmış bir araştırmaya ait verilerin yeniden kullanımı ile ilgili ilk örneğe Steinmetz (1951) tarafından yapılan bir araştırmada rastlanmıştır. Steinmetz, iribaş kurbağaların TSH ve tiroid hormonlarının büyümeleri üzerindeki etkisinin araştırıldığı ve makale olarak yayımlanmış bir çalışmanın verilerini yeniden analiz ederek çalışmayı doğrulamıştır. Tıptaki hızlı ilerleme ile tedavi yöntemleri, tıbbi müdahaleler, tanı ve test imkanları gelişirken, eski uygulamalar ile ilgili de hekimlerin bilgilerini sürekli

güncellemesi gerekmektedir. Özellikle bilginin çok hızlı aktığı Covid-19 pandemisi döneminde hekimler gelişmeleri çok hızlı takip etmeleri gerektiğini ve hem hasta bakımı hem de literatür takibinin çok yorucu olduğunu medyada sıkça dile getirmişlerdir.

Baynam ve arkadaşlarının (2015, s. 382) yaptığı bir araştırmaya göre, genotip ve fenotip verilerinin iyi tanımlanmış şekilde araştırmacıların erişimine açılması, hastalıklara doğru gen tanımlaması yapılma oranını %36'dan %81'e çıkarmıştır. Covid 19 pandemisi başladığında Çin'deki sağlık çalışanlarının ve araştırmacıların, hangi verileri yayımlayabilecekleri konusunda hükümetin sınırlamalarına maruz kaldıkları öne sürüldü. Ancak Çinli araştırmacıların SARS-CoV-2 virüsünün ilk genom dizisini “virological.org” ve “GenBank” veri havuzlarında paylaştığı 11 Ocak 2020'den itibaren, veri hacmi hızla arttı. Mart ayının sonuna gelindiğinde Nextstrain isimli virolojik analiz web sitesi, dünyanın dört bir yanındaki laboratuvarlardan günde 50 ila 200 dizi alıyor ve bu dizilerdeki genetik varyasyonlar yoluyla virüslerin yayılmasını izleyerek, birkaç saatte bir virüsün evrimini analiz edebiliyordu (Zastrow, 2020). The Diamond Light Source-PostEra COVID Moonshot, The National Center for Advancing Translational Sciences (NCATS), The Institute of Cancer Research (ICR) – canSAR, Avrupa Komisyonu tarafından oluşturulan COVID-19, The Medicines for Malaria Venture (MMV) – COVID Box, The COVID-19 Protein Portal UK, pandemi ile mücadele için araştırmacıların verilerini paylaştığı portallar olmuşlardır (Tse, Klug ve Todd, 2020).

Ross, Gross ve Krumholz (2011) sağlık araştırmalarında güçlü, kanıta dayalı uygulama, bireysel klinik kararlar, sistematik incelemeler ve meta-analizler için objektif, tarafsız araştırmaların mevcut olması gerektiğini belirtirken, sağlık sektöründe araştırmaların kalitesini düşüren ve tıbbi literatürü çarpıtan uygulamaların varlığından bahsetmektedir. Bunlar 8 başlık altında aşağıdaki gibi açıklanmıştır;

Tohumlama deneyleri: Pazarlama amaçlarını hastalara, hekimlere, düzenleyicilere veya kurumsal inceleme kurulu üyelerine açıklamadan, bir klinik araştırmaya araştırmacı olarak katılma kisvesi altında doğrudan hekimlere ilacı veya cihazı tanıtmak ve kullanımını teşvik etmek amacıyla insan katılımcılar arasında yürütülen bir ilacın veya cihazın klinik araştırmaları.

Yayın planlaması: Pazar etkisi yüksek bulgular için yüksek profilli dergileri hedefleyerek ve daha düşük profilli dergilerde stratejik olarak ilgili çok sayıda pazar odaklı makale yayınlamak, klinik arařtırmalardan maksimum ticari deęeri elde etmek için ila Őirketlerinin verilerini Őekillendirmenin ve verileri tıbbi dergi makalelerine dnřtrmenin organizasyonel ve pratik alıřması.

Anahtar mesajlar: İla satıřlarını teřvik etmesi beklenen temel mesajların veya temaların belirlenmesi ve ardından yayınlarnın bu mesajlar ve temalar etrafında planlanması.

Hayalet yazarlık: Bir makalenin arařtırılmasına veya yazılmasına nemli katkıda bulunan bir kiřinin, bu durumda bir endstri alıřanının veya harici bir tıbbi yazarın, yazar olarak belirtilmemesi.

Misafir yazarlık: Yazarlık kriterlerini karřılamayan ve endstri tarafından istihdam edilmeyen akademik bir arařtırmacının dıř tarafsızlık saęlamak amacıyla yazar olarak belirlenmesi.

Seici yayın: Bir ilacı veya cihazı desteklemeyen veya rnn ticari deęerini azaltabilecek bulgulara sahip klinik alıřmaların gecikmeli olarak yayınlaması veya yayınlanmaması.

Seici raporlama: Bir ilacı veya cihazı desteklemeyen veya rnn ticari deęerini dřrebilecek klinik arařtırma bulgularının kısmi veya eksik raporlanması.

Belirsiz raporlama: Bir ilacı veya cihazı desteklemeyen veya rnn ticari deęerini dřrebilecek klinik arařtırma bulgularının yanıltıcı veya kamuoyunun dikkatini ekme olasılıęı daha dřk olacak Őekilde raporlanması.

Piwowar, Day ve Fridsma (2007) arařtırmaya ait verilerin arařtırmayla birlikte paylařılması ile yayınlarnın atıf oranları arasında bir iliřki olup olmadıęını arařtırdıkları alıřmada, kanser ile ilgili klinik alıřmalar zerinden, verilerin paylařıldıęı arařtırmaların paylařılmayanlara gre %69 daha fazla atıf aldıklarını ortaya koymuřlardır. Ancak psikoloji alanında yapılan bir alıřmada arařtırmacıların verilerini paylařmaya istekli olup olmama durumu ile makalelerindeki istatistiksel sonuların raporlanma kalitesi arasında bir iliřkili olup olmadıęını incelemek amacıyla 49 yazardan

verilerini paylaşmaları istenmiş ve araştırmacılardan sadece 21'i (%42,9) verilerini paylaşmıştır. Belirtilen 49 makalede sunulan 1148 istatistik test tekrarlandığında verilerini paylaşan yazarların hesaplamalarında %4,3 oranında hata ile karşılaşılırken, yanlış bildirilen p değerlerinin çoğunun (%73,5) verilerini paylaşmayan yazarların makalelerinden kaynaklandığı belirtilmiştir (Wicherts, Bakker ve Molenaar, 2011). Kanser biyolojisinde klinik öncesi araştırmaların tekrarlanabilirliğini araştırmak için yapılan bir çalışmada ise 53 yüksek etkili makaleden 193 deneyi tekrarlamak isteyen araştırmacılar verilerin önemli bir kısmına ulaşamadıklarını ve 23 makaleden 50 deneyi tekrarlayabildiklerini belirtmişlerdir. Yazarlardan orijinal verileri talep etmiş ancak, deneylerin %16'sı için ham veri, %14'ü için özet veri elde ederken ve %68'i için hiçbir şey elde edemediklerini belirtmişlerdir (Errington ve diğerleri, 2021).

Tarihte "Cantekin Olayı" olarak bilinen bir başka olayda Bluestone ve Mandel'in amoksisilin tedavisinin etkinliği üzerine bir ilaç firmasından 3,5 milyar dolar destek alarak yaptıkları bir çalışma için yanlış sonuç ortaya koyduklarını iddia eden Prof. Dr. Erdem Cantekin ve çalışma grubu, araştırmanın bulgularını yorumlayarak amoksisilin tedavisinin etkisizliğini kanıtlamışlardır. Cantekin ve çalışma grubunun çalışması bilim camiası tarafından da kabul görmüştür (Glass, 1993). Klinik araştırmalarda temel zorluk, araştırmalara ayrılan büyük miktarlardaki fonların en verimli ve etkili şekilde kullanılmasını sağlamaktır. Klinik öncesi araştırmaların çoğunun, yazarların kendileri de dahil olmak üzere yeniden üretilmediği giderek daha açık hale gelmiştir (Begley ve Ioannidis, 2015, s. 117). Klinik öncesi araştırmaların yeniden üretilmezliğinin yaklaşık üçte biri biyolojik reaktifler ve referans materyallerden kaynaklanmaktadır (Nelson ve diğerleri, 2021). Meslekte yükselme için yayınların bir seçim kriteri olarak kullanılması klinik uzmanları sürekli araştırma yapmaya teşvik etmektedir ancak kalitesiz tıbbi araştırmalar, sağlık çalışanlarını ve politika yapıcıları yanlış yönlendirerek, bilime ve tıbbı olan güveni azaltarak ve kamu fonlarını israf ederek ciddi zararlara yol açmaktadır.

Goodman ve arkadaşları (2016) tarafından "yeniden üretilbilirlik" için kavramsal bir çerçeve ortaya konmuştur ve bu çerçeve üç temel kavram arasında ayırım yapmaktadır:

- Yöntemlerin yeniden üretilbilirliği, bir çalışmanın prosedürlerinin bir uzmanın aslına sadık bir şekilde tekrarlayabileceği kadar ayrıntılı bir şekilde tanımlanması,

- Genellikle "replikasyon" olarak adlandırılan ve çalışmanın teknik açıdan yetkin bir şekilde tekrarlanmasıyla ortaya çıkan sonuçların tekrarlanabilirliği,
- Çıkarımsal tekrarlanabilirlik, bir çalışmanın yeniden analizinin veya tekrarının tam olarak aynı sonuçlara olmasa da niteliksel olarak benzer sonuçlara varıp varmayacağıdır.

Tüm klinik öncesi araştırmalarda mükemmel tekrarlanabilirlik mümkün olan bir durum değildir. Tam bir yeniden üretilebilirlik elde etmeye çalışmak, bu tür çalışmaların maliyetini önemli ölçüde artıracak ve hacimlerini ciddi bir şekilde azaltacaktır (Freedman, Cockburn ve Simcoe, 2015). Sağlık verisindeki artış ve tıbbi verilerin araştırma için kullanılmasının artmasıyla birlikte, veri analitiğinin sağlık veri yönetim sistemlerine entegre edilmesi gerekli hale gelmiştir. Ancak bu durum, artan sayıda veri ihlali ve bilgisayar korsanlığı olaylarıyla başa çıkma, veri güvenliği ve gizlilik geliştirmek üzere birden fazla kaynaktan veri toplama ve ön işleme gibi yeni zorlukları da beraberinde getirmektedir (Ismail ve diğerleri, 2020). Sağlık verileri birçok farklı kaynaktan gelir, birçok farklı kişi bu verileri üretir, düzenler ve kullanır. Sistemler ve çalışmalar aynı nesnelere veya kavramları tanımlamak için farklı kelimeler kullandığında, aynı kelimeleri farklı anlamlara gelecek şekilde kullandığında veya farklı veri formatları veya yapıları kullandığında entegrasyon engellenir (Rubinstein ve diğerleri, 2020, s. 475).

### 3.2.1. Mahremiyet

Araştırma verilerinin depolanması ve iletilmesi için giderek daha karmaşık düzenlemeler gerekmektedir. Araştırma öznelerinin mahremiyetini korumak ve hassas, kişisel olarak tanımlanabilir bilgileri güvence altına almak için sağlam veri gizliliği ve güvenlik planlaması gereklidir (Harvard, 2022). Kayıtların gizliliği ve hasta bilgilerinin etik kullanımı gerekliliği sağlık hizmetlerinde veri madenciliği için büyük bir engeldir. Veri madenciliğinin daha doğru sonuç vermesi için oldukça büyük miktarda gerçek veriye ihtiyacı vardır. Sağlık verileri kişisel hassas verilerdir ancak yine de bu verilerin kullanılması tıbbın ilerlemesinde önemli yere sahiptir. Verinin açıklık düzeyinin artması ve paylaşımının yanı sıra verilerin elde edildiği kişilerin mahremiyetine her zamankinden daha fazla dikkat edilmesini gerekir (Cios ve Moore, 2002). Sağlık verileri; sağlık hizmeti verilmesi, bilimsel araştırma ve istatistiksel değerlendirmeler yapılabilmesi ve hukuki



anlaşmazlıklarda kanıt işlevi taşıma gibi bir çok sebeple işlenmektedir. Dijital sağlık pek çok anlamlı uygulamaya sahip olsa da hastaların kişisel sağlığına ilişkin hassas bilgilere olan yüksek bağımlılığı kaçınılmaz olarak veri güvenliği sorunlarını tetiklemektedir. Çoğu zaman hastaların haberi olmadan çok büyük miktarlarda kişisel sağlık verisi toplanmaktadır (Ross, 2016).

Bir sağlık veri yönetim sisteminde kişisel verilerin gizlilik gereksinimi, hastanın mahremiyetinin korunması açısından çok önemlidir. Özellikle hastaların tıbbi verilerinin analiz için kullanıldığı veri analitiğinin yükselişiyle birlikte hastanın mahremiyeti daha da önem kazanmıştır (Ismail ve diğerleri, 2020).

2018'de yürürlüğe giren Avrupa Birliği Genel Veri Koruma Yönetmeliği (GDPR), Avrupa Birliği içinde veya dışında Avrupa Birliği vatandaşlarının kişisel bilgilerini toplayan ve işleyen tüm kuruluşlar için geçerlidir. GDPR, her türlü kişisel veri işleme için geçerli olan genel kuralların yanı sıra bilimsel araştırma veya klinik yazılım geliştirme bağlamında yer alan sağlık verileri gibi özel kategorilerdeki kişisel verilerin işlenmesi için geçerli olan özel kuralları da belirlemektedir.

General Data Protection Regulation (GDPR), Avrupa'da önceki veri koruma yasalarının yeniden değerlendirilmesine ek olarak dijital sağlık çağında veri korumanın yeni zorluklarıyla başa çıkmayı amaçlamaktadır (Yuan ve Li, 2019). GDPR'de, daha önce "sağlıkla ilgili veriler" olarak geçen tanım detaylandırılarak, gerçek bir kişinin fiziksel veya ruhsal sağlığıyla ilgili geçmiş, mevcut veya gelecekte, sağlık hizmetlerine kaydı veya bu hizmetlerin sağlanması sırasında toplanan tüm veriler şeklinde tanımlanmıştır (Malgieri ve Comandé, 2017).

Ülkemizde 2015 yılında "Sağlık Bilgi Sistemleri Uygulamaları Hakkında 2015/17 Sayılı Genelge" yayımlanarak Sağlık.Net Online Sistemi kurulmuştur (Sağlık Bakanlığı, 2017). Sistemin amacı ülkedeki tüm sağlık kuruluşlarında verilen sağlık hizmeti sonucu üretilen tüm hasta verilerinin sisteme gönderilerek, Sağlık Bakanlığı bünyesinde tek bir yerde saklanmasıdır. Kişisel sağlık verilerinin korunması amacıyla yönetmelik ve yönergeler yayınlanmış ancak bunlar üzerine yapılan tartışmalar henüz son bulmamıştır. 2019'da yayınlanan en güncel yönetmeliğin "Bilimsel Amaçlarla İşleme ve Açık Sağlık Verisi" başlıklı beşinci bölümüne göre; anonim hale getirilen kişisel sağlık verilerinin, ilgili

kişilerin özel hayatın gizliliğini veya kişilik haklarını ihlal etmemek ve suç teşkil etmemek kaydıyla alınacak tedbirler doğrultusunda bilimsel amaçlarla işlenebileceği belirtilmektedir. Özellikle bilimsel araştırma alanında kişisel verilerin korunumu ile kast edilen, verinin korunumunun ötesinde o verinin kullanımında oluşabilecek kişilik hakkı ihlallerinin engellenmesidir.

### 3.2.2. Anonimleştirme

Hastalar, hastanenin güvenilir bir kurum olduğunu düşündükleri için ve bir sağlık hizmeti alabilmek için verilerini paylaşmaları zorunlu olduğundan verilerinin hastane tarafından saklanması sakınca görmezler. Pek çok kurum ve kuruluş, isim, sosyal güvenlik numarası, adres ve telefon numarası gibi açık tanımlayıcı bilgilerin kaldırılması halinde açıklanan verilerin mahremiyetinin korunmuş olduğunu düşünmektedir. Ancak, önemli araştırmalar sadece açık tanımlayıcı bilgilerin kaldırılmasının mahremiyetin korunması için yeterli olmadığını göstermiştir. Sağlık verilerinin üretildiği, depolandığı, yönetildiği ve kullanıldığı tüm bilgi sistemlerinin gelişen kalite ve güvenlik standartlarını karşılamak için sürekli olarak gelişmelere uyum sağlaması gerekir (Vayena ve diğerleri, 2018, s. 68). Etik kaygılar nedeniyle, büyük miktarda kullanılabilir sağlık verisi hastanelerin, kliniklerin kurumsal sınırları içinde ve hastaların cihazlarında hapsedilmiş durumdadır (Beyan ve diğerleri, 2020, s. 98). Veri anonimleştirme, kullanılan yöntem ve tekniklerinin etkinliğinden, anonimleştirme nedeniyle toplanan verilerin doğru analizini etkilemesine kadar birçok endişe ve soruyu gündeme getiren tartışmalı bir konudur. Bununla birlikte, toplanan veriler bireylerle ilgili olduğundan kişisel hassas verilerdir. Çalışmaya dahil olan katılımcıların istemesi halinde bu tür verilerin gizliliğinin korunması gerekli hale gelebilir. Bu amaçla, verilerin anlamını kaybetmeden anonimleştirilmesi için yöntemler türetilmelidir (Patel, 2016, s. 232). Dalenius'a (1977) göre anonimleştirme yöntemleriyle tanınmaz hale getirilen veri setine ait bilgileri çeşitli yöntemlerle edinmiş olsa bile veri sahipleriyle ilgili bilgi edinmemelidir. Samarati'ye (2001) göre verinin anonimleştirilebilmesi için öncelikle özniteliklerin belirlenmesi gerektiğini ve bu özniteliklerin dört gruba ayrıldığını varsaymaktadır;

*Açık tanımlayıcılar (Explicit Identifier):* Veri sahibinin kimlik numarası, isim, soy isim ve telefon numarası gibi doğrudan tanımlayan öznitelikleridir.

*Yarı tanımlayıcılar (Quasi Identifier):* Tek başına kayıt sahibinin kimliğini belirlemede yeterli olmamakla birlikte, diğer verilerle birleştirildiğinde bir kişiyi tanımlama potansiyeline sahip yaş, adres, cinsiyet vb. öznitelikler kümesidir.

*Hassas öznitelikler (Sensitive attributes):* Kayıt sahibine ait tıbbi bilgi, geliri, engellilik gibi kişisel hassas bilgilerdir.

*Hassas olmayan öznitelikler (Non-Sensitive attributes):* Açıklanan üç öznitelik grubu dışında kalan ve başka kişilerce bilinmesi tehlikeye sebep olmayacak bilgilerdir.

Veri işleme ve özellikle de veri aktarımında veri güvenliği konusu oldukça önemlidir. Tanımlayıcılar gizlenmeden önce verilere sadece yetkili kişiler erişebilmelidir. Hasta verilerinin tanımlanmasının dört şekli vardır (Cios ve Moore (2002, s. 9):

Anonim veriler, bilgilerin toplandığı sırada hasta kimliği kaldırılacak şekilde toplanan verilerdir. Örneğin, histoloji laboratuvarında kontrol doku bloğu olarak kullanılmak üzere belirli bir hastalığı olan bir hastanın otopsisinden bir doku bloğu alınabilir. Hastanın tanımlayıcıları numunenin toplandığı sırada kaydedilmez ve asla kurtarılamaz.

Anonimleştirilmiş veriler, başlangıçta hasta tanımlayıcıları ile toplanan ve daha sonra geri alınamaz bir şekilde kaldırılan verilerdir. Yani, hastanın kaydına geri dönme ve ek bilgi edinme olasılığı asla olamaz. Ancak, yukarıda açıklandığı gibi anonimleştirilmiş veriler yanlışlıkla çoğaltılabilir ve düzeltmeler veya ek veriler için doğrulanamaz.

Kimiksizleştirilmiş veriler, başlangıçta hasta tanımlayıcıları ile toplanan ve daha sonra kodlanan veya şifrelenen verilerdir. Hasta, yetkili bir kurum tarafından uygun görülen koşullar altında yeniden tanımlanabilir.

Tanımlanmış veriler yalnızca kurumun, federal kılavuzların vb. önemli incelemeleri altında hastanın yazılı bilgilendirilmiş onayı ile toplanabilir.

Veri kümelerinde gizliliği sağlamak için çok sayıda gizlilik modeli ve anonimleştirme yaklaşımı ortaya konmuştur. "k-anonimlik" kayıt bağlantı saldırılarını, "l-çeşitlilik" ise öznitelik bağlantı saldırılarını ölçmek için yaygın olarak benimsenen iki temel gizlilik modelidir. Çeşitli saldırı senaryoları için "t-yakınlık" ve "m-değişmezlik" gibi diğer gizlilik modelleri de önerilmiştir. Veri setlerini anonimleştirmek için bastırma,

genelleştirme, anatomizasyon, dilimleme ve ayrıştırma gibi çeşitli anonimleştirme işlemlerinden yararlanılmaktadır (Zhang ve diğerleri, 2014).

Ülkemizde Kişisel Verileri Koruma Kurumu tarafından 2017 yılında yayınlanan “Kişisel Verilerin Silinmesi, Yok Edilmesi veya Anonim Hale Getirilmesi Rehberi”, 6698 sayılı Kişisel Verilerin Korunması Kanununa uygun koşullarda anonim hale getirilmesi için gerekli yöntemleri, yapılması gerekenleri ve yasal çerçeveyi sunmaktadır. Veri anonimleştirme teknikleri tamamen kusursuz olmayabilir. Bu nedenle, hassas verilerin şifresinin çözülmesinin yasal yaptırımlarla ilişkilendirilmesi gerekmektedir. Hassas ve özel verilerin bu tür ihlallerini önlemek için cezalar katı olmalıdır.

### **3.3. VERİ OKURYAZARLIĞI**

#### **3.3.1. Tıp Eğitimi ve Veri Okuryazarlığı**

Tıp pratiği ve altında yatan bilgi tabanı o kadar hızlı değişmektedir ki, doktorların mesleki kariyerleri boyunca öğrenmeye devam etmeleri ve yaşam boyu öğrenen bireyler olmalarını sağlayacak becerileri geliştirmeleri gerekmektedir. Bununla birlikte, sürekli mesleki eğitim sadece güncel kalmaktan ibaret olmayıp, aynı zamanda yeni deneyimleri dahil etmek, mevcut durumları önceki deneyimlerle ilişkilendirmek ve bu sürece dayalı olarak mevcut deneyimleri yeniden düzenlemek için uygulamalar üzerinde düşünmeyi de gerektirir (O’Shea, 2003, s. 62). Covid 19 pandemisi döneminde tıp ve sağlık bilimcilerin sürekli öğrenme olanakları kısıtlanmış ve lisansüstü tıp eğitimi ile ilgili finansman sağlayan kurumlar tarafından fonlar sınırlandırılmıştır (Kaul ve diğerleri, 2021, s. 1950). Tıp eğitiminde önemli bir yeri olan uyarlanabilir öğrenme (adaptive learning) kavramının temelinde de, çevredeki yeni gelişmeleri öngörerek, tanıyarak ve bunlara hızla yanıt vererek elde edilen sürekli iyileştirme kavramı yer almaktadır. Veri setlerinin oluşturulması ve küresel bir kitleyle paylaşılması, uyarlanabilir öğrenenlerin eğitimi için birçok seçenek sunabilir ve uyarlanabilir müfredata entegre edilebilir. Öğrenen performansını izlemek ve öğrenme ihtiyaçlarını karşılayan vakalar önermek için uyarlanabilir öğrenme platformlarının iyileştirilmesi anlamlı klinik sonuçlar doğurabilir (Tackett, Wright ve Quirk, 2018).

Son yıllarda özellikle sosyal bilimler ve açık veri toplulukları tarafından geliştirilen yeni bir terim olan "veri okuryazarlığı" ile ilgili birçok bakış açısından farklı tanımlar yapılmaktadır (Calzada Prado ve Marzal, 2013 s. 125). Schield'e (2004) göre veri okuryazarlığı, istatistik okuryazarlığının bir parçası olarak bireylerin verilere erişim, değerlendirme, manipüle etme, özetleme ve sunma konusunda eğitilmesini içeren kısmıdır. Ridsdale (2015) veri okuryazarlığını, verileri eleştirel bir şekilde toplama, yönetme, değerlendirme ve uygulama becerisi olarak tanımlamaktadır. Eleştirel bakış açısı özellikle tedavi kararlarının hassas verilere dayandığı tıbbi araştırma ve hasta bakımı alanında oldukça önemlidir (Hoffman ve diğerleri, 2022, s. 257). Koltay'a (2017, s. 10) göre araştırma verisi okuryazarlığı, bireylerin verilere erişimlerini, yorumlamalarını, eleştirel olarak değerlendirmelerini, yönetmelerini ve etik olarak kullanmalarını sağlayarak, verileri eyleme geçirilebilir bilgiye dönüştürmelerini sağlayan belirli bir beceri seti ve bilgi tabanıdır. Kurum genelinde gerçekleşen yüksek araştırma oranı ve araştırmaların önemli bir kısmının veri yönetimi planı gereklilikleri olan fon kuruluşları tarafından finanse edildiği durumlarda veri okuryazarlığı eğitimi daha da önem kazanmaktadır (Adamick ve diğerleri, 2012, s. 180). Araştırmacılara, veri bilimini çalışmalarına uygulama konusunda destek sağlamak, kütüphanecilerin araştırmacıların yalnızca mevcut veriyi kullanan değil, aynı zamanda veriyi üreten olarak da ihtiyaçlarını anlamasını ve bunlara yanıt vermesini gerektirir (Carlson, 2014, s. 64).

### 3.3.2. Veri Okuryazarlığı Eğitimi

Veri okuryazarlığının önemli bir özelliği verinin üretilmesi gibi görünse de sadece veri tüketicisinin değil, veri üreticisinin de bakış açısını barındırma gerekmektedir (Koltay, 2014). Veri okuryazarlığı becerilerini geliştirmek isteyen araştırmacılar için eğitim kaynakları az ve nereden başlanacağını belirlemek zordur. Veri ve bilgi işlem becerilerine yönelik eğitim, lisans ve lisansüstü programlarda hala büyük ölçüde yer almamaktadır (Teal ve diğerleri, 2015). Edinburgh Üniversitesi tarafından geliştirilen MANTRA ve DataONE ile Coursera online eğitim platformu araştırmacıların veri okuryazarlığı alabileceği eğitim platformlarıdır. Ülkemizde ise TÜBİTAK tarafından 26-27 Eylül 2017 tarihleri arasında "Araştırma Verileri Yönetimi" eğitimi ile Ufuk2020 kapsamında 19-20 Nisan 2021 tarihleri arasında "Açık Bilim ve Araştırma Verilerinin Yönetimi (RDM) Eğitimi" düzenlenmiştir. TÜBİTAK ayrıca Araştırma Verilerinin Yönetimi Eğitim

Portali ile online eğitim imkanı sunmaktadır (Araştırma Verileri Yönetimi Eğitim Portalı, 2022).

Calzada Prado ve Marzal'a (2013) göre verilerin nasıl seçileceği ve sentezleneceği ve diğer bilgi kaynakları ve var olan verilerle birleştirilmesi önemlidir ve aşağıdaki beceriler gereklidir:

- Verilerin üretildiği ve yeniden kullanıldığı bağlamı tanımlamak (veri yaşam döngüsü),
- Kaynak veri değerini, türlerini ve formatlarını tanımak,
- Veriye ne zaman ihtiyaç duyulduğunu belirlemek,
- İhtiyaç duyulan bilgiye uygun veri kaynaklarına erişmek,
- Verileri ve kaynaklarını eleştirel olarak değerlendirmek,
- Uygun araştırma yöntemlerini belirlemek ve kullanmak,
- Verileri ele almak ve analiz etmek,
- Nicel bilgileri (belirli veriler, tablolar, grafikler, raporlar ve benzeri) sunmak,
- Sonuçları öğrenmeye, karar vermeye veya problem çözmeye uygulamak,
- Süreç boyunca planlama, organize etme ve öz değerlendirme yapmak.

Literatürde veri okuryazarlığı eğitimi üzerine yapılmış pek çok araştırma bulunmaktadır. Bu araştırmalar öğrencilerin eğitilmesi, kütüphanecilerin eğitilmesi ve araştırmacıların eğitilmesi üzerine çeşitlidir. Read ve arkadaşları (2019a) araştırmacılara veri okuryazarlığı eğitimi vermek üzere yedi modülden (Verilerin Hikayesi, Veri Yaşam Döngüsü, Araştırmacıları Anlamak, Araştırma Veri Yönetimi İklimi, Veri Dokümantasyonu En İyi Uygulamalar, Veri Standartları, Depolama, Muhafaza ve Paylaşım) oluşan bir online kurs içeriği hazırlamışlardır. Araştırmanın ilk kısmında yirmi yedi kütüphaneciye bu çevrimiçi eğitim verilerek modüllere yerleştirilen sorular aracılığıyla kendilerini değerlendirmeleri sağlanmıştır. Ön test-son test kullanılarak eğitim ile ilgili kendileri hakkında yaptıkları değerlendirmeler veri olarak toplanmıştır. Ayrıca, kütüphanecilerden üç tanesinin kendi kurumlarındaki araştırmacılara araştırma verileri yönetimi eğitimi vermek için kullanabilecekleri slaytlar, bir senaryo ve bir değerlendirme formundan oluşan bir "Öğretim Araç Seti" geliştirilmiştir. Araştırmaya katılan on araştırmacının eğitim ile ilgili düşünceleri ve öğrendikleri materyali eğitim sonrasında da kullanma niyetleri sorulmuştur. Derse katılan araştırmacıların veri yönetim

uygulamalarındaki gerçek deęişiklikler, dersten yedi ay sonra uygulanan bir takip anketi ile deęerlendirilmiřtir. Arařtırmacıardan beř kiřiye ulařılabilmiř ve eęitim sonucu öğrendikleri uygulamaları kullandıklarını belirtmiřlerdir.

Read (2019b) bu kez bir tıp kütüphanesinde dört farklı tarihte farklı gruplara olmak üzere bir buçuk saatlik çalıřtay (Veri Toplama Planlaması, Veri Toplama Aracı Tasarımı, Veri Standartlarının Kullanımı, Veri Kalitesi Bakımı, En İyi Veri Depolama, Aktarma ve Analiz Uygulamaları, Rol ve Sorumluluk Yönetimi) düzenlemiřtir. Arařtırma topluluęunun bu eęitim materyallerine güçlü bir ihtiyaç duyduęunu ortaya koymuřtur.

Hofmann ve arkadaşları (2022), geliřtirdikleri 28 saatlik seçmeli dersi (Tıbbın Dijitalleřmesi - Klinik Arařtırma ve Saęlık Hizmetleri İçin Veri Okuryazarlıęı), her biri iki saatlik dört çevrimiçi seminerin yanı sıra serbestçe planlanabilen haftada dört saatlik bireysel çalıřma süresi ile beř hafta boyunca yürütmüřlerdir. Kursun tasarımı 5 ařamalı HiGHmeducation Didaktik Çerçevesi'ni (Witte ve dięerleri, s. 202) takip etmiř ve öğrencilerin bir sorgulama topluluęu oluřturduęu yapılandırmacı bir yaklařım uygulamıřlardır. Her öğrenci, haftada iki ila üç göreve bölünmüř toplam 14 görev üzerinde çalıřmıřlardır. Ders sonunda öğrenciler bu eęitimi "önemli", "faydalı", "gerekli" ve "kesinlikle ilgili" olarak deęerlendirmiřtir. Tüm öğrencilerin veri okuryazarlıęı hakkında daha fazla řey öğrenmesi gerektięini ve kursun içerięinin müfredatın ayrılmaz bir parçası haline gelmesi gerektięini ifade etmiřlerdir. Ders sonunda öğrencilerde belirgin iki eğilim görölmüřtür. Bir yandan, yazılım araçlarını (örneğin R, SPSS) kullanarak veri analizi becerileri geliřtirmeye yönelik güçlü bir ilgi görölmürken, dięer yanda ise verilerin korunması, veri etięi ve bilgisayar yazılım ve donanım çözümlerinin günlük klinik hayata en iyi řekilde entegre edilmesine yönelik yaklařımlar gibi veri uygulamalarının yönleriyle daha fazla ilgilenmiřlerdir.

Ridsdale (2015) ve arkadaşları veri okuryazarlıęı eęitiminde en iyi uygulamaları dörde ayırmıřlardır. İlki olan uygulamalı öğrenme, öğrencilere teknik bir beceriyi tam olarak anlamaları için gerekli pratik deneyimi saęlarken öğrencilerin süreçleri ve yöntemleri kendi başlarına çözmeye ve kendi anlayıřlarını düzenleme řansı tanımaktadır. İkincisi olan modül bazlı öğrenme, öğrencilerin öğrenme çıktılarına ařamalı ve sistematik bir řekilde ulařmalarını saęlamaktadır. Üçüncüsü olan ardıřık veya yinelemeli öğrenme, öğrencilerin daha önce öğrendikleri beceriler üzerine yenilerini inřa etmelerini saęlar. Ezberleme veya

katı talimatları takip etme yerine süreci teşvik eder ve sonuçta alışılmadık bir kavramı öğrenmeyi daha yönetilebilir hale getirir. Dördüncüsü olan proje tabanlı öğrenme ise, ardışık öğrenme yaklaşımını uygulamak için yararlı bir yoldur. Geniş bir araştırma yelpazesi içeren ve gerçek dünyada uygulanabilirliği olan projeler, süreç/teori ve uygulama arasındaki bağlantıyı sağlamlaştıracaktır.

Strasser ve Hampton (2012) Amerika’da ekoloji alanında çalışan 48 enstitüden toplam 54 eğitmen üzerinde gerçekleştirdikleri araştırmada kurumun ve dersin özellikleri, veri ile ilgili konuların derslerine ne ölçüde dahil edildiği, bu konuları öğretmelerinin önündeki engeller ve veri yönetimi ile ilgili kendi kişisel inanç ve değerleri sormuşlardır. Veri yönetimi konularının lisans düzeyindeki ekoloji derslerinde çok çeşitli nedenlerle ele alınmadığını ve eğitmenlerin veri yönetiminin öğretilmesinin önündeki engeller olarak çoğunlukla zaman ve kaynak eksikliğini gösterdiklerini ortaya koymuşlardır. Tenopir ve arkadaşları (2016) çalışmalarında, bilim insanlarına veri yönetimi eğitim uygulamaları, veri yönetimini hangi seviyelerde öğrettikleri, hangi konuları ele aldıkları ve bu konuları öğretirken karşılaştıkları engeller hakkında anket uygulamıştır. Veri yönetimi eğitmenlerinin yarısından fazlası (%51,5) veri yönetimi konularının öğretilmesinin önündeki en büyük engellerden biri olarak zaman yetersizliğini gösterirken, %39’undan fazlası veri yönetimi eğitiminden sorumlu olmalarına rağmen söz konusu konuları öğretmek için gerekli uzmanlığa ve yeterli bilgiye sahip olmadıklarını belirtmişlerdir.



## 4. BÖLÜM

### BULGULAR

#### 4.1 ARAŞTIRMA VERİLERİNİN YÖNETİMİ FARKINDALIK ANKETİ

"Araştırma Verilerinin Paylaşımı Farkındalık Anketi" beş bölümden oluşmaktadır. İlk bölümde katılımcılara ana bilim dalı, yaşı, cinsiyeti, eğitim programı ile çalışanlar için unvanlarının sorulduğu altı soru bulunmaktadır. İkinci bölümde katılımcılara bilgi arama davranışlarını ortaya koymayı amaçlayan beş soru sorulmuştur. Üçüncü bölümde katılımcıların kullandıkları veri türleri ve formatlarını, araştırma verilerini nasıl yönettiklerini, araştırma verilerinin yönetimi uygulamalarından hangilerini kullandıklarını ortaya koymayı amaçlayan toplam on üç soru bulunmaktadır. Dördüncü bölümde katılımcıların araştırma verilerinin paylaşımına bakışı ve eğitim ihtiyaçlarını ortaya koymayı amaçlayan yedi soru bulunmaktadır. Anketin beşinci ve son bölümünde ise katılımcıların araştırma ile ilgili görüşlerinin sorulduğu açık uçlu bir soru bulunmaktadır.

Araştırmanın bu bölümünde çalışma kapsamında oluşturulan, çoktan seçmeli, 5'li likert tipi ve birden fazla seçeneğin işaretlenebildiği sorular bulunan anket ile elde edilen veriler SPSS ile analiz edilmiştir. Toplam, sıklık, ortalama, yüzde değerleri betimleyici tablolar ile gösterilmiştir. Ayrıca %95 güven düzeyinde ki-kare anlamlılık testleri yapılarak değişkenler arasında istatistiksel olarak anlamlı bir fark olup olmadığı ortaya koyulmuştur. Katılımcılara sorulan açık uçlu sorular ise ayrıca değerlendirilmiştir.

Ankete 39 ana bilim dalına kayıtlı toplam 140 öğrenci katılmıştır. Katılımcıların bölümleri fakültelere göre gruplandırıldığında toplam 6 fakülteye ayrılmaktadırlar. Dış Hekimliği Fakültesinden geri dönüş alınamamıştır. Fakülte ve bölümlere göre öğrenci dağılımı Tablo 5'te gösterilmiştir. En çok katılım 39 (%28) öğrenci ile Tıp Fakültesi'nden olmuştur. Fizik Tedavi ve Rehabilitasyon Fakültesi (FTR) ile Sağlık Bilimleri Fakültesi'nden eşit (S=33, %23,5) katılım olmuştur. Eczacılık Fakültesi'nden 15 (%10,7), Hemşirelik Fakültesi'nden ise 12 (%8,6) öğrenci araştırmaya katılmıştır. En az katılım 8 (%5,7) öğrenci ile Spor Bilimleri Fakültesi'nden olmuştur.

**Tablo 5.** Fakülte ve bölümlere göre öğrenci dağılımı

<b>Fakülte</b>	<b>Ana Bilim Dalı</b>	<b>%</b>	<b>S</b>
Eczacılık Fakültesi	Farmasötik Teknoloji		4
	Klinik Eczacılık		3
	Eczacılık Temel Bilimleri		3
	Eczacılık İşletmeciliği		2
	Radyofarmasi		1
	Farmasötik Toksikoloji		1
	Farmakoloji		1
	%10,7		15
Fizik Tedavi ve Rehabilitasyon Fakültesi	Fizyoterapi ve Rehabilitasyon		24
	Kalp ve Solunum Fizyoterapisi ve Rehabilitasyon		9
	%23,5		33
Hemşirelik Fakültesi	Halk Sağlığı Hemşireliği		3
	İç Hastalıkları Hemşireliği		2
	Hemşirelikte Yönetim		2
	Hemşirelik Esasları		2
	Doğum ve Kadın Hastalıkları Hemşireliği		2
	Cerrahi Hastalıkları Hemşireliği		1
	%8,6		12
Sağlık Bilimleri Fakültesi	Ergoterapi		8
	Çocuk Gelişimi		7
	Odyoloji		6
	Beslenme Bilimleri		5
	Diyetetik		4
	Toplu Beslenme Sistemleri		2
	Dil ve Konuşma Terapisi		1
	%23,5		33
Spor Bilimleri Fakültesi	Spor Bilimleri ve Teknolojisi		6
	Beden Eğitimi ve Spor		2
	%5,7		8
Tıp Fakültesi	Halk Sağlığı		10
	Tıp Eğitimi ve Bilişimi		4
	Psikiyatri Hemşireliği		4
	Temel Onkoloji		3
	Biyoistatistik		3
	Biyoinformatik		3

Sağlık Bilimlerinde Simülasyon	2
Pediyatrik Temel Bilimler	2
Dahili ve Cerrahi Araştırmalar	2
Tıbbi Biyoloji	1
Sosyal Pediyatri	1
Nörolojik Ve Psikiyatrik Temel Bilimler	1
Kök Hücre Bilimleri	1
Aşı Çalışmaları	1
Analitik Kimya	1
%28	39
Toplam	14
%100	0

Katılımcıların kayıtlı oldukları öğretim programı Yüksek Lisans için 70 (%50), Doktora için 70 (%50) kişi olmak üzere eşit dağılım göstermektedir (Tablo 6). Dönemlerine bakıldığında tablo 3'te görüldüğü üzere yüksek lisans ders döneminde 22 (%15,7), tez döneminde 48 (%34,3) öğrenci bulunmaktadır. Doktorada ise ders döneminde 28 (%20), tez döneminde 42 (%30) öğrenci bulunmaktadır.

**Tablo 6.** Eğitim düzeyleri ve dönemlerine göre öğrenci dağılımı

	Ders Dönemi		Tez Dönemi		Toplam
	S	%	S	%	
Yüksek Lisans	22	15,7	48	34,3	70
Doktora	28	20	42	30	70
Toplam	50	35,7	90	64,3	140

Katılımcıların çoğunluğu (S=87, %62,2) 20 ile 30 yaş aralığında bulunan öğrencilerden oluşmaktadır. 31 ile 40 yaş aralığında 37 (%26,4) kişi, 41 ile 50 yaş aralığında ise 16 (%11,4) kişi bulunmaktadır. Ankete katılan öğrencilerin yaş gruplarına göre dağılımı Tablo 7'de görülmektedir. Ankette bulunan 51 yaş ve üzeri seçeneğini işaretleyen olmadığı için tabloda belirtilmemiştir.

**Tablo 7.** Öğrenci sayısının yaşa göre dağılımı

	S	%
20 ile 30 yaş arası	87	62,2
31 ile 40 yaş arası	37	26,4

41 ile 50 yaş arası	16	11,4
Toplam	140	100

Ankette bulunan “Unvanınız nedir? (Çalışanlar için)” sorusuna 95 (%67,9) kişi cevap vermiştir. Cevap verenler arasında Uzman/Araştırmacı olduğunu belirtenler (S=30, %31,6) ile Akademik Personel olduğunu belirtenler (S=29, %30,5) neredeyse eşittir. Teknik Personel olduğunu belirtenler 20 (%21,1) kişi, İdari Personel olduğunu belirtenler ise 16 (%16,8) kişidir. Soruya cevap verenlerin unvanlarına göre toplam katılımcı içindeki yüzdeleri de ayrıca Tablo 8’de verilmiştir.

**Tablo 8.** Katılımcılar arasında çalışanların unvanları

Unvanlar	S	%	Toplam %
Uzman/Araştırmacı	30	31,6	21,4
Teknik Personel	20	21,1	14,3
İdari Personel	16	16,8	11,4
Akademik personel	29	30,5	20,7
Toplam	95	100	67,9

Akademik Personel olduğunu belirten katılımcılar ile Uzman/Araştırmacı olduğunu belirten katılımcılar (S=59), toplam katılımcı sayısının yaklaşık %40’ını oluşturmaktadır.

#### 4.1.1. Bilgi Arama Davranışları

Katılımcıların bilgiyi öncelikle hangi kaynaktan sağladığını öğrenmek amacıyla bilgi arama yaklaşımlarını kullanma sıklıkları 5’li likert tipi soru ile sorulmuştur. İfadelere verilen cevapların fakültelere göre ortalamaları ve genel ortalaması Tablo 9’da görülmektedir. Katılımcıların yaklaşık yarısı (S=75, %53,6) basılı kütüphane kaynaklarını “ara sıra” kullandıklarını belirtmişlerdir. 40 kişi (%28,6) “hiçbir zaman” seçeneğini işaretlemiştir. Basılı kütüphane kaynaklarını kullanımı için genel ortalama 2,08 olarak hesaplanmıştır. Katılımcıların çoğunluğunun bilgiye erişimde basılı kütüphane kaynaklarını kullanmayı tercih etmediklerini (S=140,  $\bar{x}$ =2,08) söyleyebiliriz. Kullanım ortalaması en yüksek olan Spor Bilimleri Fakültesi’dir (S=8,  $\bar{x}$ =2,75) ancak

istatistiksel olarak anlamlı bir farklılık göstermemektedir. Katılımcılardan 43 kişi (%31,4) “çoğu zaman” elektronik kütüphane kaynaklarını kullandıklarını belirtirken, “her zaman” (S=30, %21,9) ve “genellikle” (S=33, %24,1) seçeneğini işaretleyen katılımcı sayısı neredeyse eşittir. 25 kişi (%18,2) “ara sıra” seçeneğini işaretlerken, en az işaretlenen seçenek “hiçbir zaman” (S=6, %4,4) olmuştur. Katılımcıların büyük çoğunluğu (S=106, %72,4) bilgiye erişimde elektronik kütüphane kaynaklarını kullanmaktadır ve fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

**Tablo 9.** Bilgi arama davranışları

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	
Basılı kütüphane kaynakları	38	2,03	30	2,13	14	1,64	12	2,50	32	1,97	8	2,75	<b>2,08</b>
Elektronik kütüphane kaynakları	38	3,39	33	3,73	14	3,71	12	4,00	33	2,97	7	3,86	<b>3,48</b>
Google Scholar vb akademik internet platformları	39	4,44	33	4,45	15	3,87	12	4,42	33	4,33	8	4,50	<b>4,36</b>
Sağlık konusunda uzman veri tabanları	38	3,66	32	3,34	15	3,20	12	4,00	33	3,76	8	3,13	<b>3,56</b>
Sağlık Bakanlığının sağladığı içerikler	38	2,45	31	2,26	14	2,43	12	2,83	31	2,03	7	2,57	<b>2,35</b>
Diğer araştırmacıların paylaşımları	38	3,21	31	3,06	15	2,67	12	3,17	31	2,81	8	3,75	<b>3,05</b>
Sahip olduğu kendi koleksiyonu	38	2,55	31	2,39	15	2,27	12	2,08	30	2,00	7	2,29	<b>2,30</b>

Katılımcıların yarıdan fazlası (S=82, %58,6) Google Scholar vb akademik internet platformlarını kullanım sıklığı için “her zaman” seçeneğini işaretlemişlerdir. Aynı ifade için 32 kişi (%22,9) “çoğu zaman”, 21 kişi (%15) “genellikle” seçeneğini işaretlemiştir. En az işaretlenen seçenekler “ara sıra” (S=4, %2,9) ile “hiçbir zaman” (S=1, %0,7) seçenekleri olmuştur. Katılımcıların neredeyse tamamı bilgiye erişimde (S=135, %96) Google Scholar vb akademik internet platformlarını tercih etmektedir (S=140,  $\bar{x}$ =4,36) ve fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılardan 40 kişi (%29) sağlık konusunda uzman veri tabanlarını kullanma sıklığı için “her zaman” seçeneğini işaretlemişlerdir. “Çoğu zaman” (S=33, %23,9) seçeneği ile “genellikle” (S=35, %25,4) seçeneğini işaretleyen katılımcı sayısı neredeyse eşittir. 24

kişi (%17,4) “ara sıra” seçeneğini işaretlerken, en az işaretlenen seçenek (S=6, %4,3) “hiçbir zaman” seçeneği olmuştur. Katılımcıların neredeyse beşte dördü sağlık konusunda uzman veri tabanlarını kullanmaktadır ve fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcıların neredeyse yarısı (S=64, %48,1) Sağlık Bakanlığı’nın sağladığı içerikleri kullanma sıklığı için “ara sıra” seçeneğini işaretlemiştir. 26 kişi (%19,5) “hiçbir zaman” seçeneğini, 21 kişi (%15,8) ise “genellikle” seçeneğini işaretlemiştir. Katılımcılardan 15 kişi (%11,3) “çoğu zaman” seçeneğini işaretlerken en az işaretlenen seçenek (S=7, %5,3) “her zaman seçeneği olmuştur. Sağlık Bakanlığı’nın sağladıkları içerikleri kullanımının (S=133,  $\bar{x}$ =2,35) fakültelere göre istatistiksel olarak anlamlı farklılık göstermeksizin benzer olduğu söylenebilir.

Bilgiye erişimde diğer araştırmacıların paylaşımlarını kullanma sıklığı “ara sıra” (S=41, %30,4) seçeneğini işaretleyenler ile “genellikle” seçeneğini işaretleyenler (S=41, %30,4) için neredeyse eşittir. 32 kişi (%23,7) çoğu zaman seçeneğini işaretlerken, 15 kişi (%11,1) “her zaman” seçeneğini işaretlemiştir. En az işaretlenen seçenek (S=7, %5,2) “hiçbir zaman” seçeneği olmuştur. Diğer araştırmacıların paylaşımlarını kullanma sıklığı (S=135,  $\bar{x}$ =3,05) fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir. Katılımcılardan 56 kişi (%40) bilgiye erişimde sahip olduğu kendi koleksiyonunu kullanma sıklığını “ara sıra” olarak belirtmiştir. “Hiçbir zaman” seçeneği (S=34, %24,3) ile “genellikle” seçeneğini (S=33, %23,6) işaretleyen katılımcıların sayısı neredeyse eşittir. 12 kişi (%8,6) “çoğu zaman” seçeneğini işaretlerken, en az işaretlenen seçenek “her zaman” (S=5, %3,6) olmuştur. Sahip olduğu kendi koleksiyonunu kullanma yaklaşımının katılımcıların büyük çoğunluk için az tercih edilen (S=140,  $\bar{x}$ =2,30) bir yaklaşım olduğu söylenebilir. Bu yaklaşım fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılara gereksinim duydukları bilgiyi internette ararken hangi yaklaşımları daha sık tercih ettikleri 5’li likert tipi soruyla sorulmuştur. Katılımcıların verdiği cevapların sıklığı, yüzdesi ve ortalama değerleri hesaplanarak Tablo 10’da gösterilmiştir.

**Tablo 10.** Bilgiyi internette ararken tercih edilen yaklaşımlar

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	$\bar{x}$
Aradığım konuyu doğrudan google yandex vb arama motorlarına yazar ilk gelen sayfaları incelerim	39	3,64	33	2,94	15	2,80	12	2,75	33	2,67	8	3,00	<b>3,04</b>
Konuya ilişkin bilgileri uzman sayfalardan mutlaka teyit ederim (Sağlık Bakanlığı, TÜİK, Kalkınma Bakanlığı gibi otorite kurumlar)	39	3,82	33	3,91	15	3,80	12	3,83	33	3,39	8	3,63	<b>3,73</b>
Araştırma planı yaparım	39	3,87	33	3,91	15	3,93	12	3,83	33	3,91	8	3,63	<b>3,88</b>
Bağlaçları (ve/veya/değil) kullanarak anahtar kelimeleri ilişkilendirerek arama yaparım	39	4,28	33	3,70	15	4,07	12	4,50	33	4,27	8	3,50	<b>4,09</b>
Cevaplamam gereken soruyu aynen yazarak arama yaparım	39	2,77	32	2,55	15	2,47	12	2,83	33	2,42	8	2,38	<b>2,59</b>
İnternet bilgilerini farklı kanallardan (konu uzmanı otorite kaynak, kitap vb) mutlaka doğrularım	39	4,13	33	4,15	15	3,93	12	4,00	33	3,67	8	3,75	<b>3,97</b>
Bilgiye ulaşamadığımda anahtar kelimeleri değiştirerek tekrar arama yaparım.	39	4,72	33	4,67	15	4,53	12	4,67	33	4,52	8	4,00	<b>4,59</b>

Katılımcılar arasında “aradığım konuyu doğrudan Google, Yandex vb arama motorlarına yazar ilk gelen sayfaları incelerim” ifadesi için “ara sıra” seçeneğini (S=36, %25,7) işaretleyenler ile “çoğu zaman” (S=39, %27,9) seçeneğini işaretleyenlerin sayısı neredeyse eşittir. 13 kişi (%9,3) “her zaman” seçeneğini işaretlemiştir. En çok işaretlenen (S=45, %32,1) “seyrek” seçeneği iken en az işaretlenen (S=7, %5) seçenek “hiçbir zaman” olmuştur. Fakültelerin ortalamaları 2,67 ile 3,64 arasında değişmekte olup genel

ortalama 3,04 hesaplanmıştır. Katılımcıların bu yaklaşımı ara sıra tercih ettiği söylenebilmekte olup fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcıların %41,4'ü (S=58) gereksinim duyduğu bilgiyi internette ararken “konuya ilişkin bilgileri uzman sayfalardan mutlaka teyit ederim” ifadesi için “çoğu zaman” seçeneğini işaretlemiştir. 34 kişi (%24,3) “her zaman” seçeneğini işaretlerken, 26 kişi (%18,6) “ara sıra” seçeneğini işaretlemiştir. 20 kişi (%14,3) bu ifade için “seyrek” seçeneğini işaretlerken en az işaretlenen seçenek (S=2, %1,4) “hiçbir zaman” olmuştur. Bu ifadeye katılım için hesaplanan genel ortalama 3,73 olup fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir. Katılımcılardan %40'ı (S=56) “çoğu zaman” araştırma planı yaptığını belirtmiştir. “Her zaman” seçeneğini işaretleyen 45 kişi (%32,1), “ara sıra” seçeneğini işaretleyen ise 20 (%14,3) kişidir. 15 kişi (%10,7) “seyrek” seçeneğini işaretlerken, en az işaretlenen seçenek (S=4, %2,9) “hiçbir zaman” olmuştur. Fakültelerin ortalamaları 3,63 ile 3,93 arasında değişmekte olup genel ortalama 3,88 hesaplanmıştır. Katılımcıların büyük çoğunluğu (S=121, %86) araştırma planı yapmaktadır ve bu yaklaşımı tercih edenler fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılardan 59 kişi (%42,1) bağlaçları (ve/veya/değil) kullanarak anahtar kelimeleri ilişkilendirerek arama yapma yaklaşımını kullanma sıklığı için “her zaman” seçeneğini işaretlemiştir. 50 kişi (%35,7) “çoğu zaman” seçeneğini, 19 kişi (%13,6) ise “ara sıra” seçeneğini işaretlemiştir. En az işaretlenen seçenekler “seyrek” (S=9, %6,4) seçeneği ile “hiçbir zaman” (S=3, %2,1) seçeneği olmuştur. Fakültelerin ortalamaları 3,50 ile 4,28 arasında değişmekte olup genel ortalama 4,09 hesaplanmıştır. Katılımcıların çoğunluğu bu yaklaşımı tercih etmektedir ve %95 güven düzeyinde fakültelere göre anlamlı farklılık göstermemektedir.

Katılımcılardan 2 kişi (%1,4) “cevaplamam gereken soruyu aynen yazarak arama yaparım” ifadesi için “her zaman” seçeneğini işaretlemiştir. 15 kişi (%10,7) “hiçbir zaman” seçeneğini işaretlerken, 24 kişi (%17,1) “çoğu zaman” seçeneğini işaretlemiştir. En çok işaretlenen seçenekler “seyrek” seçeneği (S=56, %40) ile “ara sıra” seçeneği (S=43, %30,7) olmuştur. Katılımcıların büyük çoğunluğunun tercih etmediği bu yaklaşım fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.



Katılımcıların yarıdan fazlası (S=88, %62,9) bilgiye ulaşamadığında anahtar kelimeleri değiştirerek tekrar arama yapma yaklaşımı için “her zaman” seçeneğini işaretlemiştir. 48 kişi (%34,3) “çoğu zaman” seçeneğini işaretlerken en az işaretlenen seçenekler “ara sıra” (S=3, %2,1) ile “seyrek” seçeneği (S=1, %0,7) olmuştur. “hiçbir zaman” seçeneğini işaretleyen olmamıştır. Katılımcıların neredeyse tamamının her zaman bilgiye ulaşamadığında anahtar kelimeleri değiştirerek tekrar arama yaptığı söylenebilir (S=140,  $\bar{x}$ =4,59). Bu yaklaşımı tercih edenler fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılara bir konuda aradıkları bilgiyi teyit amacıyla birden fazla kaynağı kullanma alışkanlıkları 5’li likert tipi soru ile sorulmuştur. Katılımcılardan “hiçbir zaman” seçeneğini işaretleyen olmamıştır (bkz. Tablo 11).

**Tablo 11.** Katılımcıların bilgiyi başka kaynaktan teyit alışkanlıkları

	Hiçbir zaman	Ara Sıra	Genellikle	Çoğu Zaman	Her Zaman	Toplam	$\bar{x}$	$\sigma$
Tıp	S 0	1	4	11	23	39	4,44	0,788
	% 0	2,6	10,3	28,2	59	100		
Sağlık Bilimleri	S 0	0	2	14	17	33	4,45	0,617
	% 0	0	6,1	42,4	51,5	100		
Eczacılık	S 0	2	1	3	9	15	4,27	1,100
	% 0	13,3	6,7	20	60	100		
Hemşirelik	S 0	0	1	3	8	12	4,58	0,669
	% 0	0	8,3	25	66,7	100		
FTR	S 0	3	5	11	14	33	4,09	0,980
	% 0	9,1	15,2	33,3	42,4	100		
Spor Bilimleri	S 0	0	1	2	5	8	4,50	0,756
	% 0	0	12,5	25	62,5	100		
Toplam	S 0	6	14	44	76	140	4,36	0,832
	% 0	4,3	10	31,4	54,3	100		

Tüm katılımcılar aradıkları bilgiyi birden fazla kaynaktan teyit ettiklerini belirtmişlerdir. Katılımcılardan 6 kişi (%4,3) “ara sıra” seçeneğini işaretlerken, 14 kişi (%10) “genellikle” seçeneğini işaretlemiştir. Katılımcıların 44’ü (%31,4) “çoğu zaman”, yarisından fazlası (S=76, %54,3) “her zaman” seçeneğini işaretlemiştir. Bu yaklaşımı tercih edenler fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılara araştırmalarında kullandıkları veriyi öncelikle hangi yolla elde ettiklerinin sorulduğu soruda (bkz. Tablo 12) Sağlık Bakanlığı veri tabanını kullanma sıklığı için 50 kişi (%35,7) “seyrek” seçeneğini işaretlemiştir. “Ara sıra” seçeneğini işaretleyen 40 kişi (%28,6) ve “hiçbir zaman” seçeneğini işaretleyen 26 kişidir (%18,6). 17 kişi (%12,1) “çoğu zaman” seçeneğini işaretlerken, en az işaretlenen seçenek “her zaman” (S=7, %5) olmuştur. Veri elde etmek için Sağlık Bakanlığı veri tabanını kullanma yaklaşımının genel ortalaması 2,49 hesaplanmış olup fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

**Tablo 12.** Veri elde etme biçimleri

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	$\bar{x}$
Sağlık Bakanlığı veri tabanından	39	2,56	33	2,48	15	2,40	12	3,17	33	2,33	8	2,00	<b>2,49</b>
Araştırma yapacağı kurumun veri tabanından	39	3,62	33	3,33	15	3,13	12	4,08	33	3,33	8	2,75	<b>3,42</b>
Veri depolarından diğer araştırmacıların verilerini kullanarak	39	3,13	33	2,58	15	3,13	12	3,75	33	2,91	8	3,12	<b>3,00</b>
Anket/ölçek vs aracılığıyla toplanan veriler	39	3,23	33	3,97	15	2,53	12	3,92	33	3,70	8	4,50	<b>3,57</b>
Muayene, tetkik, ses veya görüntü kaydı gibi çalışmalar sonucunda elde edilen veri	39	2,97	33	2,85	15	2,67	12	2,83	33	3,58	8	3,38	<b>3,06</b>
Görüşme veya içerik analizi aracılığıyla toplanan veriler	39	2,72	33	3,48	15	2,40	12	2,92	33	3,27	8	4,00	<b>3,09</b>
Basılı kaynaklar veya arşiv belgeleri üzerinde araştırma yaparak	39	3,44	33	3,15	15	3,60	12	3,67	33	3,03	8	4,25	<b>3,36</b>

Araştırmanın yapılacağı kurumun veri tabanından veri elde etme yaklaşımı için 52 kişi (%37,1) “çoğu zaman” seçeneğini işaretlemiştir. 32 kişi (%22,9) “ara sıra” seçeneğini işaretlerken 25 kişi (%17,9) “her zaman” seçeneğini işaretlemiştir. Bu yaklaşım için en az işaretlenen seçenekler “seyrek” (S=19, %13,6) seçeneği ile “hiçbir zaman” (S=12,

%8,6) seçeneği olmuştur. Hemşirelik Fakültesi'nin ortalaması 4,08 olup çoğu zaman bu yaklaşımı tercih ettikleri söylenebilir. Çoğunluğun tercih ettiği bu yaklaşımın genel ortalaması 3,42'dir ve fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Veri depolarından diğer araştırmacıların verilerini kullanma yaklaşımı için katılımcıların 42'si (%30) "ara sıra" seçeneğini işaretlemiştir. Katılımcılardan 36 kişi (%25,7) "çoğu zaman" seçeneğini, 31 kişi (%22,1) ise "hiçbir zaman" seçeneğini işaretlemiştir. 19 kişi (%13,6) "her zaman" seçeneğini işaretlerken en az işaretlenen seçenek "seyrek" (S=12, %8,6) seçeneği olmuştur. Veri elde etmek için veri depolarını kullanma yaklaşımının genel ortalaması 3,00 hesaplanmış olup, katılımcıların ara sıra bu yaklaşımı tercih ettiği söylenebilir. Veri depolarından diğer araştırmacıların verilerini kullanma yaklaşımı fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Veri elde etme yaklaşımı olarak "anket/ölçek vb aracılığıyla veri toplama" ifadesi için 46 kişi (%32,9) "çoğu zaman" seçeneğini işaretlemiştir. "Her zaman" seçeneğini 36 kişi (%25,7), "ara sıra" seçeneğini ise 30 kişi (%21,4) işaretlemiştir. Katılımcılardan 18 kişi (%12,9) "seyrek" seçeneğini işaretlerken en az işaretlenen seçenek (S=10, %7,1) "hiçbir zaman" olmuştur. Bu yaklaşımı en az tercih eden Eczacılık Fakültesi (S=15,  $\bar{x}$ =2,53), en çok tercih eden ise Spor Bilimleri Fakültesi (S=8,  $\bar{x}$ =4,50) olmuştur. Veri elde etmek için anket/ölçek vb kullanma yaklaşımı veri elde etme biçimleri arasında en çok tercih edildiği gibi fakültelere göre de istatistiksel olarak anlamlı farklılık göstermemektedir.

Muayene, tetkik, ses veya görüntü kaydı gibi çalışmalar sonucunda veri elde etme yaklaşımı için 22 kişi (%15,7) "her zaman" seçeneğini işaretlemiştir. 40 kişi "çoğu zaman" seçeneğini işaretlerken, 29 kişi (%20,7) "ara sıra" seçeneğini işaretlemiştir. "Hiçbir zaman" seçeneğini işaretleyenler (S=26, %18,6) ile "seyrek" seçeneğini işaretleyenler (S=23, %16,4) neredeyse eşittir. Muayene, tetkik, ses veya görüntü kaydı gibi çalışmalar sonucunda veri elde etme yaklaşımının genel ortalaması 3,06 olup, katılımcıların ara sıra bu yaklaşımı tercih ettiği söylenebilir. Bu yaklaşım fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

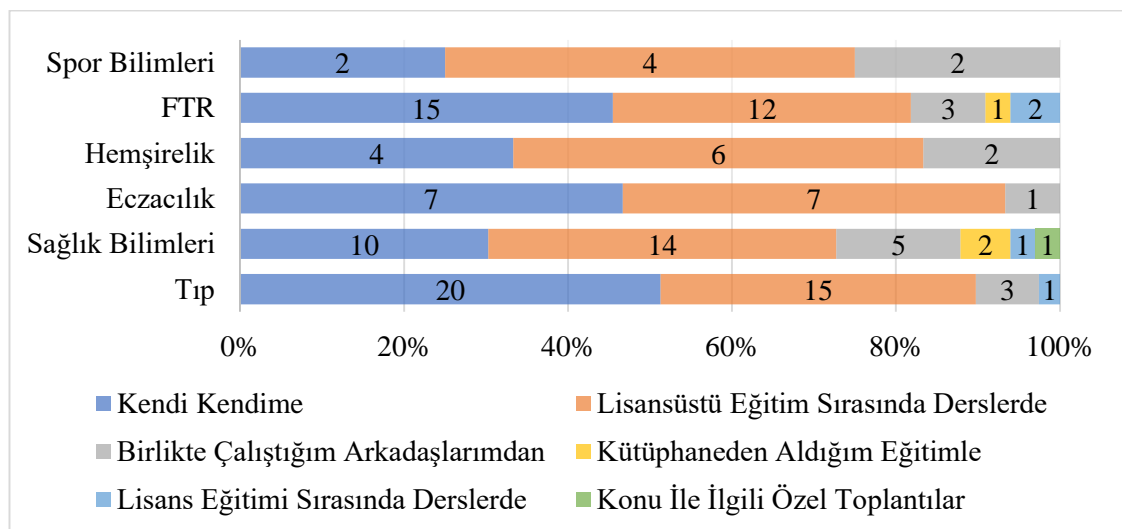
Görüşme veya içerik analizi aracılığıyla veri toplama yaklaşımı için 37 kişi (%26,4) "ara sıra" seçeneğini işaretlemiştir. 32 kişi (%22,9) "çoğu zaman" seçeneğini işaretlerken,

“seyrek” seçeneğini (S=26, %18,6), “her zaman” seçeneğini (S=24, %17,1) ve “hiçbir zaman” seçeneğini (S=21, %15) işaretleyenler neredeyse eşittir. Bu yaklaşımı en çok tercih eden Spor Bilimleri Fakültesi’nin (S=8,  $\bar{x}$ =4,00) çoğu zaman bu yaklaşımı tercih ettiği söylenebilir. Görüşme veya içerik analizi aracılığıyla toplanan verilerin kullanım ortalaması 3,09 olup, fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılar arasında basılı kaynaklar veya arşiv belgeleri üzerinde araştırma yaparak veri elde etme yaklaşımı için 44 kişi (%31,4) “çoğu zaman” seçeneğini işaretlemiştir. “Ara sıra” seçeneğini işaretleyenler (S=32, %22,9) ile “her zaman” seçeneğini işaretleyenler (S=29, %20,7) neredeyse eşittir. Benzer şekilde “hiçbir zaman” seçeneğini işaretleyenler (S=17, %12,1) ile “seyrek” seçeneğini işaretleyenler (S=18, %12,9) de neredeyse eşittir. Basılı kaynaklar veya arşiv belgeleri üzerinde araştırma yaparak veri elde etme yaklaşımı en çok Spor Bilimleri Fakültesi (S=8,  $\bar{x}$ =4,50) tarafından tercih edilmektedir ancak istatistiksel olarak fakültelere göre anlamlı farklılık göstermemektedir.

Katılımcıların bilgi arama yaklaşımlarını hangi yolla öğrendikleri çoktan seçmeli bir soruyla sorulmuştur (bkz. Tablo 13). Eşit sayıda katılımcı kendi kendine (S=58, %41,4) ve lisansüstü eğitimi sırasında derslerde (S=58, %41,4) öğrendiğini belirtmiştir. 16 katılımcı (%11,4) birlikte çalıştığı arkadaşlarından, 4 katılımcı (%3) lisans eğitimi sırasında derslerde, 3 katılımcı (%2,1) ise kütüphaneden aldığı eğitimle öğrendiğini belirtmiştir.

**Tablo 13.** Fakültelere göre bilgi arama davranışlarını öğrenme şekli

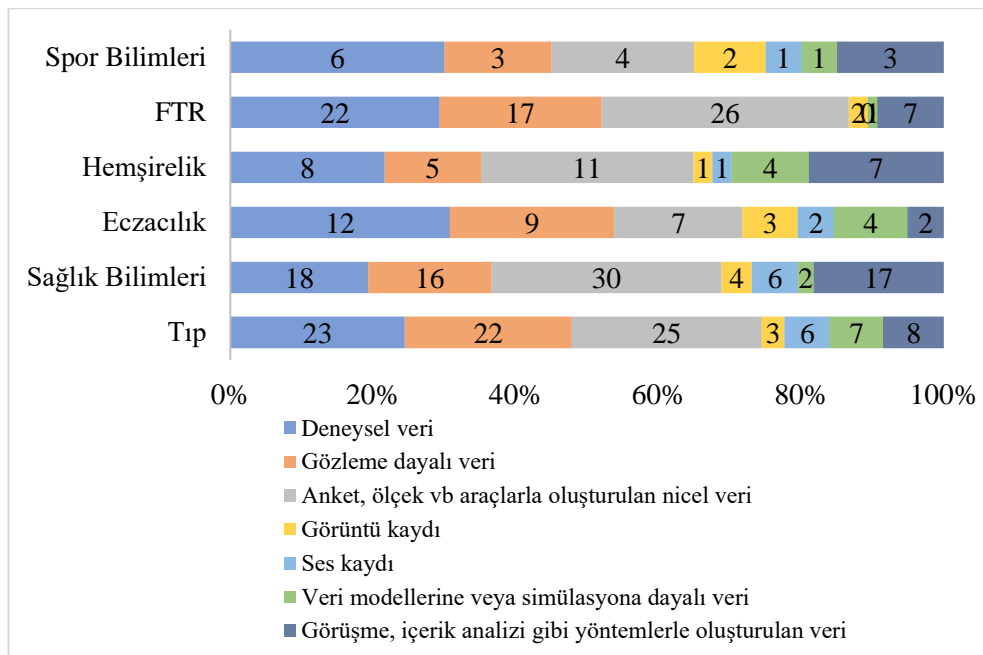


Diğer seçeneğini işaretleyen 1 katılımcı (%0,7) bilgi arama yaklaşımını konu ile ilgili özel toplantılarda öğrendiğini ifade etmiştir. Katılımcıların bilgi arama yaklaşımlarını öğrenme yolu fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

#### 4.1.2. Araştırma Verilerinin Yönetimi

Katılımcılara araştırmalarında çoğunlukla hangi tür veri kullandıkları birden fazla seçeneğin işaretlenebildiği bir soru formatıyla sorulmuştur. Katılımcıların fakültelerine göre kullandıkları veri türleri Tablo 14'te gösterilmiştir. 103 kişi (%73,6) anket, ölçek vb araçlarla oluşturulan veri kullandığını, 89 kişi (%64) deneysel veri, 72 kişi (%51,4) ise gözleme dayalı veri kullandığını belirtmiştir.

**Tablo 14.** Fakültelere göre kullanılan veri türleri

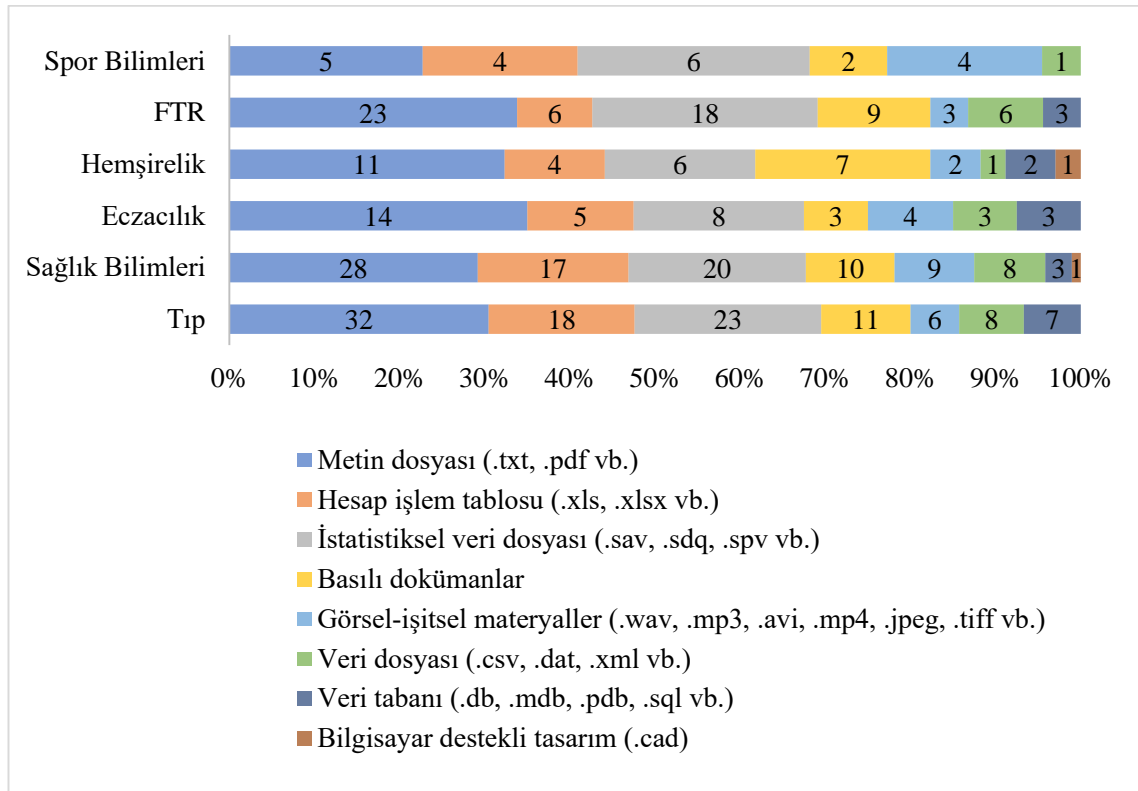


İçerik analizi gibi yöntemlerle oluşturulan veri seçeneğini 44 kişi (%31,4) işaretlemiştir. En az işaretlenen veri türleri veri modellerine veya simülasyona dayalı veri (S=19, %13,6), ses kaydı (S=16, %11,4) ve görüntü kaydı (S=15, %10,7) olmuştur. Katılımcıların çeşitli veri türlerini bir arada kullandıkları söylenebilir. Tıp, Sağlık Bilimleri ve FTR alanında eğitim alan katılımcıların kullandıkları veri türleri neredeyse

aynıdır. Kullanılan veri türü fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılara kullandıkları veri formatı da birden fazla seçeneğin işaretlenebildiği soru formatıyla sorulmuştur (bkz. Tablo 15). En çok işaretlenen seçenek (S=113, %80,7) metin dosyası (.txt, .pdf vb.) seçeneği olmuştur. İstatistiksel veri dosyası (.sav, .sdq, .spv vb.) seçeneğini 81 kişi (%57,9) işaretlemiştir. 54 kişi (%38,6) hesap işlem tablosu (.xls, .xlsx vb.) seçeneğini, 42 kişi (%30) ise basılı dokümanlar seçeneğini işaretlemiştir.

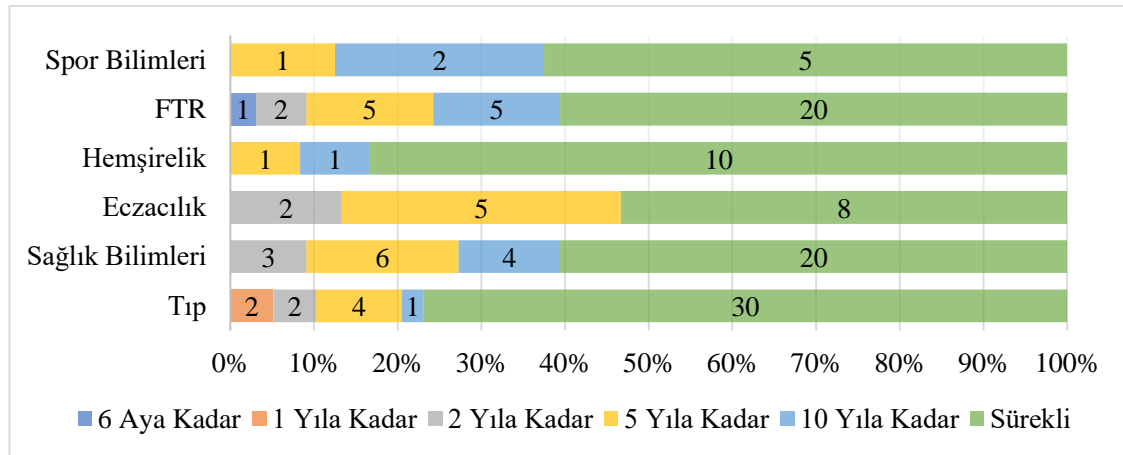
**Tablo 15.** Fakültelere göre kullanılan veri formatları



Görsel-ışitsel materyaller (.wav, .mp3, .avi, .mp4, .jpeg, .tiff vb.) seçeneğini işaretleyenler (S=28, %20) ile veri dosyası (.csv, .dat, .xml vb.) seçeneğini işaretleyenler (S=27, %19,3) neredeyse eşittir. Veri tabanı (.db, .mdb, .pdb, .sql vb.) seçeneğini işaretleyenler 18 kişi (%12,9) iken en az işaretlenen seçenek (S=2, %1,4) bilgisayar destekli tasarım (.cad) seçeneği olmuştur. Fakültelere göre kullanılan veri formatları tablosuna bakıldığında Tıp alanı ile Sağlık Bilimleri alanı öğrencilerinin veri formatları kullanım alışkanlıklarının aynı olduğu söylenebilir. Kullanılan veri formatları fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

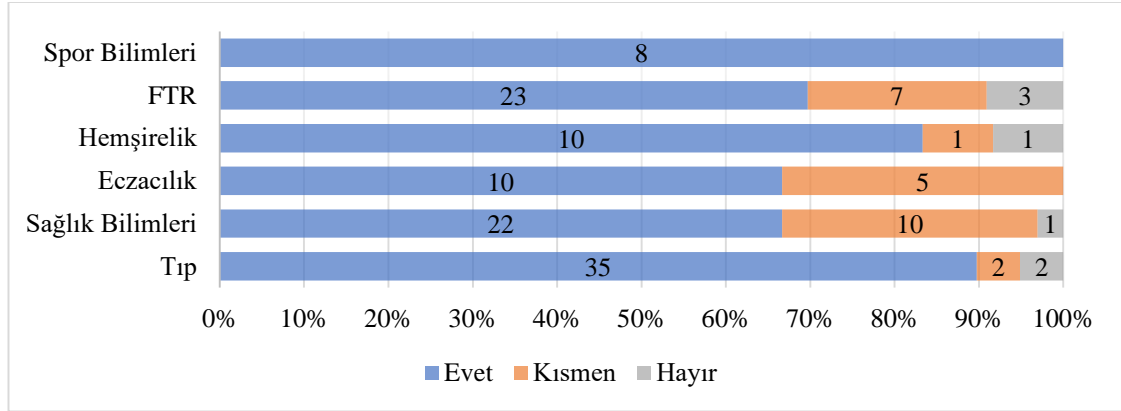
Katılımcılara araştırma tamamlandıktan sonra verilerini ne kadar süre sakladıkları sorulmuş ve sıklıklar Tablo 16’da gösterilmiştir. Katılımcılardan 93 kişi (%66,4) araştırma tamamlandıktan sonra verilerini sürekli olarak sakladığını belirtmiştir. 13 kişi (%9,4) 10 yıla kadar, 22 kişi (%15,7) ise verilerini 5 yıla kadar sakladığını belirtmiştir.

**Tablo 16.** Fakültelere göre verilerini saklama süreleri

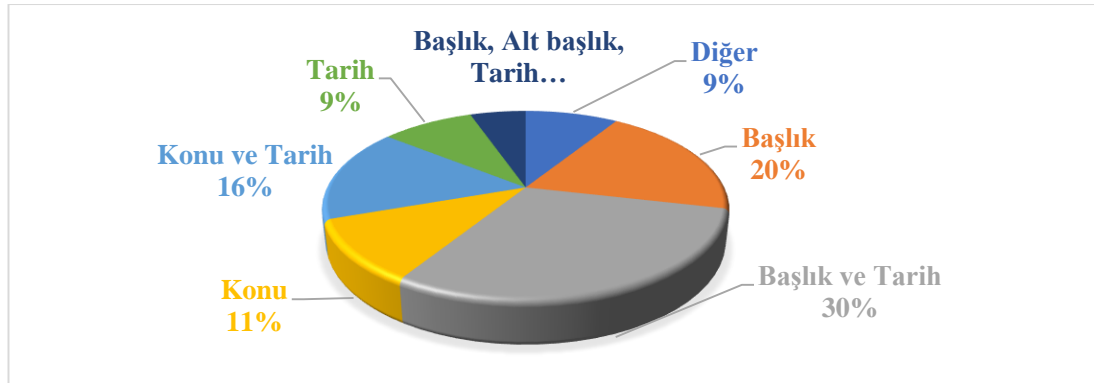


Verilerini 2 yıla kadar sakladığını belirten 9 kişi (%6,4) Tıp, Eczacılık, Sağlık Bilimleri ve FTR alanında neredeyse eşit dağılım göstermektedir. Verilerini 1 yıla kadar saklayan 2 kişi (%1,4) Tıp alanında öğrencidir. Verilerini 6 aya kadar sakladığını belirten 1 kişi (%0,7) FTR alanında öğrencidir. Katılımcıların büyük çoğunluğunun verilerini uzun sürelerle saklama eğiliminde olduğu söylenebilir. Araştırma tamamlandıktan sonra verilerini saklama süreleri fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Katılımcılara araştırma verilerini farklı bir ortamda yedekleyip yedeklemedikleri sorulmuştur (bkz. Tablo 17). Katılımcıların tamamına yakını (S=129, %92,1) verilerini yedeklediklerini belirtirken sadece 11 kişi (%7,9) yedekleme yapmadığını belirtmiştir. Büyük çoğunluğu (S=108, %77,1) verilerini saklarken veya yedeklerken tekrar erişimi kolaylaştırmak için içeriği üst veri oluşturarak tanımladığını belirtmiştir. 25 kişi (%7,9) kısmen üst veri tanımladığını, 7 kişi (%5) ise üst veri tanımlamadığını belirtmiştir. Üst veri tanımlama alışkanlığı fakültelere göre istatistiksel olarak anlamlı farklılık göstermemektedir.

**Tablo 17.** Fakültelere göre üst veri oluşturma alışkanlıkları

Verilerine üst veri tanımladığını belirten katılımcılar için herhangi bir üst veri standardı kullanıp kullanmadıkları, kullanıyorlarsa hangisi olduğu açık uçlu bir soru ile sorulmuştur (bkz. Tablo 18). Katılımcılardan 78 kişi (%55,7) bu soruya yanıt vermiştir. 22 kişi (%27,5) herhangi bir üst veri standardı kullanmadığını belirtmiştir.

**Tablo 18.** Üst veri oluşturmada kullanılan yöntemler

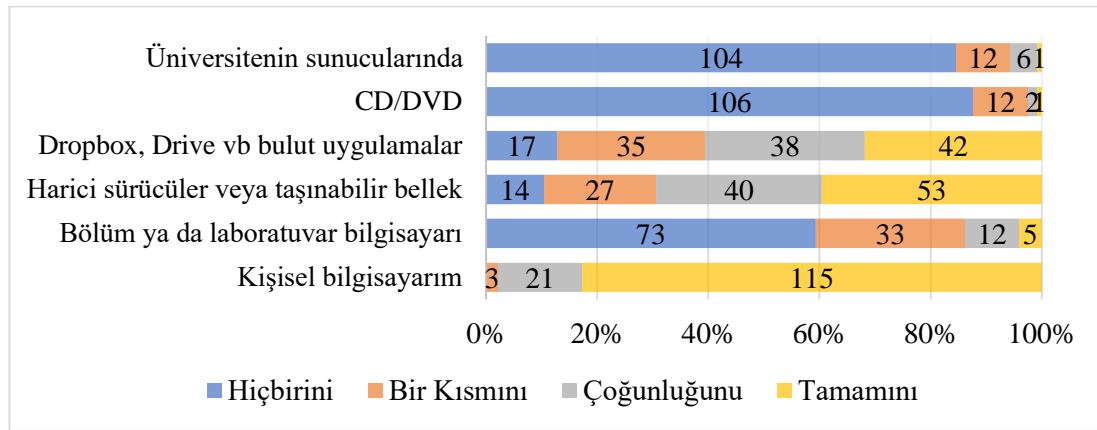
17 kişi (%30) verilerini “başlık ve tarih” kullanarak tanımladığını belirtmiştir. 11 kişi (%20) sadece “başlık” ile tanımlarken “konu ve tarih” ile tanımlayan 9 kişi (%16) vardır. Verilerini “konu” ile tanımlayan kişiler (S=6, %12) ile “tarih” ile tanımlayanların (S=5, %10) sayısı neredeyse eşittir. 3 kişi (%6) tanımlamada “başlık, alt başlık ve tarih” kullandığını belirtirken kullanılan diğer yöntemler “diğer” seçeneğinde toplanmıştır. Bu kategorideki 5 cevap ise şöyledir; “örneklem grubu ismi”, “EAD”, “gün-ay-yıl ya da yıl-yazar kodlaması yaparım”, “tanımlayıcı üst veri”, “konu veya yazar yıl kodlaması”.

Katılımcılara verilerini araştırma devam ederken ve araştırma tamamlandıktan sonra kullandıkları saklama ve yedekleme ortamları sorulmuştur. Araştırma devam ederken



veriyi saklama ortamları Tablo 19'da görülmektedir. Katılımcıların büyük çoğunluğu (S=115, %82,1) araştırma devam ederken verilerinin tamamını kişisel bilgisayarında sakladığını belirtmiştir. 53 kişi (%37,9) harici sürücüler veya taşınabilir bellekte sakladığını belirtirken, 42 kişi (%30) Dropbox, Drive vb bulut uygulamalar seçeneğini işaretlemiştir. Katılımcılardan sadece 1 kişi (%7) verilerinin tamamını CD/DVD'de, 1 kişi (%7) üniversitenin sunucularında tuttuğunu belirtmiştir.

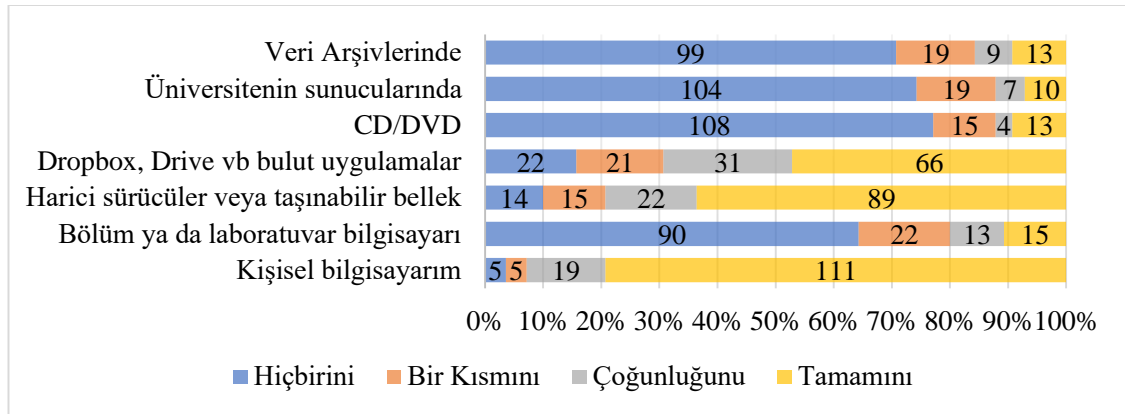
**Tablo 19.** Araştırma devam ederken veriyi saklama ortamları



Verilerinin çoğunluğunu kişisel bilgisayarında saklayan 21 kişi (%15), harici sürücüler veya taşınabilir bellekte saklayan 40 kişi (%28,6), Dropbox, Drive vb bulut uygulamalarda saklayan 38 kişi (%27,1) bulunmaktadır. 6 kişi (%4,3) üniversitenin sunucularında ve sadece 2 kişi CD/DVD'de tuttuğunu belirtmiştir.

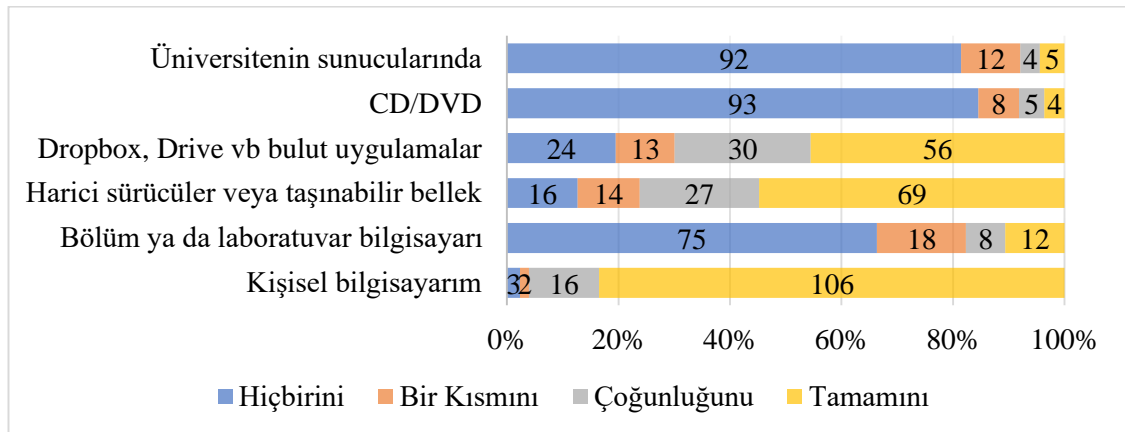
Katılımcılardan 35 kişi (%25) verilerinin bir kısmını Dropbox, Drive vb bulut uygulamalarda, 33 kişi (%23,6) ise üniversitenin sunucularında tuttuğunu belirtmiştir. Katılımcılarının veri saklama ortamı olarak ilk tercihinin kişisel bilgisayarı, ikinci olarak da harici sürücü/taşınabilir bellek ve bulut uygulamalar olduğu söylenebilir. Araştırma devam ederken tercih edilen veri saklama ortamları fakültelere göre anlamlı farklılık göstermemektedir.

Araştırma tamamlandıktan sonra katılımcıların veriyi saklama ortamları Tablo 20'de görülmektedir. Katılımcılardan 111 kişi (%79,3) araştırma tamamlandıktan sonra verilerinin tamamını kişisel bilgisayarında sakladığını belirtmiştir. 89 kişi (%63,6) harici sürücüler veya taşınabilir bellekte sakladığını, 66 kişi (%47,1) ise Dropbox, Drive vb bulut uygulamalarda sakladığını belirtmiştir.

**Tablo 20.** Araştırma tamamlandıktan sonra veriyi saklama ortamları

CD/DVD’de sakladığını belirtenler (S=13, %9,3) ile veri arşivlerinde sakladığını belirtenlerin (S=13, %9,3) sayısı eşittir. 10 kişi ise (%7,1) üniversitenin sunucularında sakladığını belirtmiştir. 31 kişi (%22,1) verilerinin çoğunluğunu Dropbox, Drive vb bulut uygulamalarda, 22 kişi (%15,7) ise harici sürücü/taşınabilir bellekte sakladığını belirtmiştir. 19 kişi (%13,6) kişisel bilgisayarım, 13 kişi (%9,3) ise bölüm ya da laboratuvar bilgisayarı seçeneğini işaretlemiştir. Dropbox, Drive vb bulut uygulamalar 31 kişi (%22,1) tarafından, harici sürücüler veya taşınabilir bellek seçeneği 22 kişi (%15,7) tarafından işaretlenmiştir. En az tercih edilen seçenekler CD/DVD (S=4, %0,03), veri arşivleri (S=9, %0,06) ve üniversitenin sunucuları (S=7, %0,05) olmuştur. Araştırma tamamlandıktan sonra tercih edilen veri saklama ortamları fakülterele göre istatistiksel olarak anlamlı farklılık göstermemektedir.

Araştırma devam ederken katılımcıların veriyi yedekleme ortamları Tablo 21’de görülmektedir.

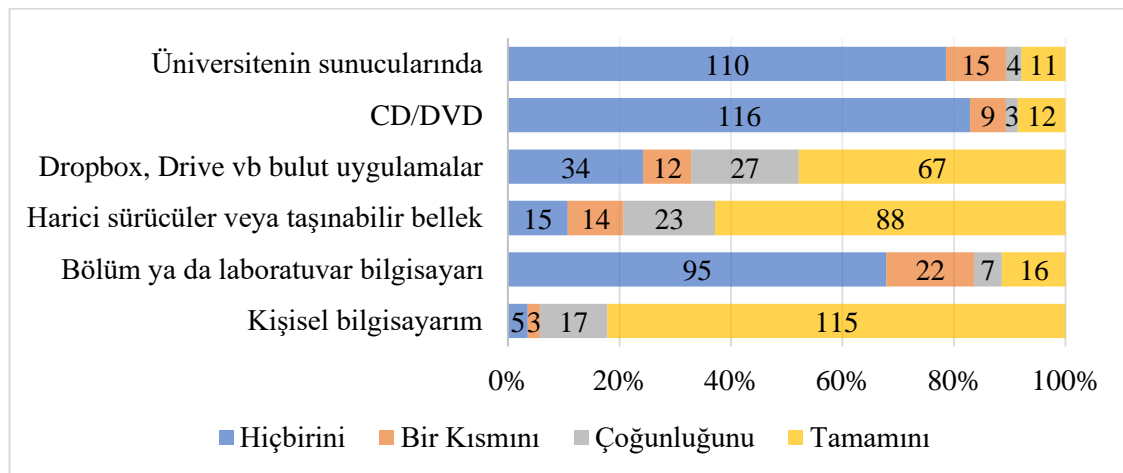
**Tablo 21.** Araştırma devam ederken veriyi yedekleme ortamları

Katılımcılar araştırma devam ederken verilerinin tamamını en çok (S=106, %75,7) kişisel bilgisayarlarında yedeklediklerini belirtmişlerdir. 69 kişi (%49,3) harici sürücüler veya taşınabilir bellek, 56 kişi (%40) Dropbox, Drive benzeri bulut uygulamaları tercih ettiğini belirtmiştir. Bölüm ya da laboratuvar bilgisayar, CD/DVD ve üniversitenin sunucuları seçenekleri en az tercih edilen seçenekler olmuştur.

Verilerinin çoğunluğunu Dropbox, Drive vb bulut uygulamalarda yedekleyenler 30 kişi (%21,4), harici sürücü veya taşınabilir bellekte yedekleyenler ise 27 kişidir (%19,3). 16 kişi kişisel bilgisayarında yedeklediğini belirtirken, en az tercih edilenler bölüm ya da laboratuvar bilgisayar, CD/DVD ve üniversitenin sunucuları olmuştur. Katılımcıların yarısından çoğu verilerinin hiçbirini üniversitenin sunucularında (S=92, %65,7), CD/DVD'de (S=93, %66,4) ve bölüm ya da laboratuvar bilgisayarında (S=75, %53,6) yedeklemediğini belirtmiştir. Araştırma devam ederken tercih edilen veri yedekleme ortamları fakülterlere göre anlamlı farklılık göstermemektedir.

Araştırma tamamlandıktan sonra verilerini kalıcı olarak yedekleme tercihlerine bakıldığında (bkz. Tablo 22) verilerinin tamamını kişisel bilgisayarında yedeklediğini belirten 115 kişi (%82,1) vardır. 88 kişi (%62,9) harici sürücü ve taşınabilir bellek tercih ederken, 67 kişi (%47,9) Dropbox, Drive vb bulut uygulamaları tercih ettiğini belirtmiştir.

**Tablo 22.** Araştırma tamamlandıktan sonra veriyi yedekleme ortamları



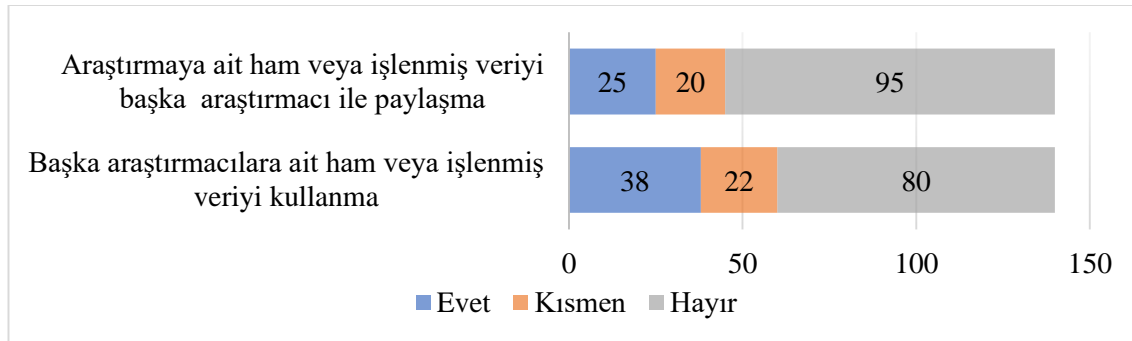
Verilerinin çoğunluğunu bulut uygulamalarda yedeklemeyi tercih eden 27 kişi (%19,3), harici sürücü veya taşınabilir bellekte yedeklemeyi tercih eden ise 23 kişi (%16,4) kişi vardır. Katılımcıların büyük çoğunluğu verilerinin hiçbirini üniversite sunucularında

(S=110, %78,6), CD/DVD'de (S=116, %82,9) veya bölüm ya da laboratuvar bilgisayarında (S=95, %67,9) yedeklemediğini belirtmiştir. Araştırma tamamlandıktan sonra verileri harici sürücüler veya taşınabilir bellekte yedekleme davranışı fakültelere göre istatistiksel olarak anlamlı farklılık göstermektedir. Katılımcıların veriyi saklarken ve yedeklerken genel tercihinin, verinin tamamı için öncelikle kişisel bilgisayarları olmak üzere, ikincil olarak harici sürücü/taşınabilir bellek ile bulut uygulamalar olduğu söylenebilir. Veri saklama alışkanlıklarına bakıldığında araştırma tamamlandıktan sonra verilerini kişisel bilgisayarda saklayan katılımcı sayısı belirgin düşüş göstermezken verinin tamamı için diğer veri saklama ortamlarında belirgin artış görülmektedir. Aynı eğilim veriyi yedekleme ortamlarında da görülmektedir.

Katılımcılara daha önce yurtdışında eğitim görüp görmedikleri, fon desteği alarak araştırma yürütüp yürütmedikleri ile veri yönetim planı hazırlayıp hazırlamadıkları sorulmuştur. Katılımcılardan 28 kişi (%20) bir eğitim veya değişim programıyla yurtdışında eğitim gördüğünü belirtirken, 112 kişi (%80) görmediğini belirtmiştir. Katılımcılar arasında daha önce fon desteği alarak bir araştırma yürüttüğünü belirten 24 kişi (%17,1), daha önce veri yönetim planı hazırladığını belirten ise 11 kişi (%7,9) vardır. Daha önce fon desteği alanlar ile daha önce veri yönetim planı hazırlayanlar arasında ise istatistiksel olarak anlamlı bir ilişki vardır ( $\chi^2(4) = 15.762$ ;  $p = 0,012$ ).

#### **4.1.3. Araştırma Verilerinin Paylaşımı**

Katılımcılara daha önce araştırma verilerini başka araştırmacılarla paylaşım paylaşmadıkları ve daha önce başka araştırmacıların verilerini kullanıp kullanmadıkları sorulmuştur (bkz Tablo 23). Katılımcılardan 25 kişi (%17,9) daha önce araştırmasına ait ham veya işlenmiş verileri başka bir araştırmacı ile paylaştığını, 20 kişi (%14,3) ise verilerini kısmen paylaştığını belirtmiştir. Katılımcıların büyük çoğunluğu (S=95, %67,9) ise daha önce verilerini paylaşmadığını belirtmiştir. Başka araştırmacılara ait ham veya işlenmiş verileri araştırmaları için kullanıp kullanmadıklarının sorulduğu soruya 38 kişi (%27,1) evet cevabı vermiştir. 22 kişi (%15,7) kısmen kullandığını belirtirken, çoğunluk (S=80, %57,1) bu soruya hayır cevabı vermiştir. Daha önce başka araştırmacıların verilerini kullanma davranışı fakültelere göre istatistiksel olarak anlamlı farklılık göstermektedir ( $\chi^2(4) = 15.762$ ;  $p = 0,010$ ).

**Tablo 23.** Araştırma verilerinin paylaşımı

Katılımcılara araştırma kapsamında elde ettikleri veya ürettikleri verinin genellikle paylaşım ve tekrar kullanım için uygun olup olmadığı sorulmuştur. 64 kişi (%46) evet, 57 kişi (%41) kısmen cevabı verirken, 18 kişi (%12,9) hayır cevabı vermiştir. Verilerin tekrar kullanım için uygun olmamasının olası nedenleri ise açık uçlu bir soru ile sorulmuştur. Toplam 40 kişi bu soruya cevap vererek konuyla ilgili görüşlerini belirtmiştir (Tablo 24). Görüş belirtenler arasında elde ettikleri veya ürettikleri verinin genellikle paylaşım ve tekrar kullanım için uygun olup olmadığı sorusuna evet cevabı veren 16 kişi (%40), kısmen cevabı veren 16 kişi (%40) ve hayır cevabı veren 8 kişidir (%20).

**Tablo 24.** Görüşlerdeki ifadelerin sıklığı

Yeniden yanlış yorumlanabileceği	Yeni araştırma için uygun olmaması	Kişisel verilerin gizliliği	Etik ihlal	Verinin hatalı eksik olması	Toplam
1	22	4	13	15	55

Görüş belirtenler ifadelerinde en sık geçen sebep (S=22) verilerin yeni araştırma için uygun olmayacağı sorunudur. Bunlar; değişkenlik gösterebileceği, güncelliğini kaybedeceği, verinin dar kapsamlı olması, çalışma planının doğru planlanmaması, deney formatı-anket-ölçek-örneklemin farklı olması, konuya özel olması, özgünlüğünü kaybetmesi, bayesli gerçekleştirilmiş olması, araştırmanın yanıltıcı olacağı, özel konulara yönelik yapılması, kesitsel/boylamsal araştırma olması ve geçerliliğinin yok olacağı şeklinde ifade edilmiştir.

Belirtilen görüşlerdeki ifadelerden 15 tanesi veriden kaynaklanan sorunlardır. Bunlar; standardize edilmemiş olması, basılı dokümanda tutuluyor oluşu, dağınık/düzensiz saklanması, güvenli olarak saklanmaması ihtimali, kullanılmış materyaller, veril ile ilgili yeterince bilgi olmaması, verilerin eksik ve hatalı olmasıdır. Görüşlerde 13 defa etik ihlal kabul edileceği ifadesi geçmektedir. Bunlar; etik ihlal, salamizasyon ve emek hırsızlığı olarak ifade edilmiştir. En az geçen ifadeler (S=4) kişisel verilerin gizliliği ile verilerin yanlış yorumlanabileceği (S=1) olmuştur.

Katılımcılardan araştırma verilerinin paylaşımı ile ilgili ifadelere katılım düzeylerini 1-en olumsuz, 5-en olumlu olacak şekilde işaretlemeleri istenmiştir. İfadelere verilen yanıtların güvenilirliği hesaplanmış ve ifadelerin iç tutarlılığını gösteren Cronbach's Alpha değerleri 0,668 olarak bulunmuştur. Bu ifadeler veri paylaşımı önündeki engeller, veri paylaşımı hakkındaki görüşler ve verinin yeniden kullanımı ile ilgili görüşler olmak üzere üç gruba ayrılarak tablolaştırılmış ve ifadelere verdikleri puanlar sıklık ve ortalama olarak Tablo 25'te gösterilmiştir.

**Tablo 25.** Veri Paylaşımı önündeki engeller

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	
Yeni yayın yapabileceğimi düşünerek verilerimi paylaşmak istemem	39	3,00	33	3,27	15	2,80	12	3,00	32	3,28	8	3,13	<b>3,12</b>
Verilerimi patent gibi ticari konular nedeni ile paylaşmak istemem	39	3,03	33	3,36	15	3,00	12	3,58	32	3,25	8	3,38	<b>3,22</b>
Araştırmacıların, verileri yanlış kullanabileceğini veya yorumlayabileceğini düşündüğüm için paylaşmak istemem	39	2,69	33	2,91	15	2,47	12	3,08	33	3,36	8	3,13	<b>2,94</b>
Verilerim kişisel ya da hassas veri içerdiğinden paylaşmak istemem	39	2,92	33	3,21	15	2,80	12	3,50	33	3,48	8	3,00	<b>3,16</b>
Verilerimi paylaşmak benim için ekstra zaman ve emek gerektirdiği için paylaşmak istemem	39	2,28	32	2,16	15	1,80	12	2,67	32	2,88	8	2,75	<b>2,40</b>
Verilerimi paylaşırsam, araştırma fikri ve süreçlerinin diğer araştırmacılar tarafından çalınacağını düşünüyorum	39	3,15	33	3,52	15	3,33	12	3,58	32	3,41	8	3,25	<b>3,36</b>

“Yeni yayın yapabileceğimi düşünerek verilerimi paylaşmak istemem” ifadesi (S=139,  $\bar{x}$  =3,12) ile “Verilerim kişisel ya da hassas veri içerdiğinden paylaşmak istemem” ifadesi (S=140,  $\bar{x}$ =3,16) için hesaplanan ortalamalar neredeyse eşittir. Bu ifadeler için katılımcıların kararsız olduğu söylenebilir. “Verilerimi patent gibi ticari konular nedeni ile paylaşmak istemem” ifadesi için genel ortalama 3,22 (S=140) hesaplanmıştır. Bu ifade için en düşük ortalama Eczacılık Fakültesi’ne ait olup fakültele göre istatistiksel olarak anlamlı bir farklılık göstermemektedir.

“Araştırmacıların, verileri yanlış kullanabileceğini veya yorumlayabileceğini düşündüğüm için paylaşmak istemem” ifadesi fakülteler arasında benzer dağılım göstermekte olup ortalaması 2,94 olarak hesaplanmıştır. “Verilerimi paylaşmak benim için ekstra zaman ve emek gerektireceğinden paylaşmak istemem” ifadesinin ortalaması 2,40 olup Eczacılık Fakültesi ( $\bar{x}$ =1,80) için ortalama değer belirgin düşüktür. Bu ifade fakültele göre istatistiksel olarak anlamlı farklılık göstermektedir ( $\chi^2$  (4)=42.569; p=0,002).

Katılımcıların “Verilerimi paylaşırsam, araştırma fikri ve süreçlerinin diğer araştırmacılar tarafından çalınacağını düşünüyorum” ifadesine katılım ortalaması 3,36 olarak hesaplanmış olup fakültele göre istatistiksel olarak anlamlı bir farklılık göstermemektedir.

Katılımcıların veri paylaşımı hakkındaki görüşler ile ilgili ifadelere verdikleri puanlar sıklık ve ortalama olarak Tablo 26’da gösterilmiştir. Tablodaki ifadeler arasında katılımcıların en olumlu (S=139,  $\bar{x}$ =4,00) düşündüğü ifade “verilerin paylaşımını ve doğru şekilde yeniden kullanımını düzenleyen politikalar olursa paylaşırım” olmuştur. Bu ifade için ortalamalar fakülteler arasında benzer olup istatistiksel olarak da anlamlı bir farklılık göstermemektedir. “Verilerimi, akademik açıdan ödül, puan veya atıf almam halinde paylaşabilirim” ifadesi için Sağlık Bilimleri Fakültesi’nin ortalaması diğerlerinden daha düşük ( $\bar{x}$ =2,79) olmasına rağmen istatistiksel olarak anlamlı farklılık göstermemektedir.

“Veri ile işim tamamen bittikten sonra paylaşabilirim” ifadesi ( $\bar{x}$ =3,73) ile “Verilerimi güvenilir bir veri arşivi ya da deposu bulursam paylaşabilirim” ifadesi ( $\bar{x}$ =3,58) ortalamaları neredeyse eşittir.

**Tablo 26.** Veri Paylaşımı hakkındaki görüşler

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	$\bar{x}$
Verilerin paylaşımını ve doğru şekilde yeniden kullanımını düzenleyen politikalar olursa paylaşırım	39	4,13	33	3,94	15	4,20	12	3,75	32	3,91	8	4,00	<b>4,00</b>
Verilerimi, akademik açıdan ödül/puan veya atf almam halinde paylaşabilirim	39	3,15	33	2,79	15	3,20	12	3,33	32	3,19	8	3,25	<b>3,10</b>
Veri ile işim tamamen bittikten sonra paylaşabilirim	39	4,00	33	3,73	15	3,87	12	3,75	32	3,37	8	3,63	<b>3,73</b>
Verilerimi güvenilir bir veri arşivi ya da deposu bulursam paylaşabilirim	39	3,77	32	3,47	15	3,60	12	3,08	32	3,63	8	3,63	<b>3,58</b>

Katılımcıların verinin yeniden kullanımı ile ilgili ifadelerine verdikleri puanlar sıklık ve ortalama olarak Tablo 27’de gösterilmiştir. “*Araştırma verilerinin analizinden elde edilen sonuçların yayınlar aracılığı ile paylaşılmasını yeterli buluyorum*” ifadesi ( $\bar{x}=3,70$ ) ile “*Araştırma verilerinin paylaşılmasının önemli olduğunu düşünüyorum*” ifadesinin ( $\bar{x}=3,65$ ) ortalamaları neredeyse eşittir. İki ifade aslında karşıt düşünceleri ifade ediyor olmasına rağmen ortalamalardaki farklılık tutarlıdır.

“*Araştırma verilerinin analizinden elde edilen sonuçların yayınlar aracılığı ile paylaşılmasını yeterli buluyorum*” ifadesi fakültelelere göre istatistiksel olarak anlamlı farklılık göstermektedir ( $\chi^2(4) = 31,433$ ;  $p=0,020$ ). Katılımcıların en olumlu düşündüğü ifadenin “*Araştırmamda kullandığım veriler için atf yapmaktan çekinmem*” ifadesi ( $S=139$ ,  $\bar{x}=4,34$ ) olduğu söylenebilir. Katılımcıların en olumsuz görüş belirttiği ifade ise “*Diğer araştırmacıların paylaştığı verileri araştırmalarımında kullanmaktan çekinmem*” ( $S=139$ ,  $\bar{x}=2,87$ ) olmuştur.

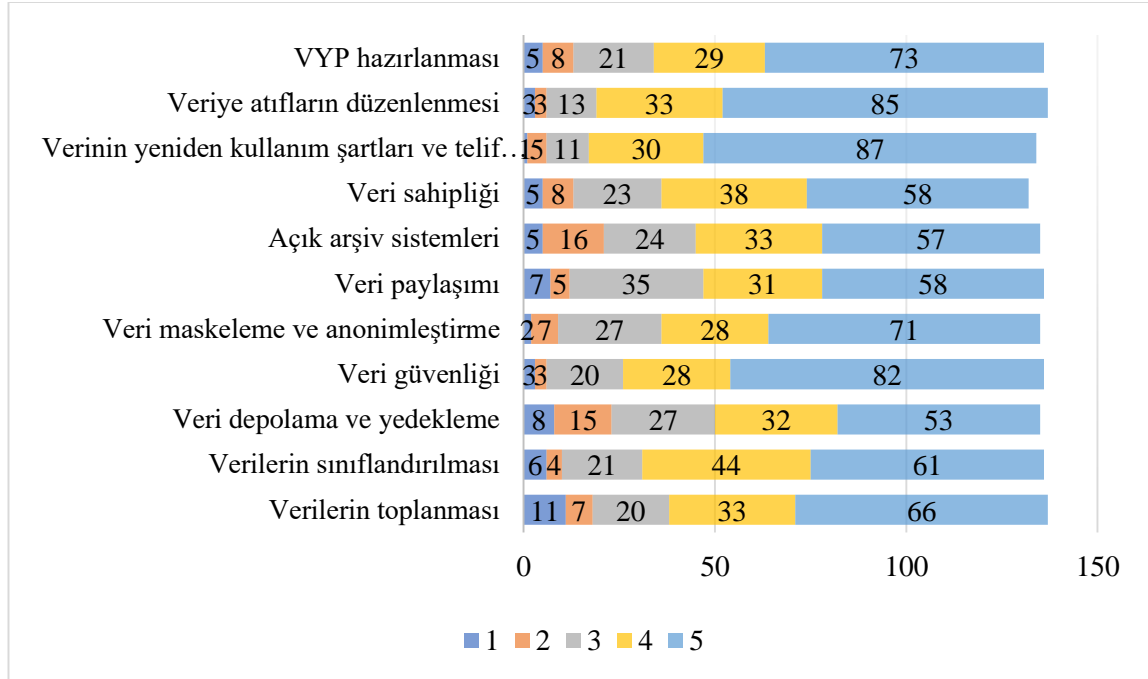


**Tablo 27.** Verilerin yeniden kullanımı ile ilgili görüşler

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	$\bar{x}$
Araştırma verilerinin analizinden elde edilen sonuçların yayınlar aracılığı ile paylaşılmasını yeterli buluyorum	38	3,08	33	3,61	15	3,80	12	4,00	32	4,31	8	3,88	<b>3,70</b>
Araştırma verilerinin paylaşılmasının önemli olduğunu düşünüyorum	38	4,00	32	3,31	15	4,07	12	3,25	32	3,59	8	3,38	<b>3,65</b>
Diğer araştırmacıların paylaştığı verileri araştırmalarımda kullanmaktan çekinmem	39	3,03	33	2,70	15	3,20	12	2,83	32	2,69	8	3,00	<b>2,87</b>
Araştırmamda kullandığım veriler için atıf yapmaktan çekinmem	39	4,46	33	4,12	15	4,67	12	4,42	32	4,31	8	4,00	<b>4,34</b>

#### 4.1.4. Katılımcıların Eğitim İhtiyaçları

Katılımcıların eğitim ihtiyaçlarını ortaya koymak amacıyla, veri okuryazarlığı ve araştırma verilerinin yönetimi eğitimi alacak olsalardı öncelikli konuların neler olmasını istedikleri sorulmuştur. Katılımcılardan konuları 1-en az öncelikli, 5-en çok öncelikli olacak şekilde işaretlemeleri istenmiştir. Katılımcıların ifadelerine verdikleri yanıtların güvenilirliği hesaplanmıştır ve ifadelerin iç tutarlılığını gösteren Cronbach's Alpha değerleri 0,878 olarak bulunmuştur. Katılımcılardan 61 (%44) ile 87 (%62) kişi tüm konularda öncelikli eğitim ihtiyacı olduğunu belirtmiştir. Araştırma grubunun genel olarak eğitim ihtiyacının yüksek olduğunu belirttiği söylenebilir (bkz. Tablo 28 ve Tablo 29).

**Tablo 28.** Katılımcıların eğitim ihtiyaçları

Katılımcılar için en öncelikli konu “*Verinin yeniden kullanım şartları ve telif hakları*” olmuştur (S=134,  $\bar{x}$ =4,47). Belirtilen konu için tüm fakültelerin ortalamaları 4’ün üzerinde hesaplanmıştır. 6 katılımcı (%4,5) az öncelikli konu olarak işaretlemiştir. Ortalaması en yüksek olan ikinci konu “*Veriye atıfların düzenlenmesi*” olmuştur (S=137,  $\bar{x}$ =4,43). İlk konuya benzer dağılım ve neredeyse eşit ortalama hesaplanmıştır. “*Veri güvenliği*” konusu için 136 katılımcı işaretleme yapmış olup ortalaması 4,33 olarak hesaplanmıştır. İlk iki konuya benzer şekilde, tüm fakülteler için ortalamaları 4’ün üzerinde hesaplanmıştır.

**Tablo 29.** Katılımcıların eğitim ihtiyaçları 2

	Tıp		Sağlık Bilimleri		Eczacılık		Hemşirelik		FTR		Spor Bilimleri		Genel
	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	S	$\bar{x}$	$\bar{x}$
Verilerin toplanması	38	3,89	33	3,82	14	4,29	12	<b>4,42</b>	32	3,94	8	<b>4,25</b>	<b>3,98</b>
Verilerin sınıflandırılması	37	3,95	37	3,95	14	4,21	12	<b>4,50</b>	32	<b>4,09</b>	8	<b>4,13</b>	<b>4,09</b>
Veri depolama ve yedekleme	37	3,76	37	3,76	14	3,79	12	<b>4,25</b>	31	3,74	8	3,87	<b>3,78</b>
Veri güvenliği	38	<b>4,16</b>	33	<b>4,21</b>	14	<b>4,64</b>	12	<b>4,83</b>	31	<b>4,42</b>	8	<b>4,25</b>	<b>4,33</b>
Veri maskeleye ve anonimleştirme	37	<b>4,30</b>	33	<b>4,21</b>	14	<b>4,64</b>	12	<b>4,83</b>	31	<b>4,42</b>	8	<b>4,25</b>	<b>4,21</b>

Veri paylaşımı	38	3,97	33	<b>4,06</b>	14	<b>4,00</b>	12	3,92	31	3,81	8	3,75	<b>3,96</b>
Açık arşiv sistemleri	38	3,95	33	<b>4,00</b>	14	<b>4,07</b>	12	<b>4,00</b>	30	3,60	8	3,88	<b>3,92</b>
Veri sahipliği	36	<b>4,08</b>	33	<b>4,12</b>	13	<b>4,23</b>	12	<b>4,17</b>	30	3,73	8	<b>4,00</b>	<b>4,08</b>
Verinin yeniden kullanım şartları ve telif hakları	37	<b>4,62</b>	33	<b>4,45</b>	14	<b>4,43</b>	12	<b>4,50</b>	30	<b>4,37</b>	8	<b>4,25</b>	<b>4,47</b>
Veriye atıfların düzenlenmesi	38	<b>4,53</b>	33	<b>4,33</b>	14	<b>4,36</b>	12	<b>4,67</b>	32	<b>4,38</b>	8	<b>4,13</b>	<b>4,43</b>
VYP hazırlanması	38	<b>4,26</b>	33	<b>4,15</b>	14	<b>4,14</b>	12	<b>4,08</b>	31	<b>4,06</b>	8	<b>4,13</b>	<b>4,17</b>

Katılımcılar “*Veri maskeleye ve anonimleştirme*” konusu ( $S=135$ ,  $\bar{x}=4,21$ ) ile “*VYP hazırlanması*” konusu ( $S=136$ ,  $\bar{x}=4,17$ ) için neredeyse eşit işaretleme yapmıştır. “*Verilerin sınıflandırılması*” ( $S=13$ ,  $\bar{x}=4,09$ ) ve “*Veri sahipliği*” ( $S=132$ ,  $\bar{x}=4,08$ ) konuları ortalamaları neredeyse eşit olmalarına rağmen fakültelerin ortalamaları arasında istatistiksel olarak anlamlı olmayan fark vardır. *Verilerin sınıflandırılması* konusu için hemşirelik alanı ( $S=12$ ,  $\bar{x}=4,50$ ) ortalaması diğer alanlardan yüksek, “*Veri sahipliği*” konusunda ise FTR alanının ( $S=30$ ,  $\bar{x}=3,73$ ) ortalaması diğer alanlardan oldukça düşüktür. “*Verilerin toplanması*” ( $S=137$ ,  $\bar{x}=3,98$ ), “*Veri paylaşımı*” ( $S=136$ ,  $\bar{x}=3,96$ ) ve “*Açık arşiv sistemleri*” ( $S=135$ ,  $\bar{x}=3,92$ ) konuları için eğitim ihtiyacı eşittir denebilir. Katılımcıların en az eğitime ihtiyaç duyduğunu belirttiği “*Veri depolama ve yedekleme*” konusunda ( $S=135$ ,  $\bar{x}=3,78$ ) hemşirelik alanında istatistiksel olarak anlamlı olmayan fark vardır.

#### 4.1.5. Katılımcıların Görüşleri

Anketin sonunda katılımcılara dilerlerse araştırma ile ilgili görüş ve önerilerini belirtebilecekler bir alan bırakılmıştır. Toplam 20 kişi bu bölümde görüş bildirmiştir. Bu katılımcılardan 10 kişi başarılar dileyerek iyi dileklerini iletmışlerdir. 4 katılımcı çalışmayı güzel, faydalı, ilginç, başarılı ve değerli bulduğunu belirten ifadeler kullanmışlardır. Bir katılımcı konu ile ilgili çok bilgisi olmadığını belirtirken, bir diğer katılımcı başka dil seçenekleri olmasının iyi olacağını belirtmiştir. Diğer katılımcıların görüşleri ise şöyledir;

- *Yeni arařtırmalarda, alıřma grubundaki verilerin gvenilir ve sađlıklı olması sreci hızlandıracaktır. Ve tekrar tekrar benzer konularda rneklem bulunmasına gerek duyulmayacaktır. Sađlıklı bir depo ile alıřmalar hızlanabilir. Teřekkrler.*
- *Yapılması gerekli olan bir arařtırma. Arařtırma sonuları ile ilgili lisansst đrencileri ve akademisyene gerekli bilgi, eđitim ve geri dntlerin yapılması gerektiđini dřnyorum.*
- *Veri paylařımı konusunda genelde ekimser kalıyoruz. Bu konulara farkındalık sađlamak aısından faydalı olacađı dřncesindeyim. Őimdiden elinize sađlık.*
- *Lisansst eđitime yeni bařladıđım iin anketle ilgili bazı sorularda ok fikir sahibi olamadıđım iin etkin bir cevaplama yapamamıř olabilirim. Yksek lisans tez dnemi đrencilerinde ve doktora đrencilerinde daha etkili olacađını dřnyorum. Kolaylıklar dilerim.*

#### **4.2. DERS İERİKLERİ ANALİZİ**

Öđrencilerin mfredatlarında arařtırma verilerinin ynetimi ve uygulamaları, veri paylařımı, veri depoları kavramlarıyla ilgili bilgi sahibi olacakları ieriklerin olup olmadıđının ortaya koyulması amacıyla ders ierikleri tematik analize tabi tutulmuřtur. Hacettepe niversitesi Sađlık Bilimleri Enstits'ne bađlı tm programlarda verilen dersler [akts.hacettepe.edu.tr](http://akts.hacettepe.edu.tr) adresi zerinden incelenerek bilimsel bir arařtırmayı planlama, yrtme ve raporlama becerilerini geliřtirmeyi hedefleyen dersler ile veriyle ilgili alıřmaların yapıldıđı dersler zerine odaklanılmıřtır. Bazı blmlerde bu ieriklerde ders olmadıđı gibi bazı blmlerde zorunlu ve semeli olarak birden fazla ders bulunmaktadır. Ders programında bu derslere ait ieriklerin yayımlandıđı toplam 128 blm vardır. Bu blmlerde belirtilen ieriklerde verilen toplam 73' semeli, 107'si zorunlu olmak zere toplam 180 ders vardır. Yksek lisans ve doktora programlarında farklı blmlerde ortak olarak verilen aynı kodlu dersler veri setinde tekrar olmaması iin bir ders olarak sayıldıktan sonra toplam 42 ders sayısına ulařılmıřtır. Analiz 42 ders zerinden yapılmıř ancak deđerlendirme, belirlenen ieriklerde dersleri olan 128 blm

üzerinden yapılmıştır. Bölümlerde ortak olarak en çok verilen iki ders Tablo 30'da, nitel araştırmanın veri seti ile ilgili genel bilgiler ise Tablo 31'de gösterilmiştir.

**Tablo 30.** Ortak verilen dersler

	Z	S	Toplam
SBE601 Araştırma Yöntemleri, Araştırma ve Yayın Etiği	64	13	77
BİS605 Biyoistatistik	24	7	31

**Tablo 31.** Nitel araştırmanın veri seti ile ilgili genel bilgiler

Dersin Kodu Ve Adı	YL/D	Z/S
BDD734 - Seminer	Z	D
BES605 - Beslenme ve Diyetetik Alanında Araştırma Yöntemleri	Z	YL
BİF661 - Biyoinformatiğe Giriş	S	YL
BİF681 - Biyolojik Veri Tabanları	S	YL
BİF686 - Uygulamalı Veri Madenciliği	S	YL
BIN704 - Araştırma Yöntemleri, Araştırma ve Yayın Etiği	Z	D
BIN 781 Biyolojik Veri Tabanları ve Veri Analiz Araçları	S	D
BIN783 - Genetik Araştırmalarda İleri İstatistik Yöntemler	S	D
BİS604 - Sağlık Bilimlerinde Araştırma Yöntemleri	S	YL
BİS605 - Biyoistatistik	S	YL
BİS662 - Sağlıkta Veri Bilimi	S	YL
BİS735 - Biyoistatistik	Z	D
BİS736 - Sağlık Bilimlerinde Araştırma Yöntemleri	S	D
BİS786 - Veri Madenciliği	S	D
BSP613 - Spor Bilimlerinde Nitel Araştırmanın Temelleri	S	YL
BYF601 - Biyofizikte Araştırma Yöntemleri	Z	YL
ÇGE652 - Çocuk Gelişimi ve Eğitiminde Araştırma Yöntem ve Teknik Uyg.	Z	YL
ÇGE752 - Çocuk Gelişimi ve Eğitiminde Araştırma Yöntem ve Teknik Uyg	Z	D
ÇOC610 - Dönem Projesi	Z	YL
DEN711 - Endodontide Araştırma Yöntemleri	Z	D
EFG623 - Eczacılıkta Literatür Tarama Teknikleri	Z	YL
ERG672 - Nitel Araştırma Yöntemleri	S	YL
HCH608 - Çevre Epidemiyolojisinde Gözlemsel Araştırma Yöntemleri	S	YL
HEP602 - Epidemiyolojinin Temel İlke ve Yöntemleri	Z	YL

HEP603 - Epidemiyolojide Araştırma Planlama ve Uygulama	S	YL
HEP722 - Sağlık Araştırmalarında Bilimsel Yazım	S	D
HHS611 - Hemşirelikte Araştırma Yöntemleri, Araştırma ve Yayın Etiği I	Z	YL
HHS717 - Hemşirelikte Araştırma Yöntemleri, Araştırma ve Yayın Etiği II		D
HİH708 - Bilimsel Makale Yazma	S	D
HİS787 - Histoloji ve Embriyolojide Araştırma Yöntemleri	Z	D
HSP704 - Çocuk Sağlığında Araştırma	S	D
KLE706 – Klinik Eczacılıkta Araştırma Yöntemleri	Z	D
MEB619 - Makale Hazırlama, Araştırma Planlama ve Yayın Etiği	S	YL
OFR714 - OFR'de Proje Araştırma, Geliştirme ve Yönetimi	S	D
SBE601 - Araştırma Yöntemleri, Araştırma ve Yayın Etiği	Z	YL
SBT720 - Spor Bilimlerinde Nitel Araştırma	S	D
SİM603 - Simülasyona Dayalı Araştırmalar ve Etik	Z	YL
TAS702 - TADS Epidemiyolojisinin İlke ve Yöntemleri	S	D
TAS711 - TAS Epidemiyolojisinde Araştırma Planlama Uygulama	S	D
TEB659 - Tıp Eğitiminde Araştırmanın Temelleri	Z	YL
TEB711 - Tıp Eğitiminde Araştırma	S	D
TNB752 - Nörolojik Bilimlerde Proje Sunumu ve Veri Analiz	Z	D

SBE601 Araştırma Yöntemleri, Araştırma ve Yayın Etiği dersi 64 bölümde zorunlu, 13 bölümde ise seçmeli olarak okutulmaktadır ve toplam 77 bölümden 15 tanesi doktora programıdır. BİS605 Biyoistatistik dersi 24 bölümde zorunlu, 7 bölümde ise seçmeli olarak okutulmaktadır. Ders içerikleri MAXQDA 2022 programında belirlenen temalara göre kodlanmıştır. Kodlama sonucu oluşan kod kitabı Tablo 32’de görülmektedir.

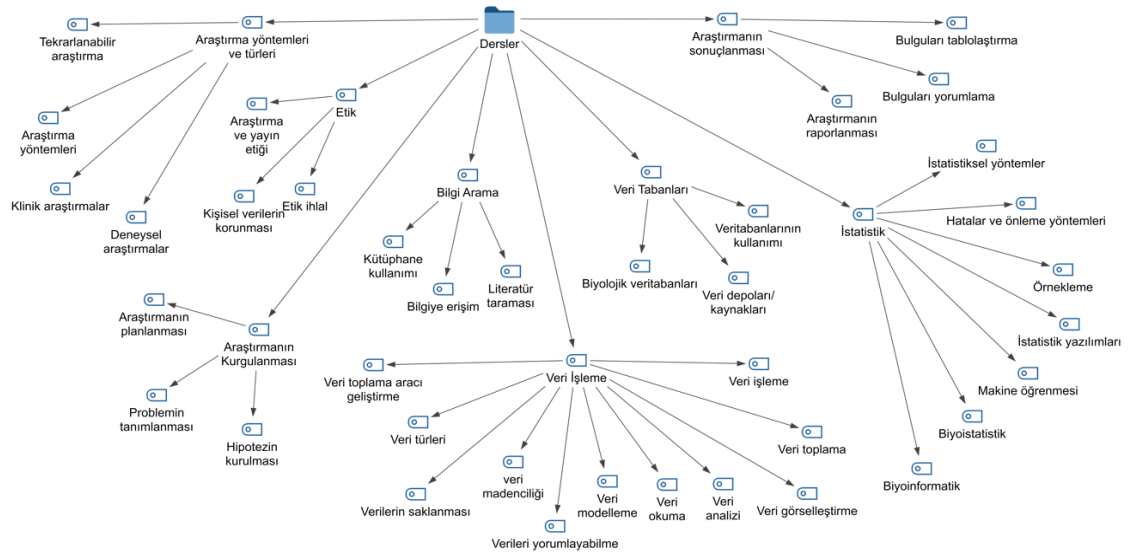
**Tablo 32.** Kod kitabı

Temalar	Kodlar	Sıklık
Araştırma yöntemleri ve türleri	Araştırma yöntemleri	30
	Tekrarlanabilir araştırma	1
	Deneysel araştırmalar	7
	Klinik araştırmalar	2
Etik	Etik ihlal	3
	Araştırma ve yayın etiği	19
	Kişisel verilerin korunması	1
Araştırmanın kurgulanması	Problemin tanımlanması	16

	Araştırmanın planlanması	26	
	Hipotezin kurulması	13	
Literatür tarama	Bilgiye erişim	8	
	Kütüphane kullanımı	2	
	Literatür taraması	15	
Veri Tabanları	Biyolojik veri tabanları	2	
	Veri depoları/kaynakları	7	
	Veri tabanlarının kullanımı	6	
Veri	Veri toplama aracı geliştirme	5	
	Veri toplama	21	
	Verilerin saklanması	3	
	Veri türleri	5	
	Veri madenciliği	4	
	Verileri yorumlayabilme	5	
	Veri okuma	2	
	Veri modelleme	1	
	Veri görselleştirme	3	
	Veri analizi	23	
	Veri işleme	8	
	İstatistik	İstatistik yazılımları	15
		Hatalar ve önleme yöntemleri	9
Biyoistatistik		2	
Biyoinformatik		1	
Makine öğrenmesi		2	
İstatistiksel yöntemler		13	
Örnekleme		13	
Araştırmanın raporlanması	Araştırmanın raporlanması	26	
	Bulguları yorumlama	18	
	Bulguları tablolaştırma	4	

Ders içeriklerinin analizi sonucunda belirlenen dersler için en çok geçen ifadeler araştırma yöntemleri (30), araştırmanın planlanması (28), veri toplama (21), veri analizi (23) ve araştırmanın raporlanması (26) olmuştur. Bunun sebebi bu kodların belirttiği süreçlerin bilimsel bir araştırma yapmanın temel adımları olması ve ders içeriklerinde daha genel ifadelerin kullanılması olabilir. Ders içerikleri kod haritası Şekil 4'te görülmektedir.

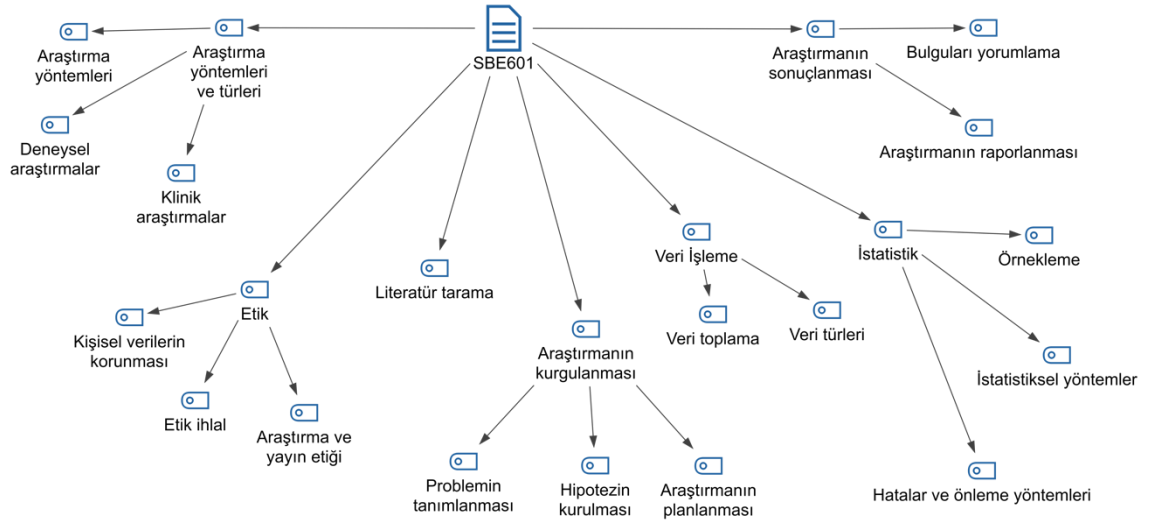
**Şekil 4.** Tüm dersler Tek vaka modeli kod haritası



Veri temasının kodlarına bakıldığında en çok geçen ifadeler veri toplama ve veri analizidir. Çalışmaya temel olan araştırma verilerinin yönetimi, veri paylaşımı, yedekleme, lisanslama, veriye atıf konuları analiz edilen derslerin hiçbirinde geçmemektedir. Veri saklama ifadesinin geçtiği üç ders içeriği bulunmaktadır. Literatür tarama temasının içerdiği kodların ders içeriklerinde sık geçmemesi bu konulara derinlemesine değinilmediğini göstermektedir. Veri tabanları temasının kodları yine ders içeriklerinde az geçen ifadeler olmuştur. Araştırma ve yayın etiği teması derslerin yarısının içeriğinde mevcuttur. En çok kodlanan ifadelere bakıldığında araştırma yöntemleri, araştırma etiği, araştırmanın planlanması, problemin tanımlanması, veri toplama, verinin analizi ve sonuçların raporlanması konuları olarak bir araştırmaya başlanmasından sonlandırılmasına kadar olan süreçlerin ifadesi olduğu görülmektedir. Sağlık Bilimleri Enstitüsü'ne bağlı programlarda ortak olarak en çok verilen ders olan "SBE601 Araştırma Yöntemleri, Araştırma ve Yayın Etiği" ders içeriğinin kod haritası Şekil 5'te gösterilmiştir.



Şekil 5. SBE601 dersi kod haritası



İlgili dersin içeriğine bakıldığında araştırma becerileri bakımından en zengin içeriklerden biri olduğu söylenebilir. Ancak veri ile ilgili kısım sadece veri türleri ve veri toplama konularıdır. Belirtilen dersin içeriğinde veriyi anlama, yorumlama, değerlendirme, düzenleme, veriyi yönetme gibi veri okuryazarlığı becerilerini geliştirecek içerikler bulunmamaktadır.

## 5. BÖLÜM

### SONUÇ VE ÖNERİLER

Araştırma verilerinin paylaşılması klinik araştırmalardan elde edilen sonuçlara duyulan güveni artırmada ve bilimsel sürecin temel ilkelerinden biri olan sonuçların bağımsız olarak doğrulanmasında önemli bir yere sahiptir. Veri yönetimi ve paylaşımı iyi yapıldığında, klinik araştırma verilerinin paylaşılması, her bir araştırmadan maksimum fayda sağlanmasına katkıda bulunmaktadır. Genç araştırmacılara müfredat içinde veri okuryazarlığı ve araştırma verilerinin yönetimi eğitimi verilmesi veya farklı kanallardan çeşitli eğitim imkanlarının sunulması, araştırmacıların iyi uygulamalar gerçekleştirebilmesinde önemli adımlardan biridir. Araştırmacıların konu ile ilgili farkındalıklarının değerlendirilmesi, mevcut alışkanlıklarının ve ihtiyaçlarının belirlenmesi ise öncelikli olarak düşünülmelidir. Bu kapsamda Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü'nde gerçekleştirilen araştırmanın sonuçları, araştırma sorularına bulunan yanıtlar ve hipotezler sonuç bölümünde değerlendirilmiştir. Değerlendirmeler sonucunda ulusal düzeyde, üniversite olarak ve araştırmacılar tarafından araştırma verilerinin yönetimi konusunda gelişme sağlanabilmesi için önerilerde bulunulmuştur.

#### 5.1. SONUÇ

Ankete katılan 140 katılımcı bölümlerinin bağlı olduğu fakültele göre ayrılarak değerlendirilmiştir. Diş Hekimliği Fakültesi'ne bağlı bölümlerden katılım olmamıştır. Yüksek lisans ve doktora programına kayıtlı öğrenci sayısı eşit olup ders veya tez döneminde olma durumları da benzerlik göstermektedir. Ankete katılanların yarısından çoğu 20 ile 40 yaş arasındadır. Katılımcıların büyük çoğunluğunun genç öğrenciler olması ve çalıştığını belirtenlerin çoğunluğunun araştırmacı olması, bu çalışma için hedef kitleye ulaşılabildiğini göstermektedir.

Katılımcıların bilgi arama davranışlarına bakıldığında fakülte farketmeksizin öncelikli tercihlerinin Google Scholar vb. akademik internet platformları olduğunu belirtmişlerdir.

Sonraki tercihleri sağlık konusunda uzman veri tabanları ve elektronik kütüphane kaynakları olmuştur. Ders içerikleri analizinde ortaya çıkan sonuçlarda da literatür taraması temasında bulunan kodlara ait konulara sıkça yer verildiği görülmektedir. Araştırmacılar literatür taraması yaparken alana özgü farklı kaynakları da kullanmaktadırlar. Bilgiye erişimde elektronik kaynaklar kullanmaları ve kütüphane kullanımının çok düşük seviyelerde kalması dikkat çekmektedir. Bunun sebepleri en güncel bilgiye hızlıca ulaşma ihtiyaçlarından dolayı basılı kaynakları öncelikli olarak tercih etmemeleri ve pandemi dolayısıyla araştırma imkanlarının da kısıtlanarak internet ortamında yoğunlaşması olabilir. Elde edilen bu bulgu Köse'nin (2011) tıp akademisyenleri üzerinde gerçekleştirdiği çalışmada elde ettiği bulgular ile benzerlik göstermektedir. Çalışmada katılımcıların %71'i bilgiye erişimde en çok elektronik kaynakları kullanmayı tercih ettiklerini belirtmişlerdir. Kullandıkları kaynaklar sorulduğunda ise %60,1'i dergileri, %60,1'i konuyla ilgili web sitelerini, %44,3'ü ise kitapları kullandıklarını belirtmişlerdir.

Araştırmacıların bilgiyi internetten arama yaklaşımları fakültelere göre benzer dağılım göstermektedir. Araştırmacıların tümü çoğu zaman araştırmaya başlamadan önce araştırma planı yapmaktadır ancak veri yönetim planı yapmamaktadırlar. Tıp araştırmacıları aradığı konuyu doğrudan Google, Yandex vb arama motorlarına yazıp ilk gelen sayfaları inceleme yaklaşımını diğer fakültelere oranla daha çok tercih etmektedirler. Ancak araştırmacıların tamamı çoğu zaman bağlaçları kullanarak anahtar kelimeleri ilişkilendirip arama yaptığını, bilgiye ulaşamadığında ise her zaman anahtar kelimeleri değiştirerek tekrar arama yaptığını belirtmiştir. Katılımcılar her zaman bilgiyi farklı kaynaklardan teyit etmekte olup bunun için de çoğu zaman Sağlık Bakanlığı, TÜİK, Kalkınma Bakanlığı gibi otorite kurumların web sayfalarını ve konu uzmanı otorite kaynak, kitap vb kanalları tercih etmektedirler. Köse'nin (2011) çalışmasında da benzer şekilde katılımcıların %72,1'i aradıkları bilginin doğruluğunu önemsediklerini belirtmişlerdir. Ders içeriklerine bakıldığında da literatür tarama temasına ait kodların sıkça geçmesi anket bulgularını desteklemektedir. Elde edilen bulgulara göre “Bilgi arama yaklaşımları fakültelere göre farklılık göstermektedir” hipotezi reddedilmiştir. Katılımcılar doğru bilgiye erişimi önemsemektedirler. Bilgiyi doğru kaynaktan edinmekte ve eriştikleri bilgiyi de farklı kaynaklardan teyit etmektedirler.

Araştırmacıların veri elde etme biçimleri istatistiksel olarak anlamlı olmayan farklılıklar göstermektedir. Tıp araştırmacıları ile hemşirelik araştırmacıları daha çok araştırma yapacağı kurumun veri tabanı ile basılı kaynaklar veya arşiv belgeleri üzerinde araştırma yaparak veri setini elde ederken, sağlık bilimleri, fizik tedavi ve rehabilitasyon ile spor bilimleri araştırmacıları en çok anket/ölçek gibi yöntemlerle veri toplamayı tercih etmektedir. Bu sonuç Krahe ve diğerlerinin (2020, s. 111) çalışmasındaki sonuçlarla benzerlik göstermektedir. Eczacılık alanındaki araştırmacıların en çok tercih ettiği yöntem basılı kaynaklar veya arşiv belgeleri üzerinde araştırma yaparak veri toplamaktır. Diğer araştırmacıların veri depolarındaki verilerini kullanmayı en çok tercih edenler hemşirelik araştırmacılarıdır. Veri toplama konusunda en öncelikli eğitim ihtiyacı olduğunu belirten araştırmacılar da yine hemşirelik araştırmacılarıdır. Bunun sebebi veri depoları ve açık veri hakkında daha çok bilgi sahibi olmak istemeleri olabilir. Bununla birlikte tüm araştırmacılar veri elde etmek için tek bir yaklaşıma bağlı kalmadan diğer yöntemleri de kullanmaktadırlar. Elde edilen bu bulgu ile “Veri elde etme biçimleri fakültelere göre farklılık göstermektedir” hipotezi reddedilmiştir. Analiz edilen ders içeriklerinin yarısında veri toplama konusu geçmektedir. Programların çoğunluğunda öğrencilerin veri toplama konusunda eğitim almalarına rağmen eğitime ihtiyaçları olduğunu belirtmeleri çalışma kapsamında uygulanan anket sonrasında araştırma verilerinin yeniden kullanımı ile ilgili sağlanan farkındalık ve veri depolarındaki verilerin kullanımına merak duymaları olabilir.

Araştırmacıların yaklaşık yarısı bilgi arama davranışlarını lisansüstü eğitim sırasında derslerde öğrenmiştir. Ders içeriklerinin analizine bakıldığında bilgiye erişim, literatür taraması, kütüphane kullanımı ve veri tabanlarının kullanımı konularının sıkça geçtiği görülmektedir. İki analiz birbirini doğrular niteliktedir. Eşit sayıda araştırmacı da bilgi arama davranışlarını kendi kendine öğrendiğini belirtmiştir. Knowles (1975), kendi kendine öğrenmeyi "bireylerin, başkalarının yardımıyla veya yardımı olmaksızın, öğrenme ihtiyaçlarını belirleme, hedefleri formüle etme, insan ve materyal kaynaklarını belirleme, uygun öğrenme stratejilerini seçme, uygulama ve öğrenme sonuçlarını değerlendirme konusunda inisiyatif aldıkları bir süreç" olarak tanımlamıştır. Kendi kendine öğrenme, yaşam boyu öğrenmenin önemli bir bileşeni olarak kabul edilir ve bu nedenle tıp fakültesi müfredatında kilit bir yetkinliktir (Eva ve diğerleri, 2004, 530). Sawatsky ve arkadaşlarına göre (2018) bu öğrenen güdümlü bir süreç olsa da, öğretim

üyeleri ve akranlar da geri bildirimde bulunarak öğrenenin yönlendirilmesinde ve desteklenmesinde önemli bir rol oynar. Katılımcıların bu dersler aracılığıyla kendi kendine öğrenmede yönlendirme ve geri bildirim desteği aldıkları ve öğrenmeyi gerçekleştirdikleri görülmektedir. Fakültelere göre benzer dağılım gösteren küçük bir grup araştırmacı bilgi arama davranışlarını birlikte çalıştığı arkadaşından öğrendiğini belirtmiştir. Yalnızca üç kişi kütüphaneden eğitim aldığını belirtmiştir. Küçük bir grup da olsa lisans eğitimi sırasında derslerde bilgi arama davranışlarıyla ilgili eğitim alan araştırmacılar vardır. YÖK Lisansüstü Eğitim ve Öğretim Yönetmeliği'nin 34. maddesinin 5. bendinde “*Bilimsel araştırma teknikleri ile araştırma ve yayın etiği konularını içeren en az bir dersin lisansüstü eğitim sırasında verilmesi zorunludur*” ifadesi yer almaktadır. Araştırmaya katılan öğrencilerin bir kısmının programda var olan dersi henüz almamış olma ihtimali bulunmaktadır.

Araştırmacıların %80'den fazlası kullandıkları veri türü için birden fazla seçeneği işaretlemiştir. Araştırmacıların çoğunluğu anket/ölçek vb araçlarla oluşturulan veri ile deneysel veri, yarısı gözleme dayalı veri, yaklaşık üçte biri içerik analizi gibi yöntemlerle oluşturulan veri kullanmaktadırlar. Yarısından fazlası en çok kullanılan bu üç veri türünü de kullanmaktadır. Katılımcıların yarıdan fazlasının nitel araştırmalar yaptığını söyleyebiliriz. Görüntü kaydı, ses kaydı ve veri modellerine ve simülasyona dayalı veri katılımcıların en az kullandığı veri türleridir. Tıp ve sağlık bilimleri alanında deneysel veri kullanımı Doğan, Taşkın ve Aydınoglu'nun (2020, s. 55) araştırması ile benzerlik gösterirken anket verileri ile gözleme dayalı veriler için farklı sonuçlar çıkmıştır. Zencir'in (2019) çalışmasında da en çok kullanılan veri türü deneysel veri iken bu çalışmada anket/ölçek vb araçlarla oluşturulan veri olarak ortaya koyulmuştur. Araştırmacıların nitel çalışmaları nicel çalışmalardan daha çok tercih etmeleri kullandıkları veri formatlarını da etkilemektedir. Katılımcıların çoğu metin dosyası (.txt, .pdf vb.), üçte biri basılı doküman kullanmaktadır. İstatistiksel veri dosyası kullananlar yarıdan fazladır. Katılımcıların üçte birinden fazlası ise hesap işlem tablosu (.xls, .xlsx vb.) kullanmaktadırlar. Diğer veri formatlarının kullanımı düşük seviyelerde kalmıştır.

Araştırmacılar verilerini uzun süreli olarak, çoğunlukla sürekli olarak saklamaktadırlar. Araştırmacılar verilerini farklı bir ortamda yedeklemekte ve verilerini saklarken veya yedeklerken tekrar erişimi kolaylaştırmak için içeriği tanımlayıcı üst veri ile

tanımlamaktadırlar. Üst veri ile tanımlayan katılımcıların neredeyse tamamının tanımlayıcı üst veri kullanıyor olması ancak standartlarla ilgili bilgi sahibi olmaması, yalnızca bir araştırmacının üst veri standardı kullandığını belirtmesi araştırmacıların bu konuda da bilgi eksikliği olduğunu göstermektedir. Araştırmacılar araştırma süresince verilerini en çok kişisel bilgisayarlarında, harici sürücülerde ve bulut uygulamalarda saklamakta ve yedeklemektedirler. Araştırma tamamlandıktan sonra da verilerini saklama ve yedekleme ortamları değişiklik göstermemektedir. Bu bulgular diğer araştırma sonuçları ile benzerlik göstermektedir (Doğan, Taşkın ve Aydınoglu, 2020, s. 57). Katılımcılar veri saklama ve veri yedekleme ilgili benzer davranışlar sergilemektedirler. Elde edilen bu bulgu ile “Verilerini saklama ve depolama alışkanlıkları fakültelere göre farklılık göstermektedir” hipotezi reddedilmiştir. Az sayıda katılımcı araştırma tamamlandıktan sonra verilerini veri arşivi veya üniversite sunucularında saklamaktadır. Araştırmacıların neredeyse tamamının verileri paylaşılmamak üzere kişisel bilgisayar, harici bellek veya bulut uygulamalarda saklanmaktadır.

Az sayıda araştırmacı araştırmasına ait veriyi başka araştırmacılarla paylaşmıştır. Ancak daha fazla sayıda araştırmacı başka araştırmacıya ait veriyi araştırması için kullanmıştır. Bu sonuç araştırmacıların verilerini diğer araştırmacılarla paylaşmaktan çok, diğer araştırmacıların verilerini kullanmaya istekli olduğunu gösteren diğer araştırma bulgularıyla uyumaktadır (Tenopir ve diğerleri, 2011, s. 11). Veri paylaşımına en istekli araştırmacı grubu istatistiksel olarak anlamlı fark gösteren tıp araştırmacılarıdır. Elde edilen bu bulgu ile “Açık veri kullanma istekliliği fakültelere göre farklılık göstermektedir” hipotezi doğrulanmıştır. Araştırmacılar verilerinin paylaşım ve tekrar kullanım için uygun olduğunu düşünmektedir. Uygun olmamasına sebep olarak en çok verilerin yeni bir araştırma için uygun olmamasını göstermektedirler. Koltay’a (2015, s. 45) göre veri içerikten/bağlamdan ayrıldığında yeniden kullanımı zordur. Araştırmacıların verilerin yeniden kullanımının uygun olmadığı düşüncelerinin temelinde bu yatıyor olabilir. Diğer sebepler verilerin eksik veya hatalı olması, etik ihlal ve kişisel verilerin gizliliğinin ihlalidir. Yalnızca bir kişi verilerin yanlış yorumlanabileceğinden paylaşımına uygun olmadığını düşünmektedir. Araştırmacılar yeni yayın yapma ihtimalini, patent gibi ticari konuları, verilerin yanlış yorumlanabileceğini, verilerin kişisel ya da hassas veri içermesini ve fikirlerinin diğer araştırmacılar tarafından çalınma ihtimalini benzer düzeylerde önemsemektedirler.

Tıp arařtırmacıları ile eczacılık arařtırmacıları arařtırma verilerinin paylařılmasının önemli olduđunu düşünürken, diđer arařtırmacılar verilerin analizinden elde edilen sonuçların yayınlara aracılıđıyla paylařılmasını yeterli bulmaktadırlar. Arařtırmacılar kullandıkları veriler için atıf yapmaktan çekinmezken, bařka arařtırmacıların verilerini kullanmaktan çekinmemektedirler. Bulgulardan hareketle sađlık bilimleri arařtırmacılarının verilerini paylařmada ve bařka arařtırmacıların verilerini kullanmada önlerindeki en büyük engelin etik ihlal yapabilecekleri kaygısı olduđunu ve verinin yeniden kullanımını düzenleyen güçlü politikalar olduđunda verilerini paylařmaya istekli olduklarını söyleyebiliriz. Arařtırmacıları verilerini paylařmak için en çok motive eden verilerin paylařımını ve dođru şekilde yeniden kullanımını düzenleyen politikalarlardır. Bunun yanında arařtırmacılar veri ile iři bittikten sonra ve güvenilir bir veri arřivi bulduđunda verilerini paylařmaya isteklidirler. Arařtırmacılar için akademik açıdan ödül, puan veya atıf almak verilerini paylařmak için öncelikli motivasyon kaynađı deđildir.

Katılımcıların daha önce veri yönetim planı hazırlayıp hazırlamamaları ile fon desteđi almaları veya yurtdıřında eđitim görmeleri arasında bir iliřki olup olmadıđı arařtırılmıřtır. Elde edilen bulguya göre daha önce fon desteđi alarak arařtırma yürüten arařtırmacıların veri yönetim planı hazırlama konusunda tecrübeleri olduđu, istatistiksel olarak da diđer arařtırmacılarla anlamlı fark olduđu ortaya koyulmuřtur. "Veri yönetim planı hazırlama ile fon destekli arařtırma yürütüp yürütmeme durumu arasında iliřki vardır" hipotezi dođrulanmıřtır.

Arařtırmacıların eđitim ihtiyaçlarına bakıldıđında da tüm konularda öncelikli ihtiyaç duyduklarını belirtmiř olup en öncelikli olan konu verinin yeniden kullanım şartları ve telif hakları olmuřtur. Katılımcıların çalıřma ile ilgili görüşlerinde de vurguladıkları nokta arařtırma verilerinin yönetimi ve veri paylařımını konusunda bilgi eksikliđidir. Fane ve arkadaşlarının (2016, s. 14) arařtırmalarında ortaya koydukları, arařtırmacıların verilerini açma kořullarını bilmedikleri ve politika eksikliđi bulgusu ile uyumaktadır. Elde edilen bulgulara göre katılımcıların belirttikleri eđitim ihtiyaçı fakülterlere göre farklılık göstermemektedir ve "Eđitim ihtiyaçı fakülterlere göre farklılık göstermemektedir" hipotezi reddedilmiřtir. Ders içerikleri analiz sonuçlarına bakıldıđında ise arařtırma becerileri ile ilgili içeriklerin sađlandıđı ancak veri ile ilgili içeriđin toplama, analiz ve yorum ařamasında kaldıđı, bu süreçler dıřında veri yönetimi ile ilgili içerik olmadıđı

ortaya koyulmuştur. Elde edilen bu bulgu ile “Eğitim programlarında veri yönetimi ile ilgili yeterli içerik yoktur” hipotezi doğrulanmıştır.

## 5.2. ÖNERİLER

Katılımcıların büyük çoğunluğunun bilgi arama davranışlarını lisansüstü derslerde ve kendi kendine öğrendiğini belirtmesi, doğru bilgiye erişimde olumlu yaklaşımları da göz önünde bulundurulduğunda bölümlerde okutulan bu derslerin öğretim hedeflerine ulaştığını göstermektedir. Ayrıca belirtilen derslere "Veri Okuryazarlığı" içeriklerinin bir modül olarak eklenmesi genç araştırmacıların araştırma verilerini yönetme becerilerini geliştirmelerine ve verilerini diğer araştırmacılarla paylaşma motivasyonu kazanmalarına önemli destek sağlayacaktır.

Araştırmaya katılan katılımcıların kullandığı veri türlerinin daha önce akademisyenler üzerinde yapılmış çalışmalardan farklılık gösterdiği görülmektedir. Bunun sebebi lisansüstü araştırmacıların eğitim sürecinde deneysel çalışmalara daha az katılmaları olabilir. Araştırmacıların deneysel çalışmalara katılımının artırılması ve teşvik edilmesi bu anlamda fayda sağlayabilir.

Bulgular ile sağlık bilimleri araştırmacılarının verilerini paylaşmada ve başka araştırmacıların verilerini kullanmada önlerindeki en büyük engelin etik ihlal yapabilecekleri kaygısı olduğunu ve verinin yeniden kullanımını düzenleyen güçlü politikalar olduğunda verilerini paylaşmaya istekli olduklarını söyleyebiliriz. Bu araştırmanın ortaya koydukları, literatürdeki diğer çalışmalarda ortaya koyulan araştırmacıların verilerini açma koşullarını bilmedikleri ve politika eksikliği bulgusu ile uyumaktadır. Üniversitenin daha güçlü bir politika oluşturması ve bu politikayı uygulamada kararlı tutum sergilemesi araştırmacıları veri paylaşımına teşvik etmeye katkı sağlayabilir.

Ders içeriklerinin çoğunda veriler bir noktada öğretim için kullanılıyor olsa da iyi veri yönetimi uygulamaları ve veri yönetiminin önemini tam olarak anlaşılması öğretilmemektedir. Lisans eğitiminin, lisansüstü öğrenciler ve çalışan araştırmacıların sahip olmaları beklenen veri toplama, literatür tarama, veri toplama, verilerin analizi, yorumlanması ve sunulması da dahil olmak üzere, araştırma sürecinin temel becerilerinin



öğrenildiği bir dönem olması gerektiği düşünülmektedir. Bu eğitimin lisans düzeyinde sınıf içi veya küçük grup çalışmaları gibi daha uygulamaya dönük çalışmalarla desteklenerek verilmesi öğrencilerin veriyi anlama, yorumlama ve bilimsel bakış açısı kazanmasında önemli rol oynayarak genç araştırmacıların yetişmesinde önemli destek sağlayacaktır. Lisansüstü eğitime geçildiğinde ise artık alana özgü araştırmalar gerçekleştirebilmeleri, hızlı ilerleyen bilimsel literatürde araştırmacılara avantaj sağlayacaktır. Eğitimli veri yönetimi uzmanları ile alanlarında uzman eğitmenlerin iş birliğine dayalı öğretim, veri yönetimi eğitiminin farklı disiplinlerin eğitim programlarına uyarlanmasında fayda sağlar.

Her ne kadar araştırmalar, veriler kamuya açık hale getirildiğinde bir yayının atıf oranının daha yüksek olduğunu gösteriyor olsa da şimdiye kadar veri üreticilerinin akademik faydası sınırlı olmuştur. Araştırmacıları verilerini paylaşmaya ikna etmek için daha güçlü teşviklere ihtiyaç vardır. Bu teşvikler akademik puan, ödül veya yükselme kriteri olabilir. Öğretim üyeliğine yükseltme ve atama kriterleri ile akademik teşvik ödeneği yönetmeliği gibi yönetmeliklerde veri paylaşımı ile ilgili madde eklenmesi araştırmacıların faydasına olabilir.

Bu çalışmada ortaya koyulan sonuçlar, Türkiye’de bir araştırma kurumunun araştırma verilerinin paylaşımı ve veri yönetim uygulamalarına bakış açısını sunarak tıp ve sağlık bilimleri araştırmacılarının veri okuryazarlığı eğitimi ihtiyaçlarının daha iyi anlaşılmasına katkıda bulunabilir.

## KAYNAKÇA

- Adamick J., Reznik-Zellen, R. ve Sheridan, M. (2012). Data management training for graduate students at a large research university. *Journal of eScience Librarianship*, 1(3), 180–188. doi: 10.7191/jeslib.2012.1022
- Akers, K. G. ve Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *The International Journal of Digital Curation*, 8(2), 5-26. doi:10.2218/ijdc.v8i2.263
- Alkan, C. (2019). Eğitim bilimlerinde araştırma. *Ankara University Journal of Faculty of Educational Sciences (JFES)*, 22(1), 23-27. doi:10.1501/Egifak\_0000000860
- Altındış, S. (2019). Büyük verinin sağlık hizmetleri kalitesindeki rolü. *Sakarya Tıp Dergisi*. 8(2), 205-213. doi:10.31832/smj.423011
- Altman, M., ve Crosas, M. (2013). The evolution of data citation: From principles to implementation. *IASSIST Quarterly*, 63.62-70. Erişim adresi: [https://iassistquarterly.com/public/pdfs/iqvol371\\_4\\_altman.pdf](https://iassistquarterly.com/public/pdfs/iqvol371_4_altman.pdf)
- Anderson, T., Gjerdinge, C., Herrman, B., Himes, K., Johnston, L.R. (2009). Implementing research cyberinfrastructure for the 21st century. University of Minnesota. Retrieved from the University of Minnesota Digital Conservancy, Erişim adresi: <https://hdl.handle.net/11299/91772>
- Araştırma Verileri Yönetimi Eğitim Portalı, (2022). Erişim adresi: <https://acikveri.ulakbim.gov.tr/>
- Aydinoğlu, 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
- Ayvaz, S., Horn, J., Hassanzadeh, O., Zhu, Q., Stan, J., Tatonetti, N.P., Vilar, S., Brochhausen, M., Samwald, M., Rastegar-Mojarad, M., Dumontier, M. ve Boyce, R.D., (2015). Toward a complete dataset of drug–drug interaction information from publicly available sources. *Journal of Biomedical Informatics*, 55, 206-217. doi:10.1016/j.jbi.2015.04.006

- Bains, K. J. (2016). Big data analytics in healthcare-Its benefits, phases and challenges. *International Journal of Advanced Research in Computer Science and Software Engineering*, 6(4), 430-435.
- Bardou, D., Zhang, K. ve Ahmad, S. (2018). Lung sounds classification using convolutional neural networks. *Artificial Intelligence in Medicine*, 88, 58-69. doi:10.1016/j.artmed.2018.04.008
- Bardyn, T. P., Resnick, T. ve Camina, S. K. (2012). Translational researchers' perceptions of data management practices and data curation needs: Findings from a focus group in an academic health sciences library. *Journal of Web Librarianship*, 6(4), 274-287. doi:10.1080/19322909.2012.730375
- Baydilli, Y.Y. (2020). Derin öğrenme yöntemleri ile periferik yayma görüntülerinin analizi ve sınıflandırılması. Doktora Tezi, Karabük Üniversitesi, Karabük.
- Baynam, G., Walters, M., Claes, P., Kung, S., LeSouef, P., Dawkins, H., Bellgard, M., Girdea, M., Brudno, M., Robinson, P., Zankl, A., Groza, T., Gillett, D. ve Goldblatt, J. (2015). Phenotyping: Targeting genotype's rich cousin for diagnosis. *Journal of Paediatrics and Child Health*, 51(4), 381-386. doi:10.1111/jpc.12705
- Begley, C. ve Ellis, L. (2012). Raise standards for preclinical cancer research. *Nature*, 483, 531–533. doi:10.1038/483531a
- Begley, C. G. ve Ioannidis, J. P. (2015). Reproducibility in science: improving the standard for basic and preclinical research. *Circulation research*, 116(1), 116-126. doi:10.1161/CIRCRESAHA.114.303819
- Bell, G., Hey, T. ve Szalay, A. (2009). Beyond the data deluge. *Science*, 323(5919), 1297-1298. doi:10.1126/science.1170411
- Berkley Library (2022). Resarch Data Management. Erişim adresi: <https://guides.lib.berkeley.edu/researchdata>

- Borghgi, J.A. ve Van Gulick, A.E., (2018). Data management and sharing in neuroimaging: Practices and perceptions of MRI researchers. *PLoS ONE*, 13(7), doi:10.1371/journal.pone.0200562
- Borgman, C.L. (2015a). Big data, little data, no data: Scholarship in the networked world. MIT Press.
- Borgman, C. L., Darch, P. T., Sands, A. E., Pasquetto, I. V., Golshan, M. S., Wallis, J. C. ve Traweek, S. (2015b). Knowledge infrastructures in science: data, diversity, and digital libraries. *International Journal on Digital Libraries*, 16, 207–227. doi:10.1007/s00799-015-0157-z
- Bosch, X. (2008). An open challenge. Open access and the challenges for scientific publishing. *EMBO Reports*, 9(5), 404-408. doi:10.1038/embor.2008.60
- Briney, K. A., Coates, H. ve Goben, A. (2020). Foundational practices of research data management. *Research Ideas and Outcomes*, 6. doi:10.3897/rio.6.e56508
- Budapeşte Açık Erişim Girişimi, (2012). Erişim adresi: <https://www.budapestopenaccessinitiative.org/boai-10-translations/turkish-translation>
- Calzado Prado, J. ve Marzal, M.Á. (2013), Incorporating data literacy into information literacy programs: core competencies and contents, *Libri*, 63(2), 123-134. doi:10.1515/libri-2013-0010
- Carlson, J. (2014). The use of life cycle models in developing and supporting data services. In J. Ray (Ed.). *Research Data Management: Practical Strategies for Information Professionals* (s. 63-86) içinde. West Lafayette: Purdue University Press. doi: 10.2307/j.ctt6wq34t.6
- Carlson, J., Nelson M.S., Fosmire, M. ve Miller, C. (2011). Determining data information literacy needs: A study of students and research faculty. *Portal-libraries and The Academy*, 11(2). 629-657. doi:10.1353/pla.2011.0022
- Chen, X. ve Wu, M. (2017). Survey on the needs for chemistry research data management and sharing. *The Journal of Academic Librarianship*. 43(4). 346-353. doi:10.1016/j.acalib.2017.06.006

- Cios, K.J. ve Moore, W. G. (2002). Uniqueness of medical data mining. *Artificial Intelligence in Medicine*, 26(1-2). doi:10.1016/S0933-3657(02)00049-0
- Coorevits, P., Sundgren, M., Klein, G.O., Bahr, A., Claerhout, B., Daniel, C., Dugas, M., Dupont, D., Schmidt, A., Singleton, P., De Moor, G. ve Kalra, D. (2013). Electronic health records: New opportunities for clinical research. *Journal of Internal Medicine*, 274(6), 547–560. doi:10.1111/joim.12119
- Corti, L., Van den Eynden, V., Bishop, L. ve Woollard, M. (2019). Managing and sharing research data: A guide to good practice. E-kitap
- Cox, A. M., Kennan, M. A., Lyon, E. J., Pinfield, S. ve Sbaffi, L. (2019). Progress in research data services: An international survey of university libraries. *International Journal of Digital Curation*, 14(1), 126-135. doi:102218/ijdc.v14i1.595
- Creative Commons (t.y.). Erişim adresi: <https://creativecommons.org/about/cclicenses/>
- Creswell, J. W. ve Clark, V. L. P. (2017). Designing and conducting mixed methods research. *Sage publications*.
- Curty, R. G. (2015). Beyond "data thrifting": An investigation of factors influencing research data reuse in the social sciences (Order No. 3713677). Available from ProQuest Dissertations & Theses Global. (1698463092). Retrieved from <https://www.proquest.com/dissertations-theses/beyond-data-thrifting-investigation-factors/docview/1698463092/se-2>
- Digital Curation Center (2019). Overview of funders' data policies. Erişim adresi: <https://www.dcc.ac.uk/guidance/policy/overview-funders-data-policies>
- DMP Tool (t.y.). Erişim adresi: <https://dmptool.org/>
- DMP Online (t.y.). Erişim adresi: <https://dmponline.dcc.ac.uk/>
- 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. doi:10.1177/0340035220917985

- Duan, Q., Wang, X. and Song, N. (2022), Reuse-oriented data publishing: How to make the shared research data friendlier for researchers. *Learned Publishing*, 35, 7-15. <https://doi.org/10.1002/leap.1444>
- Ehrenstein, V., Nielsen, H., Pedersen, A.B., Johnsen, S.P. ve Pedersen, L. (2017). Clinical epidemiology in the era of big data: New opportunities, familiar challenges. *Clinical Epidemiology*, 9, 245-250.
- Errington, T. M., Denis, A., Perfito, N., Iorns, E. ve Nosek, B. A. (2021). Challenges for assessing replicability in preclinical cancer biology. *eLife*, 10. doi:10.7554/eLife.67995
- Eva, K. W., Cunningham, J. P., Reiter, H. I., Keane, D. R. ve Norman, G. R. (2004). How can I know what I don't know? Poor self assessment in a well-defined domain. *Advances in health sciences education: theory and practice*, 9(3), 211–224. doi:10.1023/B:AHSE.0000038209.65714.d4
- Faundeen, J.L., Burley, T.E., Carlino, J.A., Govoni, D.L., Henkel, H.S., Holl, S.L., Hutchison, V.B., Martín, Elizabeth, Montgomery, E.T., Ladino, C.C., Tessler, Steven, ve Zolly, L.S., (2013). The united states geological survey science data lifecycle model. *U.S. Geological Survey Open-File Report*, 2013–1265, (2331-1258). doi:10.3133/ofr20131265
- Federer, L.M., Lou, Y.L. ve Joubert, D.J. (2016). Data literacy training needs of biomedical researchers. *Journal of the Medical Library Association*, 104(1), 52–57. doi:10.3163/1536-5050.104.1.008
- Freedman, L. P., Cockburn, I. M., Simcoe, T. S. (2015) The economics of reproducibility in preclinical research. *PLoS Biology*, 13(6), 1-9. doi:10.1371/journal.pbio.1002165
- Gewin, V. (2016). Data sharing: An open mind on open data. *Nature*, 529, 117–119. doi:10.1038/nj7584-117a
- Glass, R. M. (1993). JAMA and the "Cantekin Affair". *The Pediatric infectious disease journal*, 12(10), 891. doi:10.1097/00006454-199310000-00025
- Goodman, S. N., Fanelli, D. ve Ioannidis, J. P. (2016). What does research reproducibility mean?. *Science translational medicine*, 8(341). doi:10.1126/scitranslmed.aaf5027

- Gorgolewski K.J. ve Poldrack, R.A. (2016). A practical guide for improving transparency and reproducibility in neuroimaging research. *PLOS Biology* 14(7). doi:10.1371/journal.pbio.1002506
- Gray, J. (2010). launch of the panton principles for open data in science and ‘Is it open data?’ web service. <https://blog.okfn.org/2010/02/19/launch-of-the-panton-principles-for-open-data-in-science/>
- Green, A., Macdonald, S. ve Rice, R. (2009). Policy-making for research data in repositories: A guide. EDINA.
- Hackam, D. G. ve Redelmeier, D. A. (2006). Translation of research evidence from animals to humans. *JAMA*, 296(14), 1731–1732. doi:10.1001/jama.296.14.1731
- Harvard University (2022). Research Data Management at Harvard. <https://researchdatamanagement.harvard.edu/data-management-plans>
- Heaton, J. (2004). Reworking qualitative data. Thousand Oaks, CA: SAGE Publications.
- Hey, T., Tansley, S. ve Tolle, K. M. (2009). Jim Gray on eScience: a transformed scientific method. Erişim adresi: <http://languagelog ldc.upenn.edu/myl/JimGrayOnE-Science.pdf>
- Hoffmann, I., Behrends, M., Consortium, H., ve Marschollek, M. (2022). Data literacy in medical education—an expedition into the world of medical data. *Advances in Informatics, Management and Technology in Healthcare*, 295, 257-260. doi:10.3233/SHTI220711
- Hughes, C. S. ve Morin, G. B. (2018). Using public data for comparative proteome analysis in precision medicine programs. *PROTEOMICS–Clinical Applications*, 12(2). doi:10.1002/prca.201600179
- Hyndman, A. ve Goody, G. (2020). What is the state of open data? It all about starts with a good plan. Digital Science Report. The State of Open Data 2020: The longest-running longitudinal survey and analysis on open data içinde (ss. 17-21). Erişim adresi: [https://digitalscience.figshare.com/articles/report/The\\_State\\_of\\_Open\\_Data\\_2020/13227875](https://digitalscience.figshare.com/articles/report/The_State_of_Open_Data_2020/13227875)

- İnce, Ö., Palacı, H., Ataseven Bulun, M., Yazar, O. (2018). Sağlık kuruluşlarında görev yapan üst ve orta düzey yöneticilerin büyük veri algıları. *Journal of International Social Research*, 11(61), 848-859.
- Ismail, L., Materwala, H., Karduck, A. P. ve Adem, A. (2020). Requirements of health data management systems for biomedical care and research: Scoping Review. *Journal of medical Internet research*, 22(7), e17508. doi:10.2196/17508
- Kaptan, S. (1991). Bilimsel Araştırma ve İstatistik Teknikleri, Tekışık, Web Ofset, Ankara.
- Kaul, V., Gallo de Moraes, A., Khateeb, D., Greenstein, Y., Winter, G., Chae, J., Stewart, N. H., Qadir, N. ve Dangayach, N. S. (2021). Medical education during the COVID-19 pandemic. *Chest*, 159(5), 1949–1960. doi:10.1016/j.chest.2020.12.026
- Kayalp, D., Dogan, K., Ceylan, G., Senes, M. ve Yucel, D. (2013). Can routine automated urinalysis reduce culture requests?, *Clinical Biochemistry*, 46(13-14), 1285-1289. doi:10.1016/j.clinbiochem.2013.06.015
- Kim, J. (2015). A study on the perceptions of university researchers on data management and sharing. *Journal of the Korean Society for Library and Information Science*, 49(3), 413–436. doi:10.4275/kslis.2015.49.3.413
- Knowles M. (1975). Self-directed learning: A guide for learners and teachers. New York, Associated Press.
- Koltay, T. (2017), “Data literacy for researchers and data librarians”. *Journal of Librarianship and Information Science*, 49(1), 3-14. doi:10.1177/0961000605616450
- Köse, E. (2011). Tıp akademisyenlerinin kanıta dayalı bilgi gereksinimleri ve bilgi arama davranışları: Hacettepe Üniversitesi örneği. Yüksek Lisans Tezi, Ankara.
- Krahe, M. A., Toohey J., Wolski M., Scuffham P. A. ve Reilly S. (2020). Research data management in practice: Results from a cross-sectional survey of health and medical researchers from an academic institution in Australia. *Health Information Management Journal*, 49(2-3):108-116. doi:10.1177/1833358319831318



- Laakso, M., Welling, P., Bukvova, H., Nyman, L. ve Björk, B. C. (2011). The development of open access journal publishing from 1993 to 2009. *PLoS ONE* 6(6). doi:10.1371/journal.pone.0020961
- Lamer, A., Al Massati, S., Saint-Dizier, C., Fares, E., Chazard, E. ve Fruchart, M. (2022). Data management for health data reuse: Proposal of a standard workflow and a r tutorial with jupyter notebook. *Studies In Health Technology And Informatics*, 298, 82–86. doi:10.3233/SHTI220912
- Lamprecht, A., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Pico, E.M., Angel, V.D., Sandt, S.V., Ison, J.C., Martínez, P.A., McQuilton, P., Valencia, A., Harrow, J.L., Psomopoulos, F., Gelpi, J.L., Chue Hong, N.P., Goble, C.A. ve Capella-Gutiérrez, S. (2020). Towards FAIR principles for research software. *Data Science*, 3, 37-59. doi: 10.3233/DS-190026
- Malgieri, G. ve Comandé, G. (2017). Sensitive-by-distance: Quasi-health data in the algorithmic era. *Inf. Commun. Technol. Law*, 26(3), 229–249. doi: 10.1080/13600834.2017.1335468
- Martone, M. (2014). Data citation synthesis group: Joint declaration of data citation principles. *San Diego CA: Force11*. Erişim adresi: doi:10.25490/a97f-egykh
- Martone, M. E. ve Nakamura, R. (2022). Changing the culture on data management and sharing: Getting ready for the new NIH data sharing policy. *Harvard Data Science Review*, 4(3). doi:10.1162/99608f92.6650ce2b
- MedOANet (2022). Mediterranean Open Access Network. Erişim adresi: <http://medoanet.eu/>
- Metadata Standards Directory (2022). Erişim adresi: <http://rd-alliance.github.io/metadata-directory/standards/>
- Meyer, E. F. (1997). The first years of the Protein Data Bank. *Protein Science*, 6(7), 1591–1597. doi:10.1002/pro.5560060724
- Meystre, S. M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A. ve Lehmann, C. U. (2017). Clinical data reuse or secondary use: Current status and potential future progress. *Yearbook of medical informatics*, 26(1), 38–52. doi:10.15265/IY-2017-007

- Milewska, A., Wiśniewska, N., Cimoszko, P. ve Rusakow, J. (2021). A survey of medical researchers indicates poor awareness of research data management processes and a role for data librarians. *Health information and libraries journal*, 39(2), 132–141. doi:10.1111/hir.12403
- Morse, J. M., (1991). Approaches to qualitative - quantitative methodological triangulation. *Nursing Research*: 40(2), 120-123.
- Munafò, M., Nosek, B., Bishop, D., Button, K.S., Chambers, C.D., Sert, N.P., Simonsohn, U., Wagenmakers, E., Ware, J.J. ve Ioannidis, J.P.A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(21), 1-9. doi: 10.1038/s41562-016-0021
- Nath, A.S., Pal, A., Mukhopadhyay, S. ve Mondal, K.C. (2020). A survey on cancer prediction and detection with data analysis. *Innovations in Systems and Software Engineering*, 16, 231–243. doi:10.1007/s11334-019-00350-6
- National Academies of Sciences, Engineering, and Medicine; Policy and Global Affairs; Division on Earth and Life Studies; Division on Engineering and Physical Sciences; Board on Research Data and Information; Board on Life Sciences; Computer Science and Telecommunications Board; Committee on Applied and Theoretical Statistics; Board on Mathematical Sciences and Analytics; Committee on Forecasting Costs for Preserving and Promoting Access to Biomedical Data. (2020). Life-cycle decisions for biomedical data: The challenge of forecasting costs. *National Academies Press*. (US).
- National Science Foundation (2011). *Digital research data sharing and management*. Erişim adresi: [www.nsf.gov/nsb/publications/2011/nsb1124.pdf](http://www.nsf.gov/nsb/publications/2011/nsb1124.pdf)
- Nelson, N. C., Ichikawa K., Chung J., Malik, M. M. (2021). Mapping the discursive dimensions of the reproducibility crisis: A mixed methods analysis. *PLoS ONE* 16(7): 1-20. doi:10.1371/journal.pone.0254090
- Neuman, W. L. (2012). Toplumsal araştırma yöntemleri: Nicel ve nitel yaklaşımlar III. Cilt (5. Basım). İstanbul: Yayın Odası.

- Neylon, C. (2017). Data management plan: IDRC data sharing pilot project. *Research Ideas and Outcomes* 3. doi:10.3897/rio.3.e14672
- Nielsen, M. (2011). Reinventing discovery: the new era of networked science. Princeton, NJ: Princeton University Press.
- OECD (2007). OECD principles and guidelines for access to research data from public funding. Erişim adresi: <https://www.oecd.org/sti/inno/38500813.pdf>
- OECD (2015). Making open science a reality, *OECD Science, Technology and Industry Policy Papers*, No. 25, OECD Publishing. doi:10.1787/5jrs2f963zs1-en.
- O'Shea, E. (2003). Self-directed learning in nurse education: a review of the literature. *Journal of Advanced Nursing*, 43(1), 62-70. doi:10.1046/j.1365-2648.2003.02673.x
- Open Knowledge Foundation. (t.y.). Open Data Commons. <https://opendatacommons.org/>
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). doi:10.1126/science.aac4716
- OpenAIRE2020 (2022). Erişim adresi: <https://ufukavrupa.org.tr/tr>
- Pasek, J. E. ve Mayer, J. (2019). Education needs in research data management for science-based disciplines: Self-assessment surveys of graduate and faculty at two public universities. *University Libraries Faculty Publications and Presentations*, 88. doi: 10/29173/istl12
- PASTEUR40A (2022). Open access policy alignment strategies for european union research. Erişim adresi: <http://pasteur4oa.eu/home>
- Patel, D. (2016). Research data management: a conceptual framework. *Library Review*, 65(4/5), 226-241. doi:10.1108/LR-01-2016-0001
- Piorun, M.E., Kafel, D., Leger-Hornby, T., Najafi, S., Martin, E.R., Colombo, P. ve LaPelle, N.R. (2012). Teaching research data management: An undergraduate/graduate curriculum. *Journal of eScience Librarianship* 1(1), 46-50. doi:10.7191/jeslib.2012.1003

- Piwowar, H., Day, R. ve Fridsma, D. (2007). Sharing detailed research data is associated with increased citation rate. *PLoS One*. 2(3). doi:10.1371/journal.pone.0000308
- Piwowar, H.A. ve Vision, T.J. (2013), Data reuse and the open data citation advantage. *PeerJ* 1:e175. doi:10.7717/peerj.175
- Prinz, F., Schlange, T. ve Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets?. *Nature Reviews Drug Discovery*, 10(712). doi:10.1038/nrd3439-c1
- Rathi, V., Dzara, K., Gross, C. P., Hrynaszkiewicz, I., Joffe, S., Krumholz, H. (2012). Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ*, 345. doi:10.1136/bmj.e7570
- Read, K. B. (2019b). Adapting data management education to support clinical research projects in an academic medical center. *Journal of the Medical Library Association: JMLA*, 107(1), 89. doi:10.5195/jmla.2019.580
- Read, K.B., Larson, C., Gillespie, C., Oh, S.Y. ve Surkis, A. (2019a). A two-tiered curriculum to improve data management practices for researchers. *PLoS ONE* 14(5). doi:10.1371/journal.pone.0215509
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D.J., Kelley, D.R., Matwin, S. ve Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report. doi: 10.13140/RG.2.1.1922.5044
- 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>
- Ross, J. S. (2016). Clinical research data sharing: What an open science world means for researchers involved in evidence synthesis. *Ross Systematic Reviews*, 5(159). doi:10.1186/s13643-016-0334-1

- Ross, J. S., Gross, C. P. ve Krumholz, H. M. (2012). Promoting transparency in pharmaceutical industry-sponsored research. *American Journal of Public Health*, 102(1), 72-80. doi:10.2105/AJPH.2011.300187
- Rubinstein, Y.R., Robinson, P.N., Gahl, W.A., Avillach, P., Baynam, G., Cederroth, H., Goodwin, R.M., Groft, S.C., Hansson, M.G., Harris, N.L., Huser, V., Mascalzoni, D., McMurry, J.A., Might M., Nellake,r C., Mons B., Paltoo, D.N., Pevsner, J., Posada, M., Rockett-Frase, A.P., Roos, M., Rubinstein, T.B., Taruscio, D., van Enkevort, E. ve Haendel, M.A. (2020). The case for open science: rare diseases. *JAMIA Open*. 11;3(3):472-486. doi:10.1093/jamiaopen/ooaa030
- Safran, C. (2014). Reuse of clinical data. *Yearbook of medical informatics*, 9(1), 52-54. doi:10.15265/IY-2014-0013
- Sağlık Bakanlığı (2017). Sağlık Bilgi Sistemleri Uygulamaları Hakkında 2015/17 Sayılı Genelge. Erişim Adresi: <https://www.saglik.gov.tr/TR,11183/saglik-bilgi-sistemleri-uygulamalari-hakkinda-201517-sayili-genelge.html>
- Sawatsky, A. P., Ratelle, J. T., Bonnes, S. L., Egginton, J. S. ve Beckman, T. J. (2018). Faculty support for self-directed learning in internal medicine residency: A qualitative study using grounded theory. *Academic Medicine*, 93(6), 943-951. doi:10.1097/ACM.0000000000002077
- 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. doi:10.5860/crl-255
- Smale, N., Unsworth, K., Denyer, G. ve Barr, D. (2018). The history, advocacy and efficacy of data management plans. doi:10.1101/443499
- Steinmetz, C.H. (1951). Statistical aspects of the simultaneous detection of thyroid and thyrotrophic hormones in human sera, based on the data of Dangelo and Gordon. *Biometrics*, 7(4), 323-326. doi:10.2307/3001654

- Strasser, C. A. ve Hampton, S. E. (2012). The fractured lab notebook: Undergraduates and ecological data management training in the United States. *Ecosphere*, 3(12), 1-18. doi:10.1890/ES12-00139.1
- Sturges, P., Bamkin, M., Ander, J.H.S., Hubbard, B., Hussain, A. ve Heeley, M. (2015). Research data sharing: Developing a stakeholder-driven model for journal policies. *Journal of The Association for Information Science and Technology*. 66(12), 2445–2455. doi:10.1002/asi.23336
- Tackett, S., Wright, S. M. ve Quirk, M. (2018). Adaptive medical education research. *Medical Teacher*, 40(8), 783–785. doi:10.1080/0142159X.2018.1490705
- Tavluoğlu, C. (2022). Araştırma verilerinin yönetimi: TÜBİTAK projelerinde yer alan araştırmacılar üzerine bir değerlendirme. Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Teal, T. K., Cranston, K. A., Lapp, H., White, E., Wilson, G., Ram, K. ve Pawlik, A. (2015). Data carpentry: workshops to increase data literacy for researchers. *International Journal of Digital Curation*, 10(1). doi:10.2218/ijdc.v10i1.351
- Telenti, A. ve Jiang, X. (2020). Treating medical data as a durable asset. *Nature Genetics* 52, 1005-1010. doi:10.1038/s41588-020-0698-y
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E. ve Frame, M. (2011). Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* 6(6): e21101. doi:10.1371/journal.pone.0021101
- Tenopir, C., Allard, S., Sinha, P., Pollock, D., Newman, J., Dalton, E., Frame, M. ve Baird, L. (2016). Data management education from the perspective of science educators. *International Journal of Digital Curation*, 11(1), 232-251. doi:10.2218/ijdc.v11i1.389
- Thoegersen, J. L. (2015). Examination of federal data management plan guidelines. *Journal of eScience Librarianship. EScience in Action*. 4(1). 1-10 e1072 doi:10.7191/jeslib.2015.1072

- Toğaçar, M., Ergen, B. ve Sertkaya, M. (2019). Zatürre hastalığının derin öğrenme modeli ile tespiti. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 31(1), 223-230. <https://dergipark.org.tr/tr/pub/fumbd/issue/43638/498364>
- Tse E.G., Klug D.M. ve Todd, M.H. (2020). Open science approaches to COVID-19. *F1000Research*, 9(1043). doi:10.12688/f1000research.26084.1
- UFUK2020 (2022). Erişim adresi: <https://ufukavrupa.org/tr>
- UK Edinburgh (2020). University of Edinburgh Information Services. Research data management programme: research data management home [Internet]. Edinburgh, UK: The University; Erişim adresi: <https://www.ed.ac.uk/information-services/research-support/research-data-service/about-the-research-data-service>
- University of Edinburgh Information Services (2022). Research data management programme: research data management home (internet). Edinburgh, UK: The university. Erişim adresi: <http://www.ed.ac.uk/schools-departments/information-services>
- URAP, (2022). University Ranking by Academic Performance. Erişim adresi: <https://newtr.urapcenter.org/>
- 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
- Vayena, E., Dzenowagis, J., Brownstein, S. J. ve Sheikh A. (2018). Policy implications of big data in the health sector. *Bulletin of the World Health Organization*, 96(1), 66-68. doi: 10.2471/BLT.17.197426
- Wicherts J.M., Bakker M. ve Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS ONE* 6(11). doi:10.1371/journal.pone.0026828
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R.,

- Gonzalez-Beltran, A., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 3, 160018. doi:doi:10.1038/sdata.2016.18
- World Economic Forum (2019). 4 ways data is improving healthcare. Erişim adresi: <https://www.weforum.org/agenda/2019/12/four-ways-data-is-improving-healthcare>
- Yarkoni, T., Poldrack, R., Nichols, T., Essen D.V.C., ve Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods* 8(8). 665-670. doi:10.1038/nmeth.1635
- Yıldırım, A. ve Şimşek, H. (2008). Sosyal Bilimlerde Nitel Araştırma Yöntemleri (6. Baskı). Ankara: Seçkin Yayıncılık.
- Yıldırım, P. (2015). Association patterns in open data to explore ciprofloxacin adverse events. *Applied Clinical Informatics*, 6(4), 728–747. doi:10.4338/aci-2015-06-ra-0076
- Yuan, B. ve Li, J. (2019). The policy effect of the general data protection regulation (GDPR) on the digital public health sector in the european union: an empirical investigation. *International journal of environmental research and public health*, 16(6), 1070.1-15. doi:10.3390/ijerph16061070
- Zastrow, M. (2020). Open science takes on the coronavirus pandemic. *Nature*, 581(7806). 109-110. doi: 10.1038/d41586-020-01246-3
- Zencir, M. B. (2019). Ankara Üniversitesi akademisyenlerinin araştırma verilerinin yönetimi ile ilgili tutumları ve bir model önerisi. Doktora Tezi, Ankara Üniversitesi, Ankara.
- Zhang, X., Liu, C., Nepal, S., Yang, C., Dou, W. ve Chen, J. (2014). A hybrid approach for scalable sub-tree anonymization over big data using MapReduce on cloud. *Journal of Computer and System Sciences*, 80(5),1008-1020. doi:10.1016/j.jcss.2014.02.007



## EK 1. ORJİNALLİK RAPORU



**HACETTEPE ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU**

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

Tarih: 09 / 12 / 2022

Tez Başlığı: Sağlık Bilimlerinde Araştırma Verilerinin Yönetimi: Hacettepe Üniversitesi Lisansüstü Öğrencilerinin Farkındalıklarının Değerlendirilmesi.

Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 68 sayfalık kısmına ilişkin, 09/12/2022 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda işaretlenmiş filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 4 'tür.

Uygulanan filtrelemeler:

- 1-  Kabul/Onay ve Bildirim sayfaları hariç
- 2-  Kaynakça hariç
- 3-  Alıntılar hariç
- 4-  Alıntılar dâhil
- 5-  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:** Eda Özel

**Öğrenci No:** N19238282

**Anabilim Dalı:** Bilgi ve Belge Yönetimi

**Programı:** Yüksek Lisans

### **DANIŞMAN ONAYI**

UYGUNDUR.

Prof. Dr. Özgür Külcü



**HACETTEPE UNIVERSITY  
GRADUATE SCHOOL OF SOCIAL SCIENCES  
MASTER'S THESIS ORIGINALITY REPORT**

---

**HACETTEPE UNIVERSITY  
GRADUATE SCHOOL OF SOCIAL SCIENCES  
INFORMATION MANAGEMENT DEPARTMENT**

Date: 09 /12/2022

Thesis Title: Research Data Management in Health Sciences: Assesment of Hacettepe University Graduate Students' Awareness

According to the originality report obtained by myself/my thesis advisor by using the Turnitin plagiarism detection software and by applying the filtering options checked below on 09/12/2022 for the total of 68 pages including the a) Title Page, b) Introduction, c) Main Chapters, and d) Conclusion sections of my thesis entitled as above, the similarity index of my thesis is 4 %.

Filtering options applied:

1.  Approval and Decleration sections excluded
2.  Bibliography/Works Cited excluded
3.  Quotes excluded
4.  Quotes included
5.  Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Social Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Date and Signature

**Name Surname:** Eda Özel

**Student No:** N19238282

**Department:** Information Management

**Program:** Master's Degree

**ADVISOR APPROVAL**

APPROVED..

---

Prof. Dr. Özgür Külçü

## EK 2. ETİK KURUL İZİNİ

Tarih: 26/08/2022  
Sayı: E-35853172-300-00002358115



T.C.  
HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜ  
Rektörlük

Sayı : E-35853172-300-00002358115  
Konu : Eda ÖZEL (Etik Komisyon İzni)

26.08.2022

### SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi: 15.08.2022 tarihli ve E-12908312-300-00002338013 sayılı yazınız.

Enstitünüz Bilgi ve Belge Yönetimi Anabilim Dalı Yüksek Lisans Programı öğrencilerinden **Eda ÖZEL**'in, **Prof. Dr. Özgür KÜLCÜ** danışmanlığında hazırladığı “**Sağlık Bilimlerinde Araştırma Verilerinin Yönetimi: Hacettepe Üniversitesi Lisansüstü Öğrencilerinin Farkındahklarının Değerlendirilmesi**” başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun **23 Ağustos 2022** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Vural GÖKMEN  
Rektör Yardımcısı

Bu belge güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu: 92DC3E75-19E0-409A-BA76-D9FB7EB6C035

Belge Doğrulama Adresi: <https://www.turkiye.gov.tr/hu-ebys>

Adres: Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara  
E-posta: yazimd@hacettepe.edu.tr İnternet Adresi: www.hacettepe.edu.tr Elektronik  
Ağ: www.hacettepe.edu.tr  
Telefon: 0 (312) 305 3001-3002 Faks:0 (312) 311 9992  
Kep: hacettepeuniversitesi@hs01.kep.tr

Bilgi için: Çağla Handan GÜL  
Bilgisayar İşletmeni  
Telefon: 03123051008

