

Araştırma Verilerinin Yönetimi: Türk Araştırmacılar Verilerini Açmaya Hazır mı?*

Research Data Management: Are Turkish Researchers Prepared to Open their Data?

Yurdağül Ünal** ve Serap Kurbanoğlu***

Öz

Araştırma verilerinin erişime açılması, paylaşımı ve yeniden kullanımı bir taraftan teknik alt yapı, politikalar ve direktifleri (mandate) içeren düzenlemeler gerektirirken diğer taraftan veri yaşam döngüsü içindeki rolleri gereği araştırmacıların ve her disiplinde veri yönetimini gerçekleştirecek iş gücünün eğitilmesini gerektirmektedir. Ülkemiz üniversitelerinde çalışan araştırmacıların araştırma verilerinin oluşturulması, kullanımı, paylaşımı konularındaki uygulamaları ve araştırma verilerinin paylaşımı ile ilgili sorunlara bakış açılarının saptanması amacıyla gerçekleştirilen bu çalışmada 44 üniversiteden 540 araştırmacıya anket uygulanmıştır. Bulgular araştırmacıların kullandıkları ve ürettikleri veri türü açısından disiplinler arasında benzerlikler kadar farklılıkların da olduğunu, az sayıda araştırmacının büyük veri ile çalıştığını ve büyük veri ile çalışanların fen bilimlerinde yoğunlaştığını, araştırma verilerinin farklı kaynaklardan elde edildiğini, dış kaynaklı verinin genelde üzerinde bir miktar çaba harcadıktan sonra kullanıldığını, araştırmacıların ürettikleri veriyi genelde kendi cihazlarında depoladıklarını, depolarken verilerine ya hiç ek bilgi atamadıklarını ya da çok sınırlı ek bilgi atadıklarını göstermektedir. Bulgular ayrıca, veri paylaşımı konusuna genelde olumlu ancak temkinli yaklaştıklarına, veri paylaşımıyla ilgili endişelerinin başında yasal ve etik sorunların geldiğine, verilerini proje süresinden sonra da saklamak istediklerine, veri setlerinin depolanması ve erişime açılması konusundaki sorumluluğu öncelikle üniversitelerin yüklenmesi gerektiğini düşündüklerine, araştırma verilerinin yönetimi konusunda kapsamlı bir eğitime ihtiyaçları olduğuna ve bu konuda eğitim almaya istekli olduklarına işaret etmektedir.

Anahtar Sözcükler: Araştırma verilerinin yönetimi; araştırma verisi; açık veri; veri paylaşımı; açık erişim.

* Bu çalışma beş kıtada yürütülen çok uluslu ve geniş kapsamlı bir çalışmanın parçası olarak gerçekleştirilmiştir. Çok uluslu karşılaştırmalı veriler henüz analiz aşamasındadır ve yayımlanmamıştır. Ancak araştırma verilerinin yönetimi ile ilgili olarak farklı gelişmişlik düzeyinde bulunan sırasıyla Büyük Britanya, Fransa ve Türkiye verileri karşılaştırmalı olarak yayımlanmıştır. Yeri geldiğinde ilgili yayınlara metin içinde gönderme yapılmıştır. *This study is a part of a comprehensive multi-national research project. Comparative findings of the multi-national study is not published yet. Only a small part of it which compares data from UK, France and Turkey is published and referred within the text when it necessary.*

** Doç. Dr. Hacettepe Üniversitesi, Bilgi ve Belge Yönetimi Bölümü, 06800, Çankaya, Ankara. e-posta: unal.yurdagul@gmail.com

Assoc. Prof. Dr. Hacettepe University, Department of Information Management, 06800, Çankaya, Ankara, Turkey.

*** Prof. Dr. Hacettepe Üniversitesi, Bilgi ve Belge Yönetimi Bölümü, 06800, Çankaya, Ankara. e-posta: kurbanogluserap@gmail.com

Prof. Dr. Hacettepe University, Department of Information Management, 06800, Çankaya, Ankara, Turkey.

Geliş Tarihi - *Received*: 08.08.2018

Kabul Tarihi - *Accepted*: 13.12.2018

Abstract

Making research data open and available for easy access, sharing, and reuse requires not only building the appropriate ICT infrastructure and developing appropriate mandates and policies but also training researchers and workforce capable of managing research data in every discipline so that they can play appropriate roles throughout the research data lifecycle. Within this study, a questionnaire survey was conducted amongst university researchers, 540 researchers from 44 universities, in Turkey in order to gain an understanding of the current practices of research data creation, use and sharing amongst academic researchers, and their view on the underlying issues and challenges of research data management (RDM). The key findings from the research are that there are differences as well as similarities between the data type used and created by researchers from different disciplines, limited number of researchers work with big data and they generally come from sciences, research data is obtained from multiple sources, researchers need to spend some time to pre-process data received from outside sources, research data is generally stored on researchers' own devices, non or very limited number of tags are assigned to research data while storing, researchers feel that research data should be stored at their own universities for long-term storage and made available, they generally have positive views towards data sharing; however, they have some concerns such as data ethics, they like to keep their data after the closure of research project, they need comprehensive training on RDM, and are interested in receiving formal training on RDM related issues.

Keywords: Research data management; research data; open data; data sharing; open access.

Giriş

Bilimsel bilginin yönetimi yakın zamana kadar, daha çok araştırma sürecinin çıktısı olan makaleler ve kitaplar gibi bilimsel yayınlara odaklanmış, araştırmanın yürütülmesinde ve söz konusu çıktılarının üretilmesinde büyük katkısı olan araştırma verilerinin yönetimi ayağı genelde ihmal edilmiştir. Oysa araştırma verilerinin yaratılması/toplanması ve kullanımı araştırma faaliyetinin temelini oluşturur. Araştırmaların başarısı büyük ölçüde verinin ilgili disiplinin gereklerine ve amaca uygun olarak doğru bir şekilde yaratılması/toplanması, kodlanması, analiz edilmesi ve yorumlanmasına bağlıdır. Bu nedenle araştırma verileri literatürde araştırmanın can damarı (lifeblood of research) olarak da adlandırılmaktadır (Borgman, 2012, s. 1066).

Borgman, Wallis ve Mayernik (2012, s. 517) verinin bazen kendisini yaratan ve kullananlarca bile tam olarak anlaşılabilen karmaşık bir kavram olduğunu altını çizer. Nitekim veri kavramının farklı kişilerce ve farklı disiplinlerde farklı algılandığı görülmektedir. Bu çalışmada veri en geniş anlamıyla metin, ses, resim, hareketli resim, model, oyun ve simulasyonun dijital gösterimini, bilgisayarlar ve çeşitli yazılımlar yardımıyla kullanılabilir hale gelen laboratuvar verilerini (genomik ve elektron mikroskopu verileri gibi), gözlemsel verileri (uzaktan algılama, coğrafik ve mekansal veriler gibi), gerek bilgisayarlar gerekse insanlar tarafından toplanan sosyo ekonomik, sayısal ve başka her türlü veriyi içerir şekilde kullanılmaktadır (Borgman ve diğerleri, 2012, s. 488).

Son yıllarda bilgi ve iletişim teknolojilerinde özellikle web ve mobil teknolojiler alanında yaşanan gelişmeler, araştırma faaliyetlerini de etkilemiş, araştırmalar giderek daha veri-yoğun ve daha kolay ortak yürütülebilir hale gelmiştir. Günümüzde araştırmacılar için mevcut olan verinin miktarı, veri tufanı (deluge of data) olarak adlandırılacak kadar fazladır (Borgman, Wallis ve Enyedy, 2007; Frank ve Pharo, 2016, s. 537). Miktarı katlanarak artan araştırma verilerinin paylaşılması ve yeniden kullanımı ise son derece yararlıdır. Araştırma verilerinin paylaşılması her şeyden önce zaman, emek ve maliyet gerektiren veri toplama işleminin tekrarlanmasını engeller. Ayrıca farklı kaynaklardan veri elde edilmesine, bu verilerin karşılaştırılmasına ve farklı şekillerde ve birlikte analizine olanak tanır. Analizlerin tekrarlanabilir hale gelmesi araştırma bulgularının doğrulanması açısından önemlidir (Ball ve Duke, 2015; Beagrie ve Houghton, 2013; Lewis, 2010).

Araştırma verilerinin paylaşımının sağladığı ekonomik faydanın somut olarak ifade edildiği de görülmektedir. *Research Data Alliance Europe* (Avrupa Araştırma Verileri İttifakı) tarafından hazırlanan *The Data Harvest* (Veri Hasatı) başlıklı raporda ABD’de *Human Genome* projesi ve onu izleyen projelerin hükümete 13 milyar dolara mal olduğu, bunun karşılığında yaklaşık bir trilyon dolar ekonomik fayda sağladığı belirtilmiştir. Benzer bir şekilde Britanya’da yapılan bir çalışmada hükümet tarafından yapılan her bir sterlinlik yatırımın yaklaşık beş buçuk sterlinlik ekonomik dönüşü olduğu hesaplanmıştır (The Data Harvest, 2014).

Söz konusu gelişmelerin ışığında hükümetler ve araştırma kurumları araştırma verilerine açık erişimi desteklemeye ve teşvik etmeye başlamıştır. Örneğin, Birleşik Krallık araştırma kuruluşlarının yayınladığı antlaşmada (The Concordat on Open Research Data, 2015, s. 2), araştırma verilerinin erişime açılmasının ekonomik büyüme, kaynak verimliliğinde artış, araştırma fonlarına destek garantisi ve araştırmalara olan güvenin artması gibi sayısız toplumsal fayda getirdiği belirtilmektedir. Avrupa Komisyonu, Avrupa Birliği (AB) fonlarıyla desteklenen proje verilerinin paylaşımını desteklemekte, araştırma verilerinin bulunabilir, erişilebilir, birlikte çalışılabilir ve yeniden kullanılabilir (findable, accessible, interoperable and reusable - FAIR) olması için çalışmalar yürütmektedir. AB Ufuk 2020 programı kapsamında desteklenen projelerden üretilen verilerin erişilebilir olması için başlatılan *The Open Research Data Pilot* adlı proje 2014-2016 çalışma programlarında Ufuk 2020’nin sadece belirli alanlarını kapsarken, 2017 çalışma programının gözden geçirilmiş versiyonu ile Ufuk 2020’nin tüm tematik alanlarını kapsayacak şekilde genişletilmiştir (OpenAIRE, 2018). Verinin yenilikçi kullanımını sağlamak, ekonomik potansiyelini açığa çıkarmak ve AB kurumlarını ve diğer organlarını daha açık ve hesap verebilir hale getirmek amacıyla *The European Union Open Data Portal (EU ODP)* üzerinden AB kurumları ve diğer organları tarafından üretilmiş açık veri ücretsiz ve kolay bir şekilde erişime sunulmaktadır (EU Open Data Portal, 2018).

Yeniden kullanım için verilerin ücretsiz olarak erişime açılması fikri yeni olmamakla beraber, araştırmacılar tarafından benimsenmesi zaman almakta ve yavaş olmaktadır. Bu durum verinin başkaları tarafından keşfedilebilir, erişilebilir ve yeniden kullanılabilir hale getirilmesi için gerekli becerilerin eksikliği ile olduğu kadar güven, itibar ve etik gibi sorunlarla da ilişkilendirilmektedir (The Data Harvest, 2014; Faniel, Kansa, Whitcher Kansa, Barrera-Gomez ve Yakel, 2013; Faniel, Kriesbergand ve Yakel, 2012; Yakel ve Faniel 2014). Araştırma verilerinin paylaşımı ve yeniden kullanımı teknolojik alt yapı, politika geliştirme, kapasite oluşturma gibi çeşitli alanlarda gelişmeyi ve yatırımı zorunlu kılmaktadır.

Araştırma Verilerinin Yönetimi ile İlgili Sorunlar

Verinin yaşam süresi, yaratılmasına veya toplanmasına sebep olan araştırma projesinin yaşam süresinden daha uzundur. Dolayısıyla araştırma verilerinin yönetimi yalnızca ilgili araştırma projesinin başarıyla tamamlanması için değil aynı zamanda gelecekte yapılacak araştırmalar için de gereklidir. Araştırma verilerinin erişime açılmasının ve paylaşılmasının planlandığı durumlarda verinin tüm araştırma süreci boyunca yönetilmesi (Concordat on Open Research Data, 2015, s. 5) ve bu amaçla bir veri yönetim planı hazırlanması kritik önem taşımaktadır. Verinin açık hale getirilmesi için erişilir hale getirilmesi tek başına yeterli olmamakta, aynı zamanda keşfedilebilir ve kullanılabilir hale getirilmesi de önem taşımaktadır. Bu da ancak belli veri formatları ve standartlarının kullanımı, uygun tanımlamalarla verinin gerek kurumsal arşivlerde gerekse bilimsel dergilerin arşivlerinde depolanması ile mümkündür (Concordat on Open Research Data, 2015, s. 5).

Amerikan Ulusal Bilgi Standartları Örgütü (*US National Information Standards Organization - NISO*) tarafından geliştirilen araştırma verileri dokümanı (*Research Data Primer*) veri yönetim planlarının beş önemli unsuru tanımlaması gerektiğine işaret eder (Strasser, 2015, s. 2):

1. Bir araştırma projesi sırasında toplanacak veya üretilecek tüm veri türleri;

2. Veriyi tanımlamak için kullanılacak standartlar (üst veri);
3. Bir araştırma projesi sırasında toplanan veya üretilen veriyi ilgilendiren politikalar;
4. Proje esnasında üretilen verinin arşivlenmesi ve korunması;
5. Verinin yönetimi için gerekli kaynaklar (personel, donanım, yazılım ve finans).

İyi yönetilmiş veri bilimsel, toplumsal ve ekonomik gelişmede önemli bir rol oynar. Büyük veri setlerinin yönetilmesi ve bunlardan anlam çıkarılması için çeşitli araçlar ve teknikler kadar bilgisayar ve ağ altyapısına da gereksinim vardır. Bilgisayarlarca toplanan büyük miktarlardaki verinin depolanması, işlenmesi, analizi ve düzenlenmesi karmaşık otomasyon yöntemlerinin kullanımını gerekli kılmaktadır (Borgman, 2015). Birleşik Krallık Veri Arşivi (The UK Data Archive, 2018) altı ana aktiviteden oluşan bir araştırma verisi yaşam döngüsü öngörmektedir. Bu aktivitelerden yaratım, erişim, analiz ve yeniden kullanım aşamaları belli bir disiplindeki araştırmacılar tarafından gerçekleştirilirken veri yönetimi ve koruma aşamaları araştırma verisi yöneticilerinin faaliyet alanına girmektedir. Bu nedenle araştırma verisi yönetim aktivitelerinin gerçekleştirilmesinde bazı önemli roller üstlenen araştırmacıların konuyla ilgili sorunlar üzerinde bilgi sahibi olması ve verinin yaratımından yeniden kullanımına kadar geçen yaşam döngüsü içindeki her aşama için beceriler geliştirmesi gerekmektedir.

Araştırmacıların verilerini nasıl yönettiği ve verilerini kendi araştırma gruplarının dışındakilerle paylaşma konusundaki tutumları konusunda yapılan araştırmaların sayısı giderek artmakla birlikte henüz çok fazla şey bilinmemektedir (Kennan ve Markauskaite, 2015; Wallis, Rolando ve Borgman, 2013). Araştırmacılar genellikle verilerini yedekleme konusuna odaklanmakta ancak verilerin uzun süreli korunması konusuna kafa yormamaktadırlar (Carlson, Fosmire, Miller ve Nelson, 2011; Chowdhury, Boustany, Kurbanoglu, Ünal ve Walton, 2017). Araştırma verilerinin başkalarının kullanımına açılması ile ilgili sorunlar veri paylaşımının yeterince teşvik edilmemesi ve veri arşivlerinin geliştirilmesi ve tasarımlarında yaşanan yetersizlikler nedeniyle varlıklarını korumaya devam etmektedir (Borgman, 2012). Bunlara ek olarak verinin yeniden kullanımı süreç ile ilgili bağlamsal bilgi gerektirebilmektedir. Veri içerikten/bağlamdan ayrıldığında yeniden kullanımı zor hatta bazı durumlarda olanaksız hale gelmektedir (Koltay, 2015, s. 405). Veri paylaşımı olumlu bir davranış olarak kabul görse de bazı alanlarda talep az olduğu için araştırmacılar üst veri (metadata) ve belgelemeye çok az zaman ayırmakta ve bu durum verinin keşfedilmesini engelleyici bir unsur olarak karşımıza çıkmaktadır (Borgman, 2012; Wallis ve diğerleri, 2013). Araştırmalar az sayıda araştırmacının üst veri kullandığını (Chowdhury ve diğerleri, 2017; MacMillan, 2014) ve uluslararası standartlar yerine kendi laboratuvarlarına ait standartları kullanmayı tercih ettiklerini (Carlson ve diğerleri, 2011) göstermektedir. İstikrarlı veri paylaşımı sadece çok sınırlı alanda gerçekleşmekte ve genellikle uygulamalarda tutarsızlıklar gözlenmektedir (Borgman, 2012). Veri paylaşımı sanat ve beşerî bilimlerden çok fen bilimleri alanında gerçekleşmektedir (Akers ve Doty, 2013, s. 9; MacMillan, 2014).

Bazı araştırma kurumları ve bilimsel dergiler fon desteği alan araştırmaların bir veri yönetim planı içermesini zorunlu tutmasına rağmen veri paylaşımındaki ana problemler varlıklarını korumaktadır (Borgman, 2012). Bunun temel nedenlerinden biri verinin uygun bağlamda yeniden kullanımının (Borgman ve diğerleri, 2012) ve işbirliği yapan ekip dışındakiler tarafından da keşfedilebilirliğinin sağlanması için (Koltay, 2015; Verbaan ve Cox, 2014) veri paylaşımının standart üst veri ve dokümantasyon kullanımını gerektirmesidir. Bunlara ilave olarak, aynı araştırma ekibi içinde veri kavramı ve anlayışı birbirinden çok büyük farklılıklar gösteren disiplinlerden kişilerin olması bile veri paylaşımında sorunlara neden olabilmektedir (Borgman ve diğerleri, 2012).

Temel teknik sorunlardan biri de daha sürdürülebilir yönetim uygulamaları gerektiren veri depolama formatlarının kısalan yaşam süreleridir (MacMillan, 2014). Günümüzde araştırmacılar veri küratörlüğü becerilerinden (örneğin, yeterince anlamlı dokümantasyon

oluşturmak, standartlaşma ve göç uygulamaları hakkında bilgi sahibi olmak ve bunları gerçekleştirmek gibi) yoksundur ve bu konular yüksek öğretimde lisans eğitimleri kapsamına henüz girmemektedir (Frank ve Pharo, 2016). Veri paylaşımı uygulamaları konusunda bir akademik kredi ve ödüllendirme de söz konusu değildir (MacMillan, 2014). Mevcut akademik ödül sistemleri veri değil makale yayınlama üzerine kuruludur. Veriyi paylaşılabilir ve keşfedilebilir hale getirmek daha fazla kaynak, zaman ve emek harcanmasının (Wallis ve diğerleri, 2013) yanı sıra veri yönetimi için yaygın veri yapıları, üst veri standartları ve veri madenciliğini destekleyen ontolojiler gibi araçların kullanımını gerektirmektedir (Borgman, 2012; Koltay, 2015; Verbaan ve Cox, 2014). Diğer bir deyişle disiplinden bağımsız olarak tüm araştırmacılar tarafından paylaşılan ortak bir veri kültürüne ihtiyaç vardır.

Araştırmacıların çoğunun veri yönetimi ve kürasyonu konularında gerekli bilgi ve becerilerden yoksun olması eğitimin önemini artırmakta, ayrıca bu alanda duyulan eğitim gereksiniminin artacağı öngörülmektedir (Carlson ve diğerleri, 2011; Chowdhury ve diğerleri, 2017; Koltay, 2015; Tenopir ve diğerleri, 2016; Wallis ve diğerleri, 2013). Örneğin sağlık alanındaki araştırmacılar üzerinde yapılan bir çalışmada katılımcıların %77'sinin bu alanda hiç formel eğitim almadıkları, bilgi ve beceri düzeylerinin yetersiz olduğu saptanmıştır (Federer, Lu ve Joubert, 2016, s. 54). Kimya alanında 119 araştırmacı ve yüksek lisans öğrencisi üzerinde yürütülen bir çalışma da araştırma verilerinin yönetimi ile ilgili eğitimlere araştırmacıların fazlaca ihtiyaç duyduklarını ortaya koymuştur (Chen ve Wu, 2017). Dünyanın farklı bölgelerinden 2000'den fazla akademik personel ve araştırmacı üzerinde yapılan başka bir çalışma araştırmacıların verilerini ne kadar açtıklarını bilmediklerini göstermiştir. Katılımcıların %60'ı verilerini açma koşullarını dolayısıyla verilerinin ne kadar erişilebilir ve yeniden kullanılabilir olduğunu bilmediklerini belirtmiştir (Fane, Treadway, Gallagher, Penny ve Hahnel, 2016, s. 14). Tüm bu bulgular araştırmacıların veri paylaşımı konusunda farkındalık ve bilgi eksikliğine işaret etmekte ve eğitim konusunun vakit geçirmeden ele alınması gerektiğini göstermektedir (Calzada Prado ve Angel Marzal, 2013; Carlson ve diğerleri, 2011; Chowdhury ve diğerleri, 2017).

Türkiye'de araştırma verilerinin yönetimi ile ilgili politika, strateji ve direktiflerin olmaması nedeniyle araştırma verilerinin büyük bölümü erişilebilir ve yeniden kullanılabilir durumda değildir. TÜBİTAK tarafından hazırlanmakta olan Açık Erişim Politikası'na "araştırma verisi" ve "açık veri" ile ilgili maddelerin eklenmesi yönünde çalışmalar yapıldığı bilinmekle birlikte henüz söz konusu politika taslak halindedir. TÜBİTAK ve üniversiteler de dahil araştırma fonu sağlayan kurumların araştırma verilerinin yönetimi konusunda politikaları ve direktifleri olmadığı gibi fon sağladıkları araştırmacıardan veri yönetim planı gibi beklentileri de yoktur. Bunun sonucunda araştırma kurumları ve üniversitelerde araştırma verilerinin depolanması, analizi ve kürasyonu ile ilgili hizmetler henüz geliştirilmemiş, araştırma verilerini depolamak ve paylaşmak isteyen araştırmacılara yardımcı olabilecek birimler de henüz oluşturulmamıştır (Aydinoğlu, 2016; Tonta, 2013). Diğer taraftan 2012 yılından itibaren konunun önemi konusunda farkındalığı artırmak ve Türkiye'deki durumu yansıtabilmek amacıyla bazı girişimlerde bulunulmuştur. Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü ve Goethe Enstitüsü tarafından 2014 yılında düzenlenen uluslararası Araştırma Verilerinin Yönetimi Çalıştayı¹, yıllık Ulusal Açık Erişim Konferanslarında sunulan bildiriler (örneğin, Aydınoğlu, 2016; Gürdal ve Bitri, 2015), diğer konferanslarda sunulan bazı bildiriler (Aydinoğlu, 2014; Tonta, 2013 gibi) söz konusu girişimlere verilebilecek örneklerdendir. Aydınoğlu, Doğan ve Taşkın (2017) tarafından Türkiye'deki araştırmacılar üzerinde yürütülen çalışma araştırma verilerinin yönetimi konusunda ülkemizdeki araştırmacıların farkındalık düzeyleri ve uygulamaları üzerinde yapılan bilinen ilk iki çalışmadan biridir. Bilimsel yayın performansı açısından en üretken 25 üniversitedeki 532 akademisyen üzerinde yürütülen çalışma, araştırma verilerinin yönetimi konusunda bilgi ve

¹ <http://rdm.bilgiyonetimi.net/>

beceri eksikliğine işaret etmektedir. Diğer çalışma, bu makaleye temel oluşturan ve araştırma verilerinin yönetimi ile ilgili olarak farklı gelişmişlik düzeyinde bulunan sırasıyla Büyük Britanya, Fransa ve Türkiye'deki araştırmacılar üzerinde yürütülmüş karşılaştırmalı bir çalışmadır. Pilot çalışma sonuçları (Chowdhury ve diğerleri, 2017) ve ana çalışmanın (Chowdhury, Ünal, Kurbanoglu, Boustany ve Walton, 2018) sonuçları ayrı ayrı yayımlanmıştır/yayımlanacaktır. Karşılaştırmalarda benzerlikler bulunmuş, bazı durumlarda politikalar ve direktifler açısından daha gelişmiş ülkelerin lehine sonuçlar elde edilmiştir.

Amaç ve Yöntem

Bu çalışmanın amacı, Türkiye'deki araştırmacıların araştırma verilerinin yönetimi, paylaşımı ve yeniden kullanımı ile ilgili uygulamalarının ve farkındalık düzeylerinin belirlenmesi, konuyla ilgili varsa eğitim ihtiyaçlarının saptanması ve araştırma verilerinin yönetimi ile ilgili unsurların disipline göre istatistiksel açıdan anlamlı bir farklılık gösterip göstermediğinin test edilmesidir. Çalışmada araştırmacıların veri yönetimi ile ilgili mevcut durumlarının saptanabilmesi ve bu doğrultuda gerekli önerilerin sunulabilmesi için şu sorulara yanıt aranmıştır:

1. Araştırmacılar tarafından kullanılan ve üretilen verinin türü ve miktarı nedir?
2. Araştırmacılar araştırma verilerini nasıl temin etmektedir?
3. Araştırmacıların veri paylaşımı konusundaki uygulamaları, görüşleri ve kaygıları nelerdir?
4. Araştırmacılar tarafından toplanan/üretilen veri üzerinde ne tür etiketleme ve ön-işlem (veri temizleme gibi) yapılmaktadır?
5. Araştırma verileri nerelerde depolanmaktadır? Uzun süreli depolama için araştırmacıların görüşleri nelerdir?
6. Veri yönetim planı uygulamaları ve araştırmacıların bu konudaki görüşleri nelerdir?
7. Araştırmacıların veri yönetimi sorunları ile ilgili farkındalık düzeyi nedir?
8. Araştırmacıların veri yönetimi ile ilgili aldıkları ve almak istedikleri eğitim konuları nelerdir?
9. Veri yönetimi ile ilgili konularda araştırmacıların çalıştıkları disipline göre bir farklılık var mıdır?

Çalışma için gerekli veriler Ocak-Nisan 2017 tarihleri arasında dağıtılan çevrimiçi anket ile toplanmıştır. Anket dağıtılmadan önce Hacettepe Üniversitesinden Etik Kurul izni alınmış ve anketin diğer üniversitelere de dağıtılabilmesi için gerekli başvurular yapılmıştır. Veri toplama süresi içinde anketin dağıtılmasına izin veren üniversitelerin elektronik listeleri aracılığıyla ya da üniversite kütüphaneleri aracılığıyla araştırmacılara ulaşılmaya çalışılmıştır. Ankete katılım gönüllülük bazında gerçekleşmiş, dolayısıyla çalışmanın örneklemini de gönüllü katılım ile oluşturulmuştur.

Aydınoglu ve diğerleri tarafından 2017 yılında yayımlanan ve bilimsel yayın performansı açısından en üretken 25 üniversitedeki 532 akademisyen üzerinde yapılan çalışmadan farklı olarak bu çalışmada daha geniş bir araştırmacı kitlesine ulaşılmaya çalışılmış ve daha geniş bir yelpazede sorular sorulmuştur. Bazı üniversitelerden katılım düşük olmasına rağmen araştırma kapsamında Türkiye'deki 44 farklı üniversitedeki toplam 540 araştırmacıdan veri toplanmıştır. Örneklemin istatistiksel olarak evreni temsil gücü olmamasına rağmen katılımcı sayısı ve çeşitliliği açısından durum saptaması için eldeki sınırlı bilgiye katkı sağlayacağı düşünülmektedir. Ayrıca araştırmacılar çalışma alanlarına göre fen bilimleri, sosyal bilimler ve beşerî bilimler olarak üç kategori altında toplanarak disiplinler arası farklar test edilmiştir. Söz konusu alanlar çok geniş olmakla birlikte bilgi davranışları ile ilgili geniş literatürden ve veri davranışları ile ilgili sınırlı literatürden de bilindiği üzere bu alanlarda çalışan araştırmacılar ortak bazı özellikler sergilemektedir. Disiplinler arası farklılıkların saptanmasının planlamada faydalı olacağı düşünülmüştür.

Katılımcıların ayrı ayrı sunulan bilim dalları içerisinde kendi araştırma alanlarını seçmeleri istenmiş ancak veri analizi sırasında ilgili yanıtlar Web of Science'ın konu

sınıflaması temel alınarak fen bilimleri (doğa bilimleri, mühendislik ve teknoloji, tıp ve sağlık bilimleri, ziraat gibi), sosyal bilimler (eğitim, sosyoloji, hukuk, siyasal bilimler gibi) ve beşerî bilimler (tarih, arkeoloji, dil ve edebiyat gibi) temel başlıkları altında gruplandırılmıştır.

Katılımcılara uygulanan anket iki bölüm ve toplam 26 sorudan oluşmaktadır. Anketin ilk bölümünde; disiplin, cinsiyet, statü gibi araştırmacıyla ilgili demografik bilgileri toplamak amacıyla hazırlanmış toplam 7 soruya, ikinci bölümünde ise araştırmacıların kullandığı ve yarattığı veri, veri paylaşım uygulamaları ve bu konuda taşıdıkları kaygılar, veri yönetimi uygulamaları ve sorunlarıyla ilgili farkındalık düzeyleri ve eğitim gereksinimleri gibi veri yönetimi ile ilgili toplam 19 soruya yer verilmiştir. Soruların bir kısmı çoktan seçmelidir ve bir kısmı 5’li Likert ölçeğine göre düzenlenmiştir. Veri setinin analizi için SPSS istatistik yazılımı kullanılmış, tanımlayıcı istatistiklerle birlikte, değişkenlerin yapısına uygun olarak *disiplin* ile diğer unsurların tamamı arasında ki-kare ilişki testi uygulanmıştır. Çalışmada sadece ilişki çıkan test sonuçları rapor edilmiştir. 5’ten küçük beklenen sıklık yüzdesinin %20’yi aştığı durumlar için ki-kare test sonucu rapor edilmemiştir.

Bulgular ve Yorum

Genel Bilgiler

Araştırmaya Türkiye’deki 44 farklı üniversiteden toplam 540 araştırmacı katılmıştır. Araştırmacıların %79’u (n=429) dokuz farklı üniversiteden gelmektedir (Tablo 1). Katılımcı sayısı on ve altında olan üniversiteler diğer başlığı altında toplanmıştır ve bu üniversitelerden toplam katılım oranı %21’dir. Katılımcıların %79’u (n=426) akademik personel, %19’u (n=105) lisansüstü öğrencidir. Dokuz kişi (%2) ise kendisini diğer kategorisinde tanımlamıştır.

Tablo 1

Katılımcıların Üniversitelere göre dağılımı

Üniversite	n	%
Hacettepe	95	18
Ankara	84	16
Ankara Yıldırım Beyazıt	82	15
ODTÜ	63	12
Adnan Menderes	35	6
İstanbul	32	6
Kırklareli	15	3
Tokat Gaziosmanpaşa	12	2
Kırıkkale	11	2
Diğer ²	111	21
Toplam	540	101

Not: Yuvarlamadan dolayı toplam %100’den farklıdır.

Araştırmacıların %46’sı (n=246) fen, %38’i (n=203) sosyal ve %17’si (n=91) beşerî bilimler alanında çalışmaktadır. Beşerî bilimler alanında çalışan araştırmacıların anketi cevaplama oranlarının diğerlerine göre daha düşük olduğu görülmektedir. Bu durumun konuyu kendileri ile fazla ilişkilendirememiş olmalarından kaynaklanıyor olabileceği düşünülmektedir.

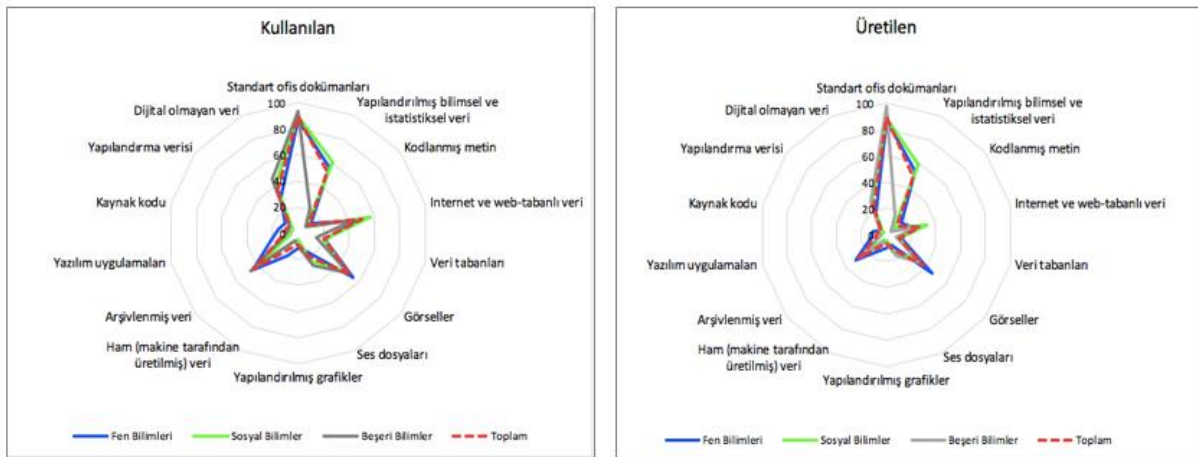
²1-5 arası: Dokuz Eylül, Düzce, Ege, Fırat, Harran, Marmara, Necmettin Erbakan, Nevşehir Hacı Bektaş Veli, İstanbul Sabahattin Zaim, Sakarya, Selçuk, Sinop, TOBB ETÜ, Yaşar, İstanbul Yeni Yüzyıl, Van Yüzüncü Yıl, Bilecik Şeyh Edebali, Erciyes, Gümüşhane, İYTE, Koç, Özyeğin, Çankırı Karatekin, Gazi, Zonguldak Bülent Ecevit, Pamukkale, Iğdır, Ondokuz Mayıs, 6-10 arası: Osmaniye Korkut Ata, Gebze Teknik, Balıkesir, Kahramanmaraş Sütçü İmam, Mimar Sinan Güzel Sanatlar, Bilkent, Bozok, *Belirtilmeyen*: 4

Araştırma Verilerinin Türü ve Miktarı

Tüm disiplinlerde katılımcıların araştırmalarında en çok kullandıkları ve araştırma sırasında en çok ürettikleri dosya türü %89 ile *standart ofis dokümanlarıdır*. Beşerî bilimlerde diğerlerine göre daha düşük olmakla birlikte yine tüm disiplinler için ikinci sırada *yapılandırılmış bilimsel ve istatistiksel veri dosyaları* (kullanım %52, üretim %49) gelmektedir. Kullanımda üçüncü sırada *internet ve web tabanlı veri* gelmektedir (%50). *Kodlanmış metin, yapılandırılmış grafik, yapılandırma verisi, kaynak kodu, ham veri, yazılım uygulamaları* gibi dosya türlerinin en az kullanılan ve üretilen dosya türleri olduğu görülmekle birlikte fen bilimlerindeki kullanım oranları, diğerlerine göre nispeten daha yüksektir (Şekil 1).

Araştırmalarda kullanılan bazı dosya türleri disipline göre farklılık göstermektedir. *Yapılandırılmış bilimsel ve istatistiksel veri* sosyal bilimlerde (%61) ve fen bilimlerinde (%56) beşerî bilimlere (%21) göre çok daha fazla kullanılmaktadır. *Internet ve web-tabanlı* verinin kullanım oranının ise sosyal bilimlerde (%57) diğerlerine (Fen: %46, Beşerî: %43) oranla biraz daha yüksek olduğu görülmektedir. *Yapılandırılmış grafikler* (Sosyal: %5, Fen: %12, Beşerî: %10), *ham veri* (Sosyal: %7, Fen: %20, Beşerî: %7), *yazılım uygulamaları* (Sosyal: %6, Fen: %20, Beşerî: %11), *kaynak kodu* (Sosyal: %5, Fen: %15, Beşerî: %8) ve *yapılandırma verisi* (Sosyal: %5, Fen: %12, Beşerî: %9) gibi teknik dosya türleri fen bilimlerinde diğerlerine oranla daha fazla, *ses dosyaları* (Sosyal: %24, Fen: %15 Beşerî: %27) ve *dijital olmayan veri* (Sosyal: %41, Fen: %30, Beşerî: %46) ise daha az kullanılmaktadır.

Araştırmalarda üretilen bazı dosya türleri de disipline göre farklılık göstermektedir. *Standart ofis dokümanları* tüm bilim dallarında en çok üretilen dosya türü olmakla birlikte en yüksek üretim oranı beşerî bilimlere aittir (%98). *Yapılandırılmış bilimsel ve istatistiksel verinin* üretim miktarı beşerî bilimlerde (%17), sosyal (%59) ve fen (%53) bilimlerine göre oldukça düşüktür. Kullanımda olduğu gibi *kodlanmış metin* (Sosyal: %8, Fen: %15, Beşerî: %4), *görsel veri* (Sosyal: %29, Fen: %46, Beşerî: %33), *ham veri* (Sosyal: %4, Fen: %14, Beşerî: %7), *yazılım uygulamaları* (Sosyal: %5, Fen: %12, Beşerî: %4) ve *kaynak kodu* (Sosyal: %3, Fen: %11, Beşerî: %5) türündeki veri dosyaları fen bilimlerinde diğerlerine oranla daha çok, *ses dosyaları* (Sosyal: %15, Fen: %8, Beşerî: %18) ise daha az üretilmektedir.

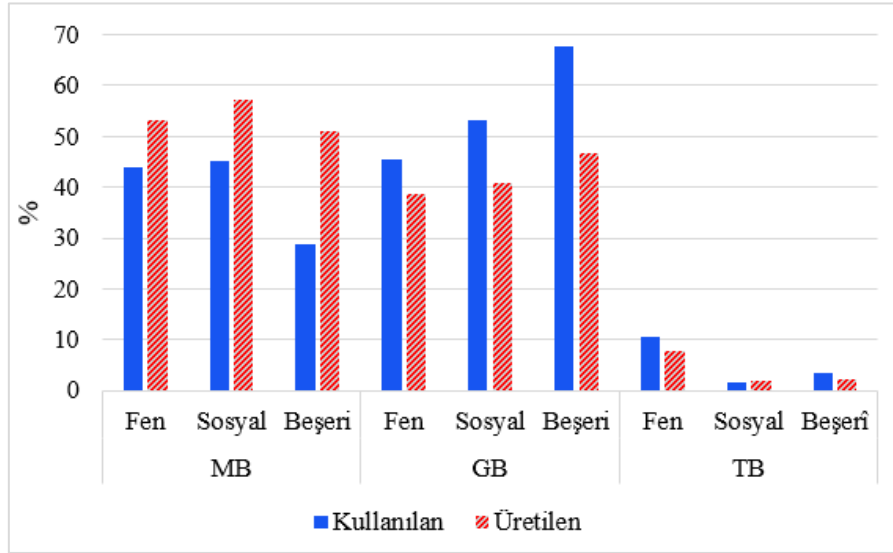


Şekil 1. Bilim alanlarına göre en çok kullanılan ve üretilen dosya türleri (%)

Kullanılan veri miktarında %52 ile GB (gigabyte), üretilen veri miktarında ise %54 ile MB (megabyte) ilk sırada gelmektedir. Toplamda kullanım (%6) ve üretim (%5) oranı oldukça düşük olan TB (terabyte) düzeyindeki verinin, fen bilimlerindeki kullanım ve üretimi diğer bilim dallarına göre biraz daha yüksektir (sırasıyla %11, %8). Bu durum büyük veriyle çalışan bilim dallarının genellikle fen bilimlerinde toplanması ile açıklanabilir. Üretilen veri miktarının en yüksek MB düzeyinde olması araştırmacılarımızın yarıdan fazlasının veri yoğun

çalışmadığının bir göstergesi olarak yorumlanabilir (Şekil 2). Araştırmalarda kullanılan veri miktarı disipline göre bazı küçük farklılıklar göstermektedir. Söz konusu farklılık toplamda en düşük kullanım oranına sahip TB büyüklüğündeki veri miktarının %11 ile en yüksek fen bilimlerinde, yine toplam içerisinde en yüksek kullanım oranına sahip GB düzeyindeki veri miktarının ise en çok (%68) beşerî bilimlerde kullanılması ile açıklanabilir.

Kullandıkları ve ürettikleri veri türü ile veri miktarı araştırmacıların veri nosyonunu etkilemekte ve söz konusu veri türünün yönetimine yönelik spesifik bilgi ve destek ihtiyaçları doğmasına neden olabilmektedir. Bu durumun gerek destek birimi hizmetlerinin planlamasında ve gerekse eğitim programlarının geliştirilmesinde göz önüne alınmasında fayda vardır.



Şekil 2. Araştırmalarda kullanılan ve üretilen veri miktarı

Verinin Kaynağı, Dış Kaynaklı Verinin Kullanımı ve Veri Depolama

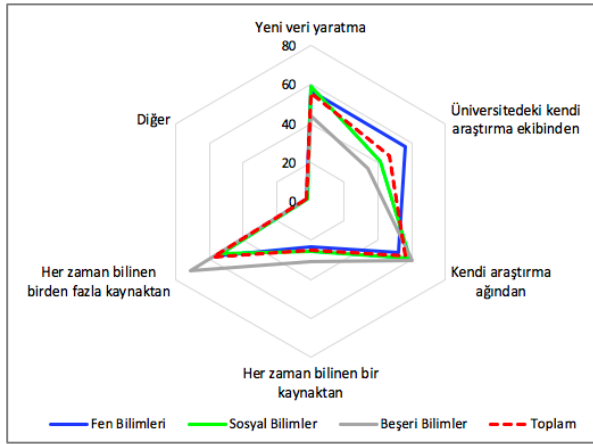
Bulgular araştırma verilerinin farklı kaynaklardan elde edildiğini ve katılımcıların genelde bilinen birden fazla kaynaktan veri elde ettiğini (%58) göstermektedir. Katılımcıların yarıdan fazlası (%55) kendi verisini ürettiğini belirtmiştir. Sahip olduğu araştırma ağı (%56), üniversite içindeki araştırma ekibi (%47) veri temin edilen diğer kaynaklar arasındadır (Şekil 3). New South Wales’da (Avusturalya) bulunan on farklı üniversitedeki akademisyenlerin araştırma verilerinin yönetimi ve paylaşımı ile ilgili uygulamaları üzerine yapılan bir çalışmada da veri sağlamada araştırmacıların %76,9’unun sıklıkla kendi verisini topladığı ya da oluşturduğu, %53,6’sının ise bir araştırma ekibinin parçası olarak veri topladığı ya da oluşturduğu görülmüştür (Kenan ve Markauskaite, 2015, s. 80). Katılımcıların sadece %10’u dış kaynaklardan veri kullanmadığını belirtmiştir. Dış kaynaklı veri kullanımında araştırmacıların %26’sı sağladıkları veriyi olduğu gibi sorunsuz bir şekilde kullandıklarını, %52’si dış kaynaklı veriyi kullanmadan önce üzerinde belli bir miktarda çaba harcadığını, %46’sı ise dış kaynaklı veriyi üzerinde yoğun çaba ve zaman harcadıktan sonra kullanabildiğini belirtmiştir³ (Şekil 4). Araştırmacıların büyük bölümünün dış kaynaklı veriyi üzerinde az veya çok çaba harcadıktan sonra kullanabildiklerini belirtmiş olması dış kaynaklı verinin doğrudan kullanılabilirlik açısından bazı sorunlar taşıdığını göstermektedir.

Araştırmalarda kullanılan verinin sağlanma şekillerinden bazıları %95 güven düzeyinde disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir. Beşerî bilimlerdeki araştırmacılar diğer araştırmacılara oranla yeni veri yaratma ($\chi^2(2) = 6,077; p=0,048$) ve üniversitedeki kendi araştırma ekibinden ($\chi^2(2) = 17,604; p=0,000$) veri sağlamayı daha az, her

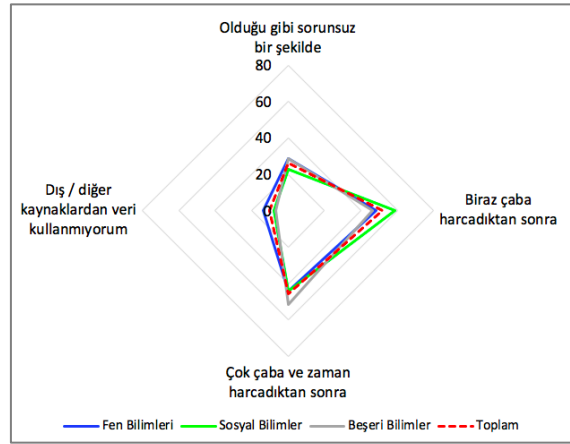
³ Bu bölümdeki sorularda birden fazla seçenek işaretlenebildiği için toplamlar %100’ün üzerindedir.

zaman bilinen birden fazla kaynaktan veri sağlamayı ($\chi^2(2) = 8,960; p=0,011$) daha çok tercih etmektedirler. Bulgular bir yandan beşerî bilimlerdeki araştırmacıların bireysel çalışma eğilimlerini destekler görünürken diğer yandan kendi verilerini üretmek yerine bildikleri farklı kaynaklardan veri sağlamayı tercih ediyor olmaları veri (data) kavramının bilgi (information) ile karıştırılmış olabileceği kaygısını yaratmaktadır. Görüşme yoluyla yapılacak başka çalışmalarla kavram kargaşasından kaynaklandığı düşünülen bu durumun açıklığa kavuşturulması faydalı olacaktır.

Dış kaynaklardan sağlanan veri üzerinde *belli bir miktar çaba harcadıktan sonra kullanım* ile disiplin arasında bir ilişki görülmekte ve bu durumun sosyal bilim araştırmacılarından kaynaklandığı anlaşılmaktadır. Dış kaynaklardan sağlanan veriyi kullanmadan önce üzerinde temizleme ya da değişiklik için *biraz çaba harcadığını* belirtenlerin oranı %59 ile sosyal bilimlerde diğerlerinden daha yüksektir.



Şekil 3. Verinin kaynağı



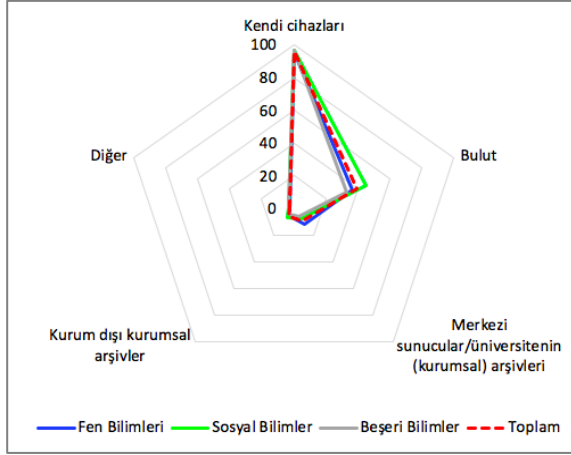
Şekil 4. Dış kaynaklı verinin kullanımı

Araştırmacıların neredeyse tamamı (%96) ürettikleri veriyi kendi cihazlarında saklamaktadır. Çok yüksek oranlarda olmamakla birlikte bulutta depolamak (%39) ikinci sırada gelmektedir ve sosyal bilimlerde diğerlerine göre nispeten biraz daha yüksektir (%45). Bulgular araştırma verilerinin depolanmasında birden fazla yöntemin aynı anda kullanıldığını göstermektedir. Merkezi sunucu ya da üniversitelerin kurumsal arşivleri (%9) ile kurum dışı kurumsal arşivlerin (%6) kullanım oranları ise oldukça düşüktür (Şekil 5). Konuyla ilgili literatürdeki benzer araştırmalar da (Akers ve Doty, 2013; Aydınoglu ve diğerleri, 2017; Chen ve Wu, 2017; Kenan ve Markauskaite, 2015) araştırmacıların veriyi depolamak/saklamak için ilk sıralarda kendi kişisel bilgisayarlarını ya da bilgisayarların sabit ya da harici disklerini (USB sürücüler de dâhil) tercih ettiklerini, kurumsal arşivlerin (ulusal, disiplin temelli, kurumsal, ticari) tercih oranlarının ise oldukça düşük olduğunu göstermektedir. Literatürdeki bulgularla da örtüşen bu durum ülkemizde kurumsal arşivlerin henüz yeterince gelişmemiş olması ile ilişkilendirilebileceği gibi kurumsal arşivlerin daha çok araştırma sürecinin son ürünü olan bilimsel yayımları depolamak amacıyla kullanılması ve veri depolamanın ayrı bir alt yapı ve uzmanlık gerektirmesi ile de açıklanabilir.

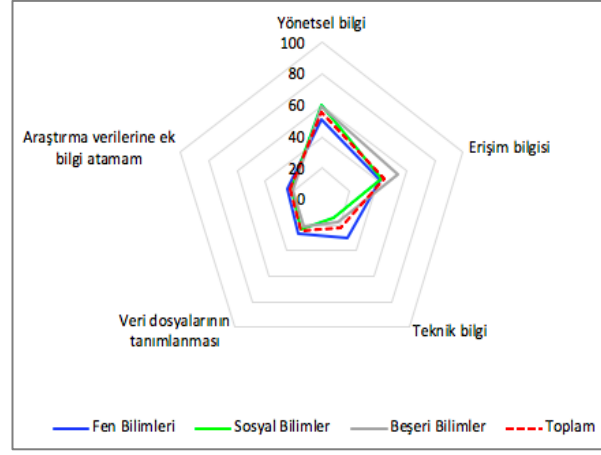
Araştırmacıların %23'ü depolarken verilerine hiçbir ek bilgi atamadıklarını belirtmişlerdir. Yarıdan fazlası (%56) veriye yönetsel bilgi atarken, %44'ü erişim bilgisi atamaktadır. Teknik bilgi ve veri dosyalarının tanımlanması türündeki ek bilgi atamalarının daha düşük olduğu görülmektedir (sırasıyla %22, %24). Teknik bilgi atama fen bilimlerinde, erişim bilgisi atama ise beşerî bilimlerde diğerlerine göre nispeten biraz daha yüksektir (Şekil 6). Araştırma verisine teknik ek bilgi atama disipline göre istatistiksel açıdan düşük düzeyde bir farklılık göstermekte ($\chi^2(2) = 16,778; p=0,000$) ve farklılığın fen bilimlerindeki

araştırmacıların diğerlerine oranla biraz daha yüksek oranda (%29) bu ek bilgi türünü atamasından kaynaklandığı anlaşılmaktadır.

Araştırma verilerinin depolanması ile ilgili bulgular açık erişim ve paylaşılabilirlik açısından önem taşıyan erişilebilirlik, keşfedilebilirlik ve kullanılabilirlik konularında sorunlar olduğunu göstermektedir.



Şekil 5. Araştırma verilerini depolama ortamları



Şekil 6. Araştırma verilerine ek bilgi atama

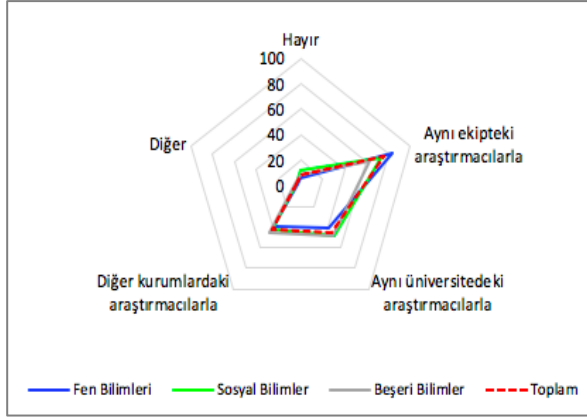
Araştırma Verilerinin Paylaşımı ve Erişimi

Araştırmacıların %9'u veri paylaşımında işbirliği yapmadıklarını belirtmişlerdir. Sosyal bilimlerde (%12) veri paylaşımında işbirliği yapmadıklarını belirtenlerin oranı fen (%7) ve beşerî (%9) bilimlere oranla biraz daha yüksektir. Araştırmacıların veri paylaşımında işbirliği konusuna genelde olumlu olmakla birlikte biraz temkinli yaklaştıkları görülmektedir. Katılımcıların %76'sı kendi ekibindeki diğer araştırmacılarla, yaklaşık yarısı da (%45) aynı üniversitedekilerle veri paylaşımında işbirliği yaptığını belirtmiştir. Diğer kurumlardaki araştırmacılarla veri paylaşımında işbirliği yaptığını belirtenlerin oranı ise daha düşüktür (%42) (Şekil 7). Aynı ekipteki araştırmacılarla işbirliği yapma %95 güven düzeyinde disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir ($\chi^2(2) = 16,964; p=0,000$). Farklılık fen bilimlerindeki araştırmacıların (%83) beşerî bilimlerdeki araştırmacılara göre (%63) daha yüksek oranda işbirliği yapması ile, bu da genelde fen bilimlerinde daha çok ekip çalışması yapılması ile açıklanabilir. Türkiye'deki en üretken 25 üniversitedeki akademisyenler üzerinde yapılan Aydinoglu ve diğerlerinin (2017, s. 278) çalışmasında veri paylaşımı yapmayacağını belirtenlerin oranı daha yüksektir (ilgili soruya yanıt verenlerin %37,4'ü).

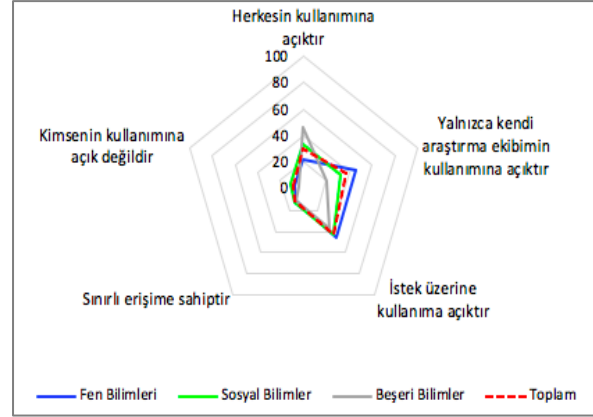
Veri paylaşımında işbirliği yapma ile ilgili bulguları doğrular şekilde araştırmacıların %8'i verilerinin kimsenin erişimine/kullanımına açık olmadığını belirtmiştir. Verilerin en çok (%44) istek üzerine başkalarının erişimine ve kullanımına açıldığı anlaşılmaktadır. Verilerinin herkesin kullanımına açık olduğunu belirtenlerin oranı %30'dur (Şekil 8). Bu konudaki en yüksek oranın beşerî bilimlerde (%46), en düşük oranın ise fen bilimlerinde (%22) olduğu görülmektedir ve bu farklılık disipline göre istatistiksel açıdan anlamlıdır ($\chi^2(2) = 20,783; p=0,000$). Kurumsal arşivlerin henüz veri depolama konusunda devreye girmediği, veri depolama konusunda yönlendirme yapan ve destek veren birimlerin üniversitelerde henüz oluşmadığı, araştırmacıların verilerini ağırlıklı olarak kişisel bilgisayarlarında sakladığı ve verilerine sınırlı ek bilgi atadıkları bir ortamda verilerin herkesin kullanımına açılması ile araştırmacıların ne kastettiği ve hangi koşullarda verilerini açtıkları ayrıca araştırılması gereken bir durumdur. Benzer şekilde Emory Üniversitesindeki akademisyenlerin büyük çoğunluğunun araştırma grubu dışındaki kişilerle veri paylaşımı yapmadığı, çok az sayıda araştırmacının geniş bir kitle ile veri paylaşımına istekli

olduğu ancak disiplin bazında değerlendirildiğinde herkes ile paylaşma konusunda beşerî bilimlerdeki araştırmacıların diğerlerine kıyasla veri paylaşımında daha istekli olduğu belirlenmiştir (Akers ve Doty, 2013, s. 9-10).

Araştırmacıların %12'si verilerinin sınırlı erişime açık olduğunu belirtmiştir (Şekil 8). Verilerin yalnızca araştırmacıların kendi araştırma ekibinin kullanımına açık olması da ($\chi^2(2) = 19,989$; $p=0,000$) disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir. Farklılık, muhtemelen daha çok ortak çalışma yaptıkları için, fen bilimlerindeki araştırmacıların (%46), sosyal (%33) ve beşerî (%21) bilimlere göre kendi araştırma ekibine verilerini açması gerekliliğinin daha yüksek olması ile açıklanabilir.

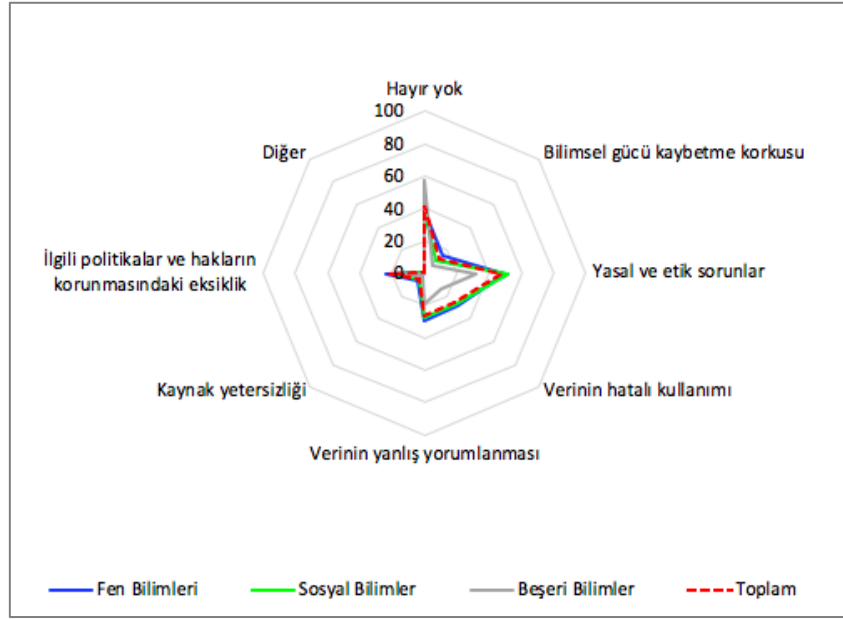


Şekil 7. Veri paylaşımında işbirliği



Şekil 8. Araştırma verilerine erişim

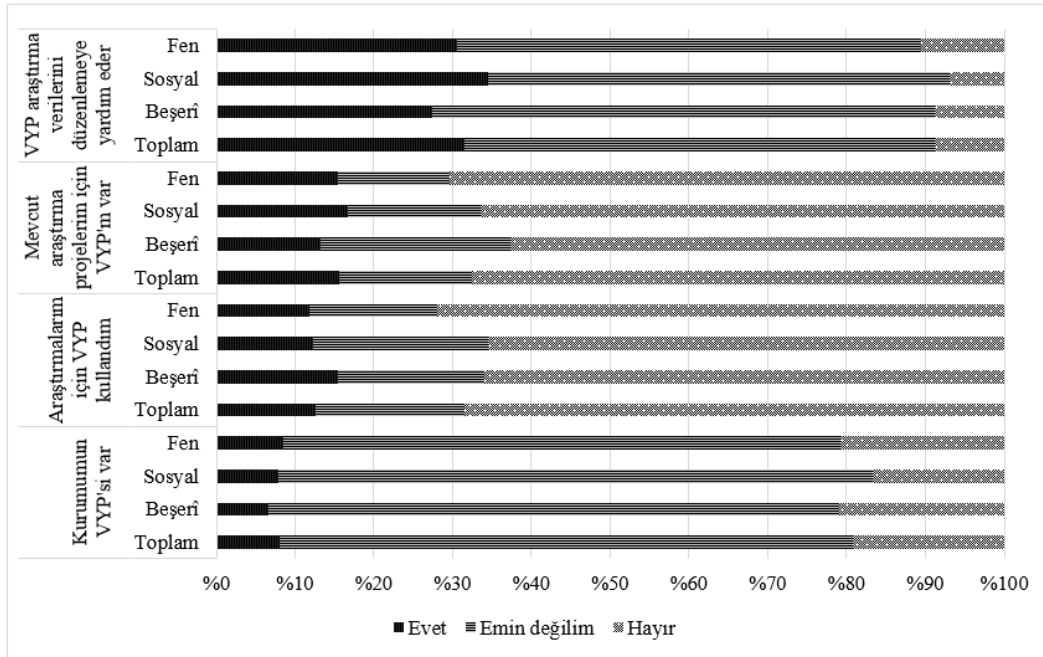
Araştırmacıların yarısından fazlasının veri paylaşımıyla ilgili birtakım endişelerinin olduğu görülmektedir. Bu endişelerin başında %47 ile *yasal ve etik sorunlar* gelmektedir. *Verilerin yanlış yorumlanması* (%27) ve *hatalı kullanımı* (%26), ilgili politika eksikliği ve buna bağlı olarak hakların korunması konusundaki eksiklikler (%22) diğer endişeler arasındadır. Araştırmacıların %42'sinin ise bu konuda herhangi bir endişesinin olmadığı görülmektedir. En az endişe duyan grup %57 ile beşerî bilimlerdir (Şekil 9). Veri paylaşımı konusundaki endişelerden *bilimsel gücü kaybetme korkusu* ($\chi^2(2) = 6,378$; $p=0,041$), *yasal ve etik sorunlar* ($\chi^2(2) = 10,295$; $p=0,006$), *verinin hatalı kullanımı* ($\chi^2(2) = 7,557$; $p=0,023$) ile bu konuda *herhangi bir endişe olmaması* ($\chi^2(2) = 10,787$; $p=0,005$) %95 güven düzeyinde disipline göre istatistiksel açıdan anlamlı farklılık göstermektedir. Beşerî bilimlerdeki araştırmacıların verilerini herkese açmak konusunda daha olumlu eğilim göstermesi ve daha az endişe taşıması, çalıştıkları veri türü ve veri kavramının farklı yorumlanması ile açıklanabilir. Bu konuda yorum yapılabilmesi için daha fazla araştırmaya gereksinim vardır. Veri paylaşımı ile ilgili endişe duyanlar sadece Türkiye'deki araştırmacılar değildir. Örneğin Oxford Üniversitesindeki araştırmacıların yanıtladığı anketin sonuçları da göstermiştir ki yasal, etik ve ticari kaygılar birçok araştırma projesinde veri paylaşımı ile ilgili önemli bir engel oluşturmaktadır (University of Oxford Research Data Management Survey 2012: The Results, 2013).



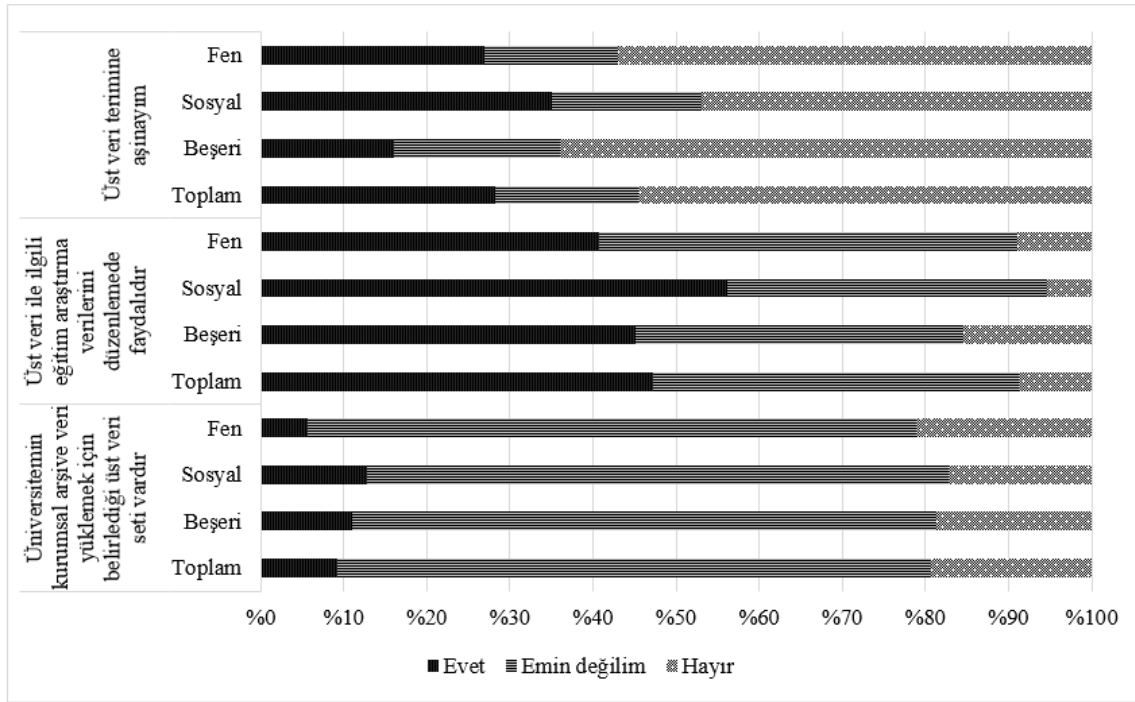
Şekil 9. Veri paylaşımı ile ilgili endişeler

Araştırma Verilerinin Yönetimi: Uygulamalar ve Farkındalık

Katılımcıların %60'ı veri yönetim planının (VYP) araştırma verilerini düzenlemede yardımcı olacağı konusunda emin olmadığını belirtmiştir. Söz konusu oranın yüksekliği VYP ile ilgili bilgi eksikliğinden kaynaklanıyor olabilir. Bilim dalları arasında konuyla ilgili farkındalığın birbirine çok yakın olması da ayrıca dikkat çekicidir. Araştırmacıların sadece %16'sı mevcut araştırma projeleri için bir VYP'sinin olduğunu, %13'ü de geçmiş araştırmalarında VYP kullandığını belirtmiştir. Söz konusu oranlarda bilim dallarına göre belirgin bir farklılık görülmemektedir. Kurumlarının bir VYP'si olduğunu belirten araştırmacı oranı %8, kurumlarının bir VYP'si olup olmadığından emin olmayanların oranı ise %73'tür (Şekil 10). Aydınoglu ve diğerlerinin (2017) çalışmasında da benzer bir sonuç elde edilmiş, araştırmacıların sadece %6,1'i kurumlarında VYP'nin zorunlu olduğunu belirtmiştir.



Şekil 10. Veri yönetimi planı (VYP) ile ilgili farkındalık

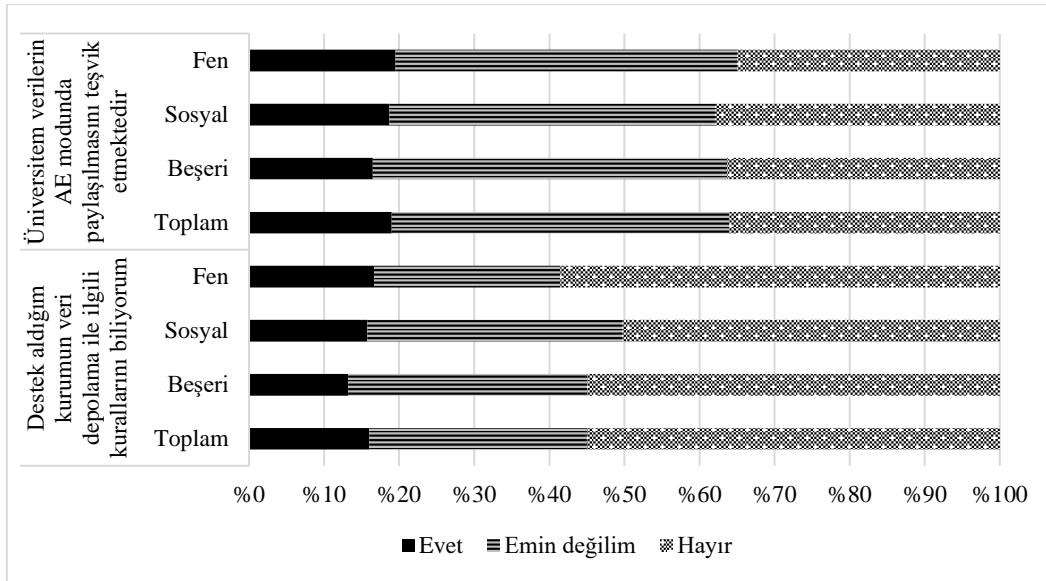


Şekil 11. Üst veri ile ilgili farkındalık

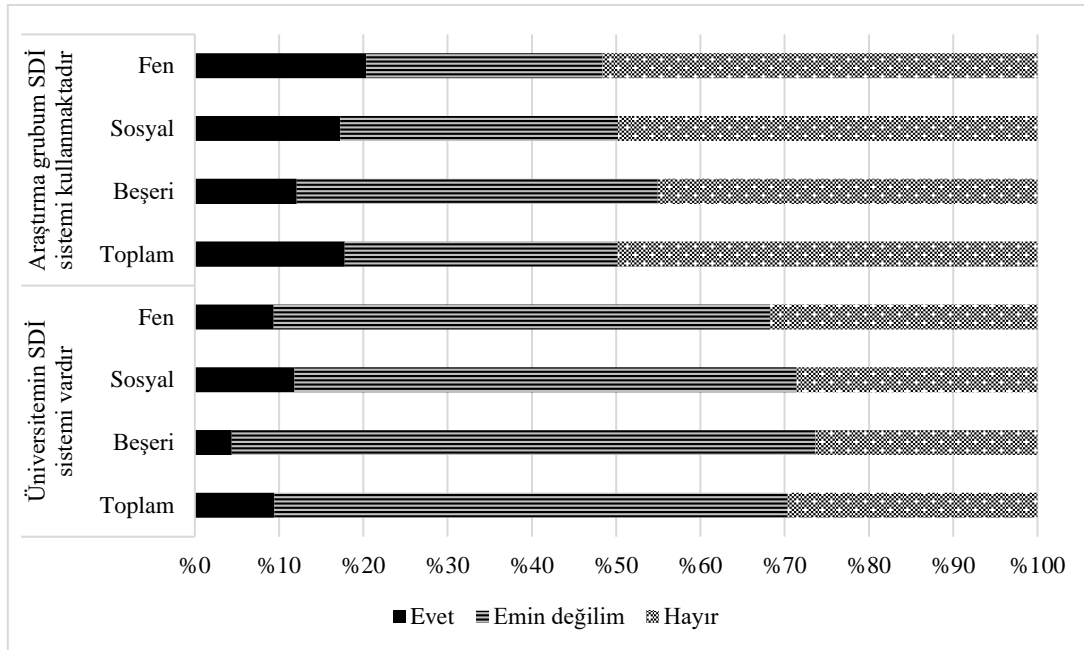
Araştırmacıların %54'ü üst veri terimine aşına olmadığını belirtirken, %17'si emin olmadığını ve sadece %28'i terime aşına olduğunu belirtmiştir. En düşük aşinalık oranı beşeri bilimlerde (%16) ve bu durum disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir ($\chi^2(4) = 12,046; p = 0,017$). Aydınoglu ve diğerlerinin (2017) araştırmasında üst veri hakkında hiçbir şey bilmediğini belirtenlerin oranı %27,1 ile çok daha düşüktür. Veri toplanan üniversitelerin bazıları her iki çalışmada da ortak olmasına rağmen, Aydınoglu ve diğerlerinin (2017) çalışmasında üst veri vb. konularda bilinçli ya da bilgili araştırmacı sayısının daha yüksek olması hem ilgili araştırmanın örnekleminin farklı olması hem de sadece Türkiye'deki en üretken 25 üniversitede yürütülmüş olması ile açıklanabilir.

Araştırmacıların yaklaşık yarısı (%47) üst veri konusundaki bir eğitimin araştırma verilerini düzenleme konusunda faydalı olacağını düşünürken, %44'ü emin olmadığını belirtmiştir. Emin olmama oranının bu kadar yüksek olması muhtemelen bu araştırmacıların kavram ile aşına olmamalarından kaynaklanmaktadır. Konuyla ilgili en yüksek olumlu görüş (%56) sosyal bilimlerdeki araştırmacılara aittir ve üst veri eğitimi konusundaki görüş disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir ($\chi^2(4) = 17,102; p = 0,002$). Katılımcıların büyük çoğunluğu (%71) üniversitelerinin kurumsal arşive veri yüklemek için belirlediği bir üst veri standardı olup olmadığını bilmediklerini belirtirken %19'u bu soruya olumsuz yanıt vermiştir (Şekil 11).

Araştırmacıların %36'sı üniversitelerinin verilerin erişime açılmasını teşvik etmediğini, %45'i ise bu konuda bilgisi olmadığını belirtmiştir. Araştırmacıların %55'i araştırma desteği aldığı kurumun veri depolama ile ilgili kurallarını bilmediğini, %29'u ise bundan emin olmadığını beyan etmiştir. Her iki konuyla ilgili olarak bilim dallarına göre çok büyük farklılıktan söz etmek mümkün değildir (Şekil 12). Ülkemiz üniversitelerinde kurum çalışanlarınca yapılan yayınlara açık erişim tartışılırken, bu konudaki gelişmelerin yavaş seyrettiği ve verilere açık erişim konusunun henüz bu tartışmalarda ve özellikle uygulamalarda çok fazla yer bulamadığı da bilinmektedir. Ayrıca ulusal düzeyde araştırma desteği veren kurumların da henüz veri depolama konusunda kurallar koymadığı ve yaptırımlar uygulamadığı bilinmektedir (Aydınoglu, 2016; Tonta, 2013). Dolayısıyla burada verilen olumlu yanıtların yorumlanması ancak farklı veri toplama teknikleri ile konunun irdelenmesi ile mümkün olacaktır.



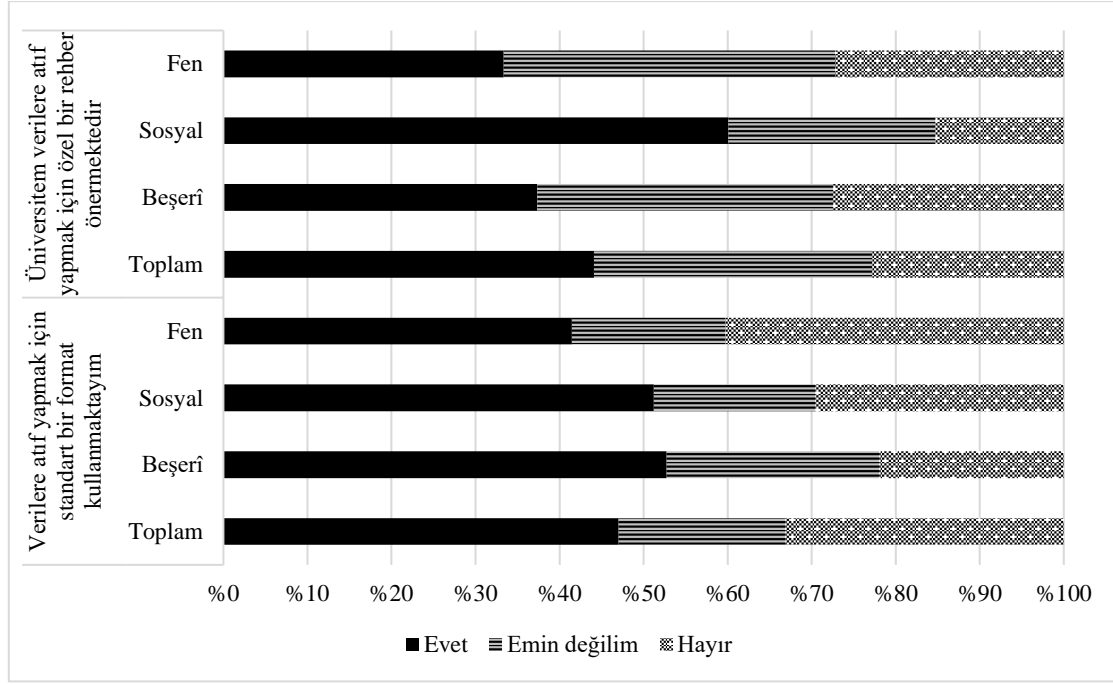
Şekil 12. Açık erişim politikalarında veri paylaşımı ve veri depolama ile ilgili farkındalık



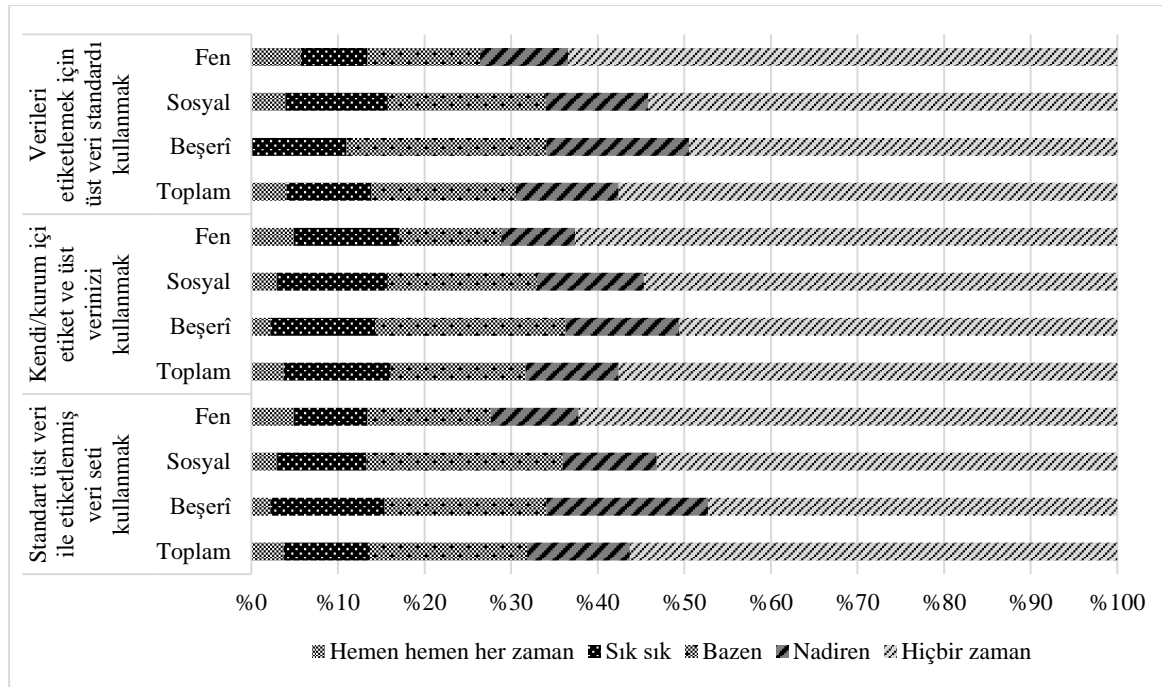
Şekil 13. Standart dosya isimlendirme (SDİ) ile ilgili farkındalık

Araştırmacıların sadece %18'i araştırma grubu içinde bir standart dosya isimlendirme (SDİ) sisteminin kullanıldığını belirtmiştir. Üniversitelerinin SDİ sistemi olduğunu belirten araştırmacı oranı ise sadece %9'dur ve büyük çoğunluk (%61) bu konuda bilgi sahibi değildir (Şekil 13). Katılımcıların yarıdan fazlasının VYP'nin fonksiyonunu ve üniversitelerinin SDİ sistemi olup olmadığını bilmemesi, üst veri terimine aşına olmaması ve veri etiketlemek için üst veri standardı kullanmaması bu konularda bilgi ve farkındalık eksikliğine işaret etmektedir.

Araştırmacıların %47'si verilere atıf yapmak için standart bir format kullandığını belirtmiştir. Üniversitesinin bu konuda özel bir rehber/kılavuz önerdiğini belirten araştırmacı oranı ise toplamda %44'dür (Şekil 14). Ancak burada açık veri, veri paylaşımı ve veriye atıf konularının henüz konuşulmaya başlandığı bir ortamda bu soruda yer alan atıfın daha genel anlamıyla veriden çok bilgiye atıfla karıştırılmış olabileceği düşünülmektedir.



Şekil 14. Atıf standardı kullanımı



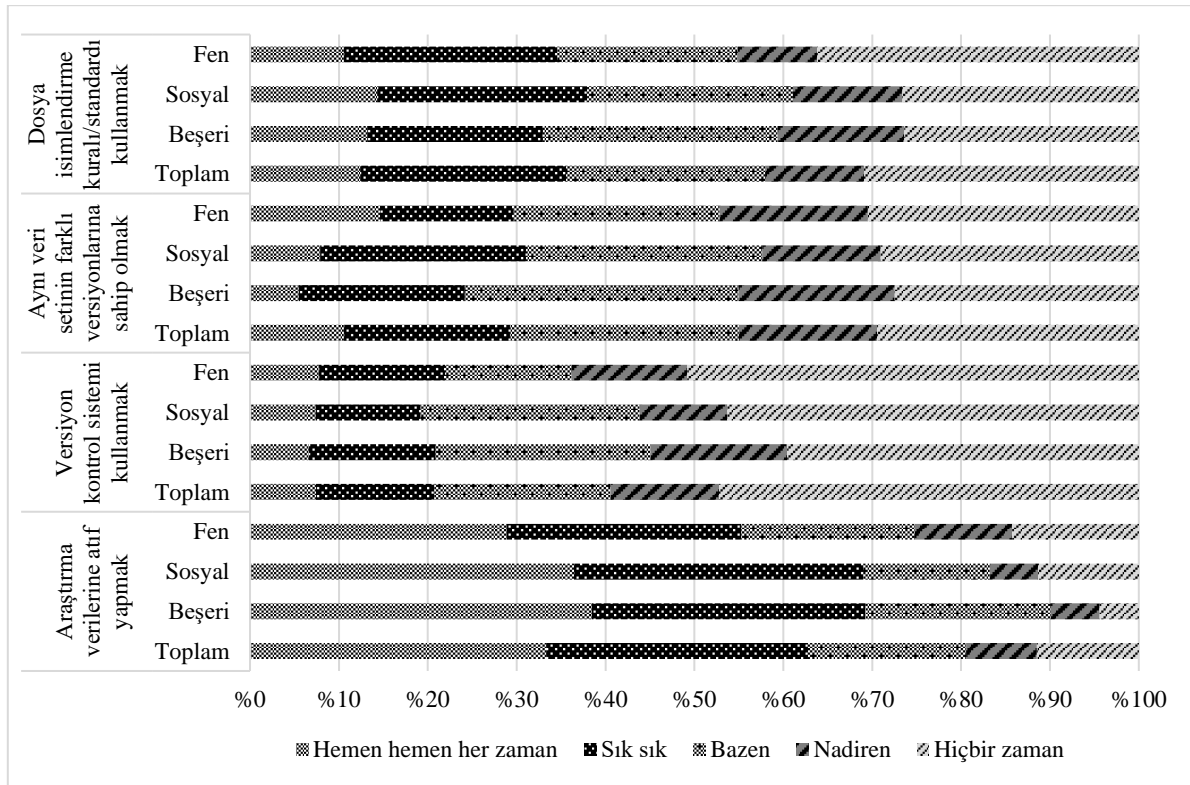
Şekil 15. Üst veri standartlarının kullanımı

Araştırmacıların yarısından fazlası (%58) veri etiketlemek için hiçbir zaman herhangi bir üst veri standardı kullanmadığını belirtmiştir (Aydınoglu ve diğerlerinin (2017) çalışmasında üst veri kullanım oranı %36,4 olarak belirlenmiştir) ve kullanım disiplini göre istatistiksel açıdan anlamlı bir farklılık göstermektedir ($\chi^2(8) = 17,012; p=0,030$). Herhangi bir üst veri standardını hiçbir zaman kullanmadığını belirten araştırmacıların oranı fen bilimlerinde (%63) sosyal (%54) ve beşerî (%49) bilimlere göre daha yüksektir. Aynı şekilde araştırmacıların %58'i hiçbir zaman ne kendi ne de kurum içi etiket ve üst veri standardını kullanmadığını belirtmiştir ve bu oran fen bilimlerinde (%63) sosyal (%55) ve beşerî (%51) bilimlere göre

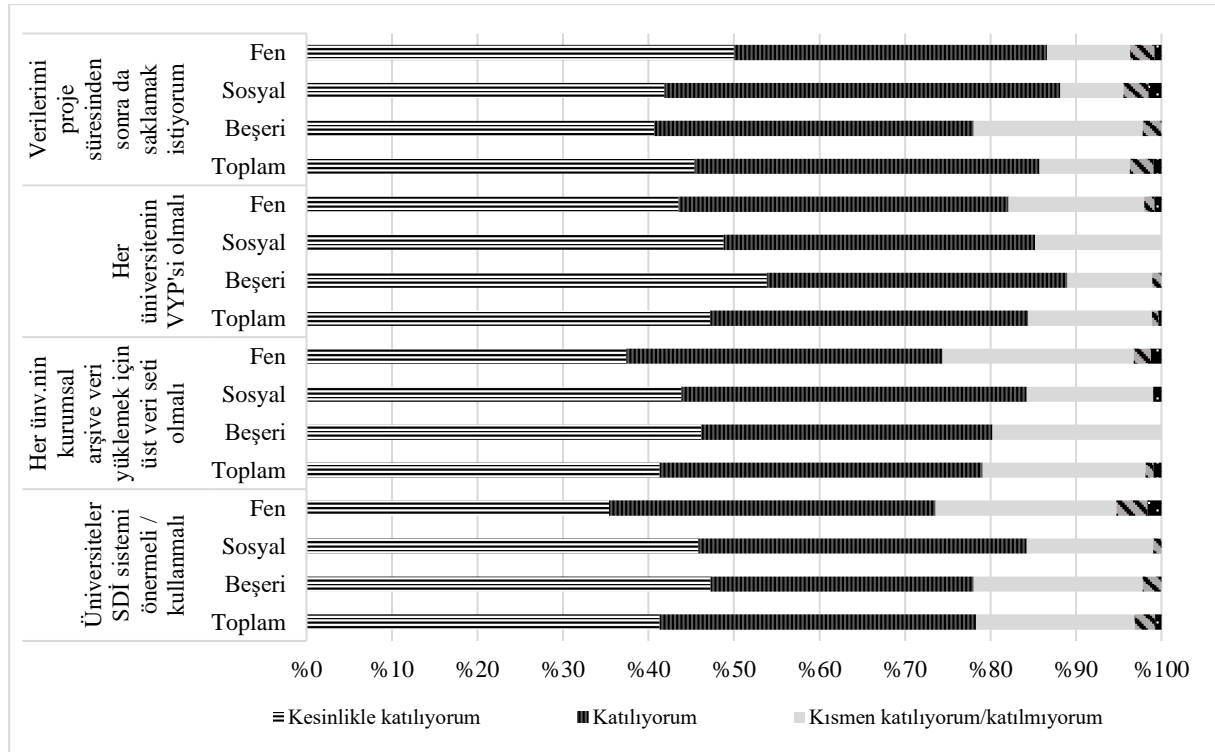
nispeten daha yüksektir. Araştırmacıların yarısından fazlasının veri etiketlemek için herhangi bir üst veri standardı kullanmadığı gibi *standart bir üst veri ile etiketlenmiş veri setlerini* de kullanmadıkları (%56) ve bu oranın fen bilimlerinde (%62) beşerî bilimlere göre (%47) daha yüksek olduğu görülmektedir (Şekil 15).

Araştırmacıların %35'i hemen hemen her zaman ya da sık sık dosya isimlendirme standardı kullandığını belirtmiştir. Hiçbir zaman dosya isimlendirme standardı kullanmadığını belirten araştırmacı oranı fen bilimlerinde (%36) daha yüksektir (Şekil 16). Hemen hemen her zaman ya da sık sık aynı veri setinin farklı versiyonlarına sahip olan araştırmacı sayısı toplamda %30'dur. Araştırmacıların %29'u ise kullandıkları veri setleri için hiçbir zaman farklı versiyon oluşturmadıklarını belirtmiştir. Her zaman farklı versiyon oluşturma yoluna giden araştırmacı oranı fen bilimlerinde (%15) sosyal (%8) ve beşerî (%5) bilimlerden daha yüksektir. Araştırmacıların sadece %20'si hemen hemen her zaman ya da sık sık versiyon kontrol sistemi kullandığını belirtirken yaklaşık yarısı (%47) hiç versiyon kontrol sistemi kullanmadığını belirtmiştir (Şekil 16).

Araştırmacıların %62'si hemen hemen her zaman ya da sık sık araştırma verilerine atıf yaptığını belirtmiştir. Bu oranın yüksek olması, diğer sorulara verilen cevaplara bakıldığında yukarıda da belirtildiği gibi araştırmacıların, araştırma verisine atıf yapma ile bilgi kaynaklarına atıf yapma arasında bir fark görmediği şeklinde yorumlanabilir. Araştırma verilerine atıf yapmak %95 güven düzeyinde disipline göre istatistiksel açıdan anlamlı bir farklılık göstermektedir ($\chi^2(8) = 17,303; p = 0,027$). Farklılık hiçbir zaman atıf yapmadığını belirtenlerin oranının fen bilimlerinde %14, sosyal bilimlerde %11 iken beşerî bilimlerde sadece %4 olması ile açıklanabilir (Şekil 16).



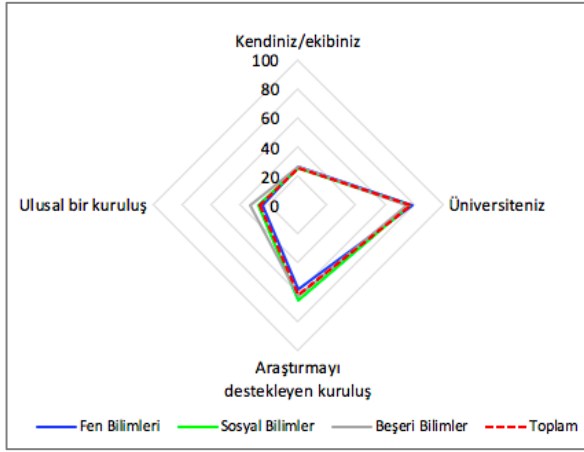
Şekil 16. Dosya isimlendirme, versiyon ve atıf uygulamaları



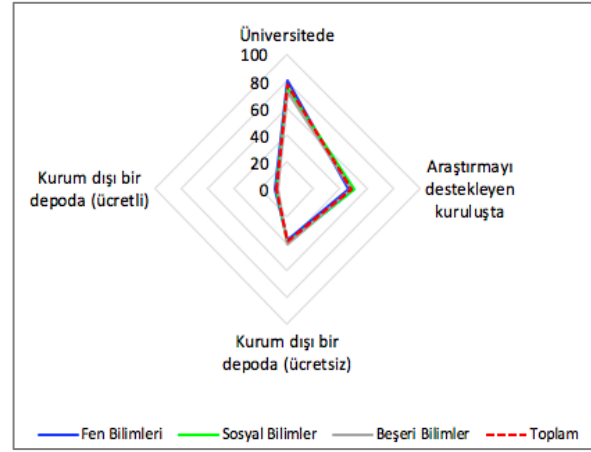
Şekil 17. Veri yönetim planı (VYP), üst veri, standart dosya isimlendirme (SDİ)

Araştırmacıların %85'i (kesinlikle katılıyorum ve katılıyorum seçeneklerini işaretleyenler) verilerini proje süresinden sonra da saklamak istemektedirler. Katılımcıların büyük bir kısmı üniversitelerin veri yönetimi ile ilgili bazı sorumlulukları yerine getirmesi gerektiğini düşünmektedir. Örneğin %84'ü her üniversitenin VYP'si olması, %79'u her üniversitenin kurumsal arşive veri yüklemek için üst veri seti olması ve %78'i üniversitelerin SDİ sistemi önermesi/kullanması gerektiğini (kesinlikle katılıyorum ve katılıyorum seçeneklerini işaretleyenler) belirtmiştir (Şekil 17).

Araştırmacıların %78'i veri setlerinin depolanması ya da erişime açılması konusundaki sorumluluğu üniversitelerin, %62'si ise araştırmayı destekleyen kuruluşun alması gerektiğini düşünmektedir. Sorumluluğu kendilerinin ya da ulusal bir kuruluşun alması gerektiğini düşünenlerin oranı ise oldukça düşüktür (sırasıyla %26, %27). Konuyla ilgili görüşlerde bilim dallarına göre önemli bir fark görülmemektedir (Şekil 18). Araştırmacıların sorumluluk konusunda olduğu gibi, uzun süreli depolama yeri olarak da tercihleri üniversitelerden yanadır (%77). Yaklaşık yarısı ise (%49) araştırma verilerinin uzun süreli olarak araştırmayı destekleyen kuruluşta depolanması gerektiği görüşüne sahiptir. Katılımcıların %39'u kurum dışında ücretsiz depolamaya olumlu bakarken, kurum dışı ücretli depolamanın tercih edilme oranı sadece %8'dir. Depolama yeri konusunda da bilim dalları arasında önemli bir farklılık görülmemektedir (Şekil 19).



Şekil 18. Veri setlerini depolama sorumluluğu ile ilgili görüşler



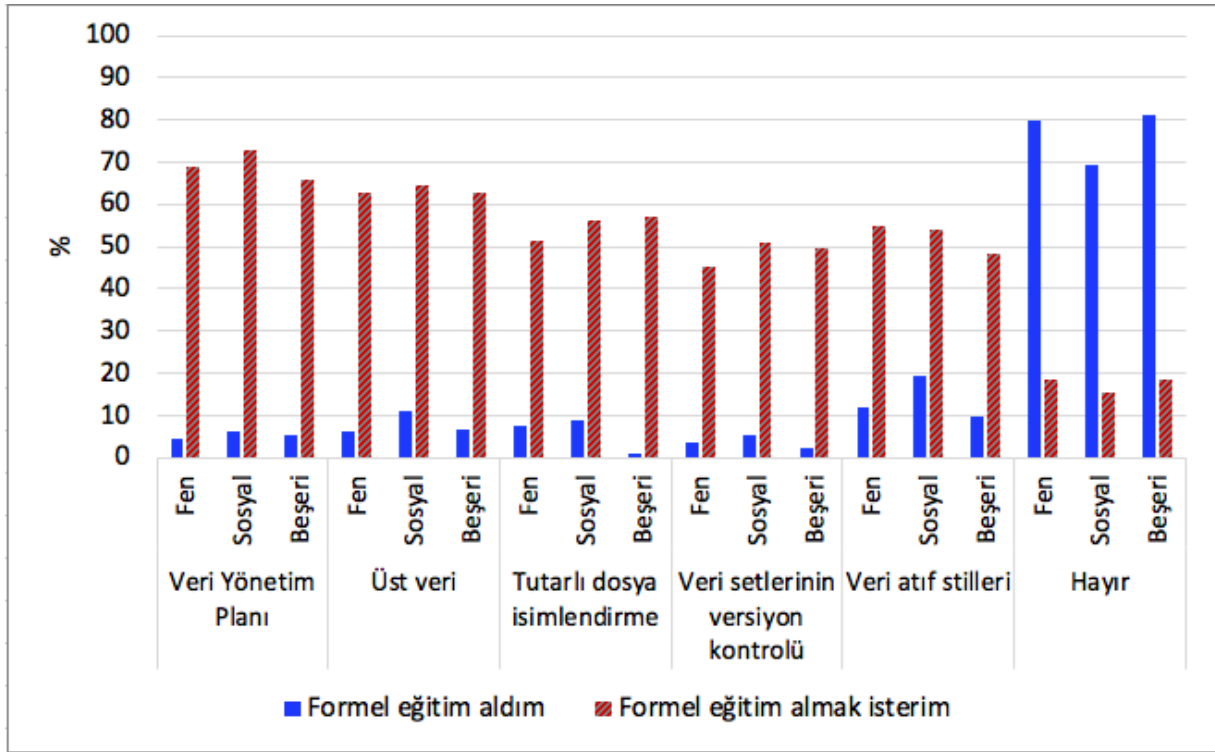
Şekil 19. Uzun süreli erişim için verileri depolama yeri ile ilgili görüşler

Araştırma Verilerinin Yönetimiyle İlgili Eğitim Gereksinimi

Katılımcıların araştırma verilerinin yönetimi ile ilgili konularda aldıkları eğitimler araştırılmıştır. Araştırmacıların büyük bir kısmı (%76) herhangi bir formel eğitim almadığını belirtmiştir. VYP (%5), üst veri (%8), tutarlı dosya isimlendirme (%7), veri setlerinin versiyon kontrolü (%4) konularında eğitim alanların oranları ise oldukça düşüktür ve bilim dallarına göre önemli bir farklılık görülmemektedir. En çok eğitim alınan konu %14 ile veri atıf stilleridir. Eğitim almayanların oranının fen (%80) ve beşerî bilimlerde (%81) sosyal bilimlerden (%69) daha yüksek olduğu görülmektedir.

Katılımcıların eğitim almaya en çok gereksinim duydukları konular %70 ile VYP, %63 ile üst veridir. Araştırmacıların %17'si araştırma verilerinin yönetimiyle ilgili konularda herhangi bir formel eğitim almak istemediğini belirtmiştir (Şekil 20). Averkamp, Gu ve Rogers (2014) tarafından Iowa Üniversitesi'nde yürütülen çalışmada fakülte üyelerinin araştırma verilerinin yönetimi konusundaki bilgi gereksinimlerinin en çok verilerin depolanması konusunda olduğu saptanmıştır.

Bulgular araştırmacılarımızın araştırma verilerinin yönetimi kapsamındaki pek çok konuda eğitime ihtiyacı olduğunu göstermektedir. Bu durum konunun henüz Türkiye'de nispeten yeni ele alınması ve dolayısıyla konuyla ilgili farkındalık ve bilgi eksikliği ile açıklanabilir. Söz konusu eksiklik ise büyük ölçüde ne üniversitelerin ne de fon sağlayıcı kurumların destekledikleri araştırmalar için VYP, verileri diğer araştırmacılara açma gibi taleplerinin olmamasından ve ilgili politika ve direktiflerin eksikliğinden kaynaklanmaktadır.



Şekil 20. Alınan ve alınmak istenen eğitim konuları

Sonuç ve Öneriler

Araştırma verilerinin erişime açılması ve paylaşımı verilerin farklı araştırmacılar tarafından farklı amaçlarla yeniden kullanımına olanak tanımakta, bir taraftan zaman, emek ve maliyetten tasarruf sağlarken diğer taraftan araştırma bulgularının doğrulanmasına fırsat yaratmaktadır. Araştırma verilerinin başkaları tarafından keşfedilebilir, erişilebilir ve yeniden kullanılabilir hale getirilmesi süreci ise hem yaratım, erişim, analiz ve yeniden kullanım aşamalarını gerçekleştiren/gerçekleştirecek araştırmacıların hem de veri yönetimi ve koruma gibi diğer aşamaları gerçekleştiren/gerçekleştirecek araştırma verisi yöneticilerinin bir takım beceriler geliştirmesini zorunlu kılmaktadır.

Araştırma verilerinin erişime açılması fikri yeni olmamakla birlikte, gerek ilgili politika ve direktiflerin eksikliği gerekse güven, itibar kaybı ve etik gibi kaygılar nedeniyle araştırmacılar tarafından benimsenmesi zaman almakta ve yavaş olmaktadır. Araştırmacıların verilerini nasıl yönettikleri ve verilerini paylaşma konusundaki tutumları konusunda henüz çok fazla şey bilinmemekle birlikte bu konuda yapılan araştırmaların sayısı giderek artmaktadır.

Ülkemizde araştırma verilerinin büyük bölümü erişilebilir ve yeniden kullanılabilir durumda değildir. Araştırma verilerinin yönetimi ile ilgili politika, strateji ve direktifler henüz oluşturulmamıştır. Türkiye'deki araştırmacıların araştırma verilerini nasıl yönettiği konusunda yapılmış çok az araştırma vardır. Bu çalışmanın bulgularının eldeki sınırlı bilgiye katkı sağlayacağı düşünülmektedir. Çalışmanın ana bulguları literatür kısmında ele alınan diğer çalışmaların bulguları ile büyük ölçüde benzerlik göstermektedir.

Bulgular araştırmacıların araştırmalarında kullandıkları ve araştırmaları sonucunda ürettikleri veri türü açısından disiplinler arasında benzerlikler kadar farklılıklar da olduğunu, az sayıda araştırmacının büyük veri ile çalıştığını ve büyük veri ile çalışanların fen bilimlerinde yoğunlaştığını göstermektedir. Katılımcıların yarısından fazlası kendi verisini üretmekte, bunun yanı sıra verilerini ağırlıklı olarak sahip oldukları araştırma ağından ve üniversite içindeki araştırma ekibinden temin etmektedir. Araştırmacıların büyük bölümü dış kaynaklı veri üzerinde belli miktarda bir çaba harcadıktan sonra bu veriyi kullanabilmekte, neredeyse tamamı

ürettikleri veriyi kendi cihazlarında saklamakta ve bulut ortamını depolama için ikinci sırada tercih etmektedir. Araştırmacıların yaklaşık dörtte biri depolarken verilerine hiçbir ek bilgi atamadıklarını, az sayıda da olsa bazı araştırmacılar veri paylaşımında işbirliği yapmadıklarını ve verilerini kimseye açmadıklarını belirtmektedir. Araştırmacıların veri paylaşımı konusuna genelde olumlu ancak temkinli yaklaştıkları görülmektedir. Veri paylaşımıyla ilgili endişelerin başında yasal ve etik kaygılar gelmektedir. Veri paylaşımı konusunda en az endişe duyan grup beşerî bilimlerdir. Katılımcıların yarıdan fazlası VYP'nin fonksiyonunu, üniversitelerinin SDİ sistemi olup olmadığını, üst veri terimini bilmemekte, veri etiketlemek için üst veri standardı kullanmamaktadır. Araştırmacıların yarıdan fazlası araştırma verilerine atıf yaptığını ve yarıya yakını verilere atıf yapmak için standart bir format kullandığını belirtmiştir. Araştırmacıların büyük bölümü verilerini proje süresinden sonra da saklamak istemekte ve üniversitelerin veri setlerinin depolanma ve erişime açılması, VYP hazırlamak, kurumsal arşive veri yüklemek için üst veri seti ve SDİ sistemi geliştirmek/önermek gibi veri yönetimi ile ilgili sorumlulukları olduğunu düşünmektedir. Çoğunluk veri setlerinin depolanması ya da erişime açılması konusundaki sorumluluğu sırasıyla üniversitelerinin ve araştırmayı destekleyen kuruluşun alması gerektiğini düşünmektedir. Uzun süreli depolama yeri olarak da araştırmacıların tercihleri önce üniversitelerden ikinci sırada araştırmayı destekleyen kuruluştan yanadır. Bulgular araştırmacıların araştırma verilerinin yönetimi konusunda kapsamlı bir eğitime ihtiyacı olduğunu göstermektedir. Araştırmacıların büyük bölümü eğitim almaya isteklidir ve eğitim almaya en çok gereksinim duyduklarını belirttikleri konular ise VYP ve üst veridir.

Eğitim ihtiyacının belirlenmesi kadar söz konusu eğitimi kimin vereceği de önemlidir. Bu konuda başta kütüphaneciler olmak üzere konu üzerinde çalışan uzmanlara önemli görevler düşmekte, çok gecikmeden yeni bir sorumluluk alanı olarak araştırma verilerinin yönetimi konusunun kütüphanecilerin eğitimine de yansıtılması gerekmektedir.

Araştırma verilerinin başarıyla yönetilmesinin ve sürdürülebilirliğinin temel şartları gerekli politika, direktif ve düzenlemelerin hem ulusal hem kurumsal düzeyde geliştirilmesi, teknik alt yapının sağlanması, destek birimlerinin oluşturulması ve destek hizmetlerinin verilmesidir. Bunun yanı sıra araştırma verilerinin yönetimi yaşam döngüsü içinde üzerlerine düşeni yapabilmeleri için hem araştırmacıların hem de araştırma verisini yönetecek iş gücünün eğitilmesi gerekmektedir. Gerek araştırmacıların gerekse ilgili iş gücünün eğitilmesinde disiplinlere yönelik farklılıkların göz önüne alınması önemlidir. Söz konusu farklılıkların daha ayrıntılı belirlenebilmesi için daha spesifik alanlarda yapılacak çalışmalara gereksinim vardır.

Araştırma verilerinin yönetimi sorumluluğunu üstlenmek için kütüphaneciler en uygun meslek grubu olarak görünmektedir. Ancak araştırma verisi yönetimi ile ilgili konuların çok gecikmeden bilgi ve belge yönetimi bölümleri programlarına entegre edilmesi gerekmektedir. Bunun yanı sıra mesleğe atılmış kütüphanecilerin eğitimi de meslek içi eğitim programları çerçevesinde ele alınmalı, kütüphaneler veri yönetimi birimleri oluşturmak için gerekli alt yapıyı hazırlamak üzere girişimlere başlamalıdır. Tüm bu girişimler ancak politika ve direktiflerin desteği ile gerçekleştirilebileceğinden öncelikle bu alanda gelişmeye gereksinim vardır. TÜBİTAK açık erişim politikası taslağında açık verinin de ele alınmış olması iyi bir başlangıç olarak düşünülebilir. Ancak araştırma verilerinin başarıyla yönetimi için daha detaylı politika belgelerine gereksinim vardır. Söz konusu koşullar sağlandıktan sonra veri paylaşımı konusunda özendirme ve ödüllendirme sistemlerinin geliştirilmesi de faydalı olacaktır.

Kaynakça

Akers, K. G. ve Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, 8(2), 5-26. doi:10.2218/ijdc.v8i2.263

- Aydinoglu, A. U. (2014, Kasım). *The social, technical, and policy landscape of in environmental sciences in Turkey*. 5. Uluslararası Değişen Dünyada Bilgi Yönetimi Sempozyumu'nda sunulan bildiri, Antalya, Türkiye. Erişim adresi: <http://imcw2014.bilgiyonetimi.net/wp-content/uploads/sites/12/2015/01/Aydinoglu.pdf>
- Aydınoglu, A. U. (2016, Ekim). *Araştırma verileri yönetimi: Türkiye*. 5. Ulusal Açık Erişim Konferansı'nda sunulan bildiri, Ankara, Türkiye. Erişim adresi: http://ae2016.acikerisim.org/wp-content/uploads/sites/2/2016/11/arsev_aydinoglu.pdf
- Aydinoglu, A. U., Dogan, G. ve Taskin, Z. (2017). Research data management in Turkey: Perceptions and practices. *Library Hi-Tech*, 35, 271-289. doi:10.1108/LHT-11-2016-0134
- Ball, A. ve Duke, M. (2015). *How to track the impact of research data with metrics*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Erişim adresi: <http://www.dcc.ac.uk/resources/how-guides>
- Beagrie, N. ve Houghton, J. (2013). *The value and impact of data sharing and curation: A synthesis of three recent studies of UK research data centres*. Jisc. Erişim adresi: http://repository.jisc.ac.uk/5568/1/iDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63, 1059-1078. doi:10.1002/asi.22634
- Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world*. Cambridge, MA, The MIT Press.
- Borgman, C. L., Wallis, J. C. ve Enyedy, N. (2007). Little science confronts the data deluge: Habitat ecology, embedded sensor networks, and digital libraries. *International Journal of Digital Libraries*, 7, 17-30. doi:10.1007/s00799-007-0022-9
- Borgman, C. L., Wallis, J. C. ve Mayernik, M. S. (2012). Who's got the data? Interdependencies in science and technology collaborations. *Computer Supported Cooperative Work*, 21, 485-523. doi:10.1007/s10606-012-9169-z
- Calzada Prado, J. ve Ángel Marzal, M. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63, 123-134. doi:10.1515/libri-2013-0010
- Carlson, J., Fosmire, M., Miller, C. C. ve Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *Libraries and the Academy*, 11, 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, 346-353. doi:10.1016/j.acalib.2017.06.006
- Chowdhury, G., Boustany, J., Kurbanoglu, S., Ünal, Y. ve Walton, G. (2017). Preparedness of research data sharing: A study of university reserachers in three European countries. S. Choemprayong, F. Crestani ve S. J. Cunningham (Ed.). *Digital Libraries: Data, Information and Knowledge for Digital Lives: 19th International Conference on Asia Pacific Digital Libraries, ICADL 2017, Bangkok, Thailand, Nowember 13-15, 2017: Proceedings* içinde (ss. 104-116). Cham, Switzerland: Springer.
- Chowdhury, G., Ünal, Y., Kurbanoglu, S., Boustany, J. ve Walton, G. (2018). *Research data management and data sharing behaviour of university researchers*. ISIC: The Information Behaviour Conference, 9-11 Ekim 2018, Kraków, Polonya. (Tam metni kabul edilmiş bildiri).
- Concordat on open research data*. (2015). Versiyon 10, 17 Temmuz 2015. Erişim adresi: <http://www.oulu.fi/yliopisto/sites/default/files/content/ConcordatOpenResearchData.pdf>
- EU Open Data Portal: Access to European Union open data*. (2018). Erişim adresi: <https://data.europa.eu/>

- Fane, B., Treadway, J., Gallagher, A., Penny, D. ve Hahnel, M. (2016). Open season for open data: A survey of researchers. *Digital Science Report. The state of open data: A selection of analyses and articles about open data*, curated by Figshare içinde (ss. 12-19). Erişim adresi: https://figshare.com/articles/The_State_of_Open_Data_Report/4036398
- Faniel, I. M., Kriesberg, A. ve Yakel, E. (2012). Data reuse and sensemaking among novice social scientists. *Proceedings of the American Society for Information Science and Technology*, 49, 1-10. doi:10.1002/meet.14504901068
- Faniel, I., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J. ve Yakel, E. (2013). The Challenges of digging data: A study of context in archaeological data reuse. *JCDL 2013 Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* içinde (ss. 295-304). New York, NY: ACM. doi:10.1145/2467696.2467712. Ön baskı erişim adresi: <https://www.oclc.org/content/dam/research/publications/library/2013/faniel-archae-data.pdf>
- Federer, L. M., Lu, Y. L. ve Joubert, D. J. (2016). Data literacy training needs of biomedical researchers. *Journal of the Medical Library Association*, 104, 52-57. doi:10.3163/1536-5050.104.1.008
- Frank, E. P. ve Pharo, N. (2016). Academic librarians in data information literacy instruction: A case study in meteorology. *College & Research Libraries*, 77, 536-52. doi:10.5860/crl.77.4.536
- Gürdal, G. ve Bitri, E. (2015). *Araştırma verisi yönetimi, açık veri ve Avrupa Birliği bilimsel veri altyapısı: OpenAIRE2020*. XVII. Akademik Bilişim Konferansı. Erişim adresi: <http://ab.org.tr/ab15/ozet/124.html>
- Kenan, M. A. ve Markauskaite, L. (2015). Research data management practices: A snapshot in time. *International Journal of Digital Curation*, 10(2), 69-95. doi:10.2218/ijdc.v10i2.329
- Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71, 401-415. doi:10.1108/JD-02-2014-0026
- Lewis, M. J. (2010). Libraries and the management of research data. McKnight, S. (Ed.) *Envisioning Future Academic Library Services. Facet Publishing, London* içinde (ss. 145-168). Erişim adresi: http://eprints.whiterose.ac.uk/11171/1/LEWIS_Chapter_v10.pdf
- MacMillan, D. (2014). Data sharing and discovery: What librarians need to know. *The Journal of Academic Librarianship*, 40, 541-549. doi:10.1016/j.acalib.2014.06.011
- Mayernik, M. S., Wallis, J.C. ve Borgman, C. L. (2013). Unearthing the infrastructure: Humans and sensors in field-based scientific research. *Computer Supported Cooperative Work*, 22, 65-101. doi:10.1007/s10606-012-9178-y
- OpenAIRE. (2018). What is the Open Research Data Pilot? Erişim adresi: <https://www.openaire.eu/what-is-the-open-research-data-pilot>
- Strasser, C. (2015). *Research data management: A primer publication of the National Information Standards organization*. Baltimore, MD: NISO. Erişim adresi: https://groups.niso.org/apps/group_public/download.php/15375/PrimerRDM-2015-0727.pdf
- 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
- The Data harvest: How sharing research data can yield knowledge, jobs and growth*. (2014). An RDA Europe Report, Aralık 2014. Erişim adresi: <https://rd-alliance.org/sites/default/files/attachment/The%20Data%20Harvest%20Final.pdf>
- The UK Data Archive*. (2018). Create & manage data. Erişim adresi: <http://www.data-archive.ac.uk/create-manage>

- Tonta, Y. (2013). *Açık erişimin geleceği ve araştırma verilerine açık erişim*. Bilkent Üniversitesinde Kütüphanecilik Seminerleri'nde sunulan bildiri, 17 Aralık 2013, Ankara. Erişim adresi: <https://library.bilkent.edu.tr/tr/acik-erisim-farkli-bakis-acilari-vi-seminer/>
- University of Oxford Research Data Management Survey 2012: The Results*. (2013). Erişim adresi: <http://blogs.it.ox.ac.uk/damaro/2013/01/03/university-of-oxford-research-data-management-survey-2012-the-results/>
- Verbaan, E. ve Cox, A. M. (2014). Occupational sub-cultures, jurisdictional struggle and third space: Theorising professional service responses to research data management. *The Journal of Academic Librarianship*, 40, 211-219. doi: 10.1016/j.acalib.2014.02.008
- Yakel, E. ve Faniel, I. (2014). *Virtuous circles: Circulating old data through new collaborations*. 17th ACM Conference on Computer Supported Cooperative Work and Social Computing Workshop: Sharing, Re-Use and Circulation of Resources in Cooperative Scientific Work. Baltimore, MD, 15 Şubat 2014.
- Wallis, J. C., Rolando, E. ve Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE* 8(7), e67332. doi: 10.1371/journal.pone.0067332

Summary

Making research data open and available for easy access and sharing allows reuse of data by different researchers for different purposes and makes research findings verifiable while eliminating expenditure of time, effort and funds. It is evident that in order to make research data accessible, discoverable and reusable by others, development is required in several areas such as building the appropriate ICT infrastructure; building capacity for managing research data (workforce for managing and preserving data); developing appropriate regulations and policies governing sustainable data management practices; and training researchers who take part in the first phases of research data lifecycle such as creation, collection and analysis of data.

Whilst the idea of free and open data for reuse is not new, the adopting and take-up amongst researchers has been low because of issues related to trust (such as issues surrounding intellectual property concerns), reputation, and ethics as well as lack of policies and mandates. Today, little is known about how researchers go about managing their data and whether they are willing to share it with others. However, body of research regarding researchers' data management practices is growing.

Government bodies and funding agencies around the world have recently advocated and introduced requirements for open access and sharing of research data. In Turkey, due to lack of necessary policies, strategies and regulations for research data management (RDM), the majority of research data is not archived and cannot be accessed and re-used. Neither the national funding agency for scientific research (TUBITAK= The Scientific and Technological Research Council of Turkey) nor other funding agencies or universities have any RDM policy and/or mandate, and do not require a data management plan from research they fund. As a result, none of the research institutions have yet implemented services for research data storage, analysis and curation. There are no units within research institutions which provides support to researchers who would like to store and share their research data. However, despite the lack of policy development and service implementation for RDM, starting from 2012, there have been several initiatives which aim to increase awareness towards the importance of the subject and address the current situation in Turkey.

There are only few researches carried on Turkish researchers' RDM practices. This study is an attempt to gain an understanding of the current practices of research data use, creation and sharing amongst academic researchers, and their view on the underlying issues and challenges of RDM. With this aim, an online questionnaire survey was conducted amongst university researchers (540 researchers from 44 universities) in Turkey. There were 26 questions to collect data on: researcher information for example role, discipline, gender, experience; nature and volume of data collected and created; data sharing practices, concerns; familiarity with data management practices; knowledge of data management practices such as tagging, use of metadata; and needs for training. SPSS was used to code and analyse the dataset, and Chi-Square tests were conducted to find out correlation between researchers' behaviour in different areas of RDM and their discipline. The following classification of disciplines was used in this study: Sciences (natural sciences, engineering and technology, medical and health sciences, agricultural sciences), social sciences (for example education, sociology, law and political science), humanities (for instance history, archaeology, language and literature).

The key findings from the research are that there are differences as well as similarities between the data type used and created by researchers from different disciplines, limited number of researchers work with big data and they generally come from sciences, research data is obtained from multiple sources, researchers need to spend some time to pre-process data received from outside sources, research data is generally stored on researchers' own devices, non or very limited number of tags are assigned to research data while storing, researchers feel that research data should be stored at their own universities for long-term storage and made available, they generally have positive views towards data sharing however they have some concerns such as data ethics, they like to keep their data after the closure of the research project, they need comprehensive training on RDM and are interested in receiving formal training on RDM related issues.