

ARAŞTIRMA VERİLERİNİN YÖNETİMİ ANKET RAPORU



Araştırma Verilerinin Yönetimi Anket Raporu

Elektronik ortamda LimeSurvey yazılımı aracılığıyla hazırlanan anket ARBİS'e kayıtlı 37.223 akademisyene gönderilmiştir. Ankete 1577 kişi dönüş yapmıştır. Anketin en temel üç sorusu katılımcılara araştırma verisi kullanıp kullanmadıklarının, daha önce hiç araştırma verisine atıf yapıp yapmadıklarının ve araştırma verisi üretip üretmediklerinin sorulduğu üç sorudur. Bu üç sorunun yanıtlanmadığı 269 anket analize dâhil edilmemiş, analizler geriye kalan 1308 anket üzerinden gerçekleştirilmiştir.

Formül 1 ve Formül 2'ye göre (Cochran, 1963, s. 75) 37.223 kişilik evren %99 güven düzeyinde $e=0,04$ ile 996 kişilik ve $e=0,03$ ile 1736 kişilik bir örneklem ile temsil edilebilir ($z=2,56$; $p=0,5$; $q=0,5$).

$$n_0 = \frac{z^2 pq}{e^2} \quad (\text{Formül 1})$$

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \quad (\text{Formül 2})$$

Formüllerde kullanılan simgelerin anlamları aşağıdaki gibidir:

N: Evren büyüklüğü

*n*₀: Örneklem büyüklüğü

n: Düzeltilmiş örneklem büyüklüğü

z: Seçilen güven düzeyine karşılık gelen *z* tablo değeri

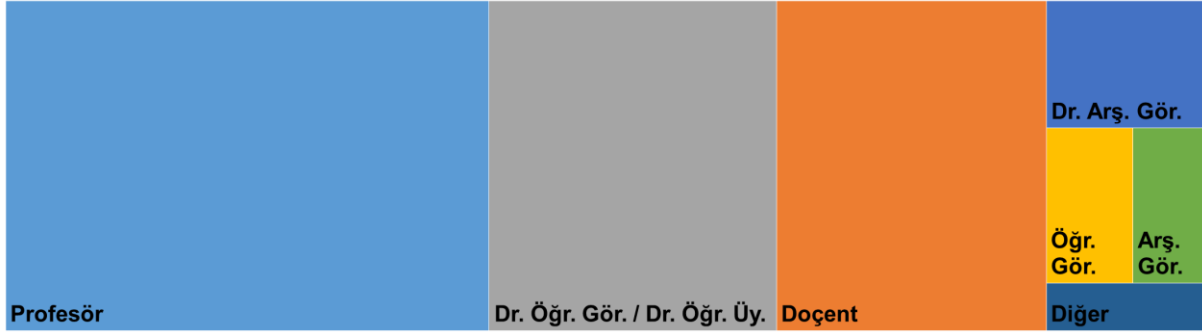
p: İncelenen olayın görülme oranı

q: İncelenen olayın görülmemesi oranı ($q = 1 - p$)

e: Örnekleme hatası

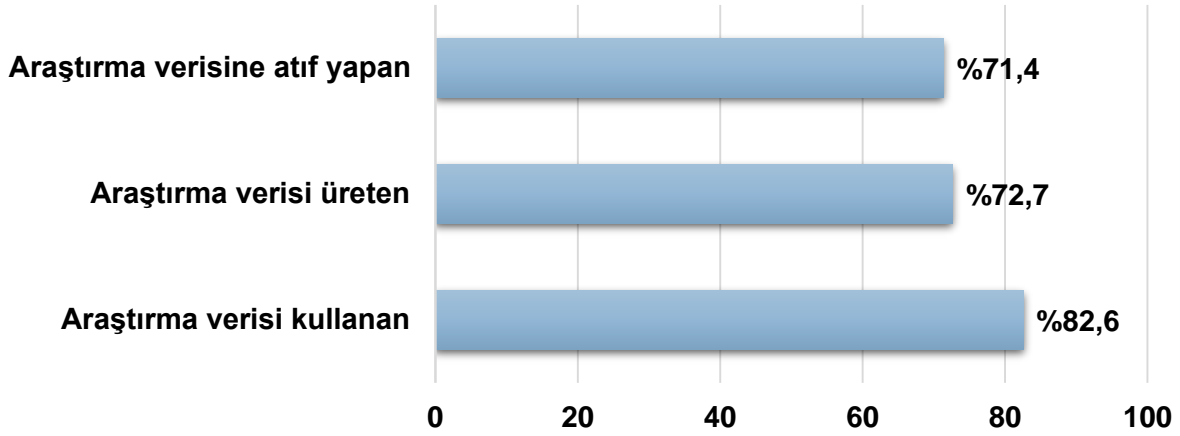
1308 kişinin unvanlarına göre dağılımları Şekil 1'de gösterilmektedir. Yaklaşık %40'lık bir oranla çoğunluğu profesörler oluşturmaktadır. Birbirine yakın oranlarla doçentler (%22,5) ve Doktor Öğretim Görevlileri ile Doktor Öğretim Üyeleri grubu (%24) takip etmektedir. Tümünü birlikte anketi yanıtlayanların yaklaşık %87'sine karşılık gelmektedir. Bunların dışında kalan %13'lük kısımda

Öğretim Görevlileri (%3,4), Doktor Araştırma Görevlileri (%5,1), lisansüstü eğitime devam eden Araştırma Görevlileri (%2,8) yer almaktadır.



Şekil 1. Anketi yanıtlayanların unvanlarına göre dağılımı (%)

Bu kişiler arasında araştırma verisi kullanım oranı yaklaşık %83'tür. Beklenmedik bir biçimde, yaklaşık %73 oranında daha önce araştırma verisine atıf yapıldığı belirtilmiş, "veriye atıf yapıldığını bilmiyordum" seçeneği kimse tarafından işaretlenmemiştir. Katılımcıların içerisinde yaklaşık %71'lik bir grup araştırma verisini kendisi üretmektedir (Şekil 2).



Şekil 2. Araştırma verisi üreten, kullanan ve araştırma verisine atıf yapan kişilerin yüzdeleri

En çok üretilen/kullanılan veri türü her iki kişiden biri tarafından belirtilen deneysel veridir (%50). Bunu benzer oranlarla metin verisi (%24), anket (%22,5) ve grafik (%22,2) takip etmektedir. Bu üç veri türü ise her dört kişiden biri tarafından üretilmekte ya da kullanılmaktadır (Tablo 1). Diğer seçeneğini işaretleyen yaklaşık %4'lük kısım içerisinde en dikkat çekici seçenekler beşer kişi tarafından belirtilen arazi verileri ile gözlem verisi olmuştur.

Tablo 1. Kullanılan/üretilen veri türlerinin yüzdeleri

Veri türü	%
DeneySEL veri	50,1
Metin verisi	23,9
Anket	22,5
Grafik	22,2
İşlenmemiş ham veri	13,1
Lab defteri	10,9
Veri modelleri	8,9
Ses kaydı	5,1
Video	4,3
Uzaktan algılama	3,3

Not: Bu soru için birden çok seçenek işaretlenebilmiştir.

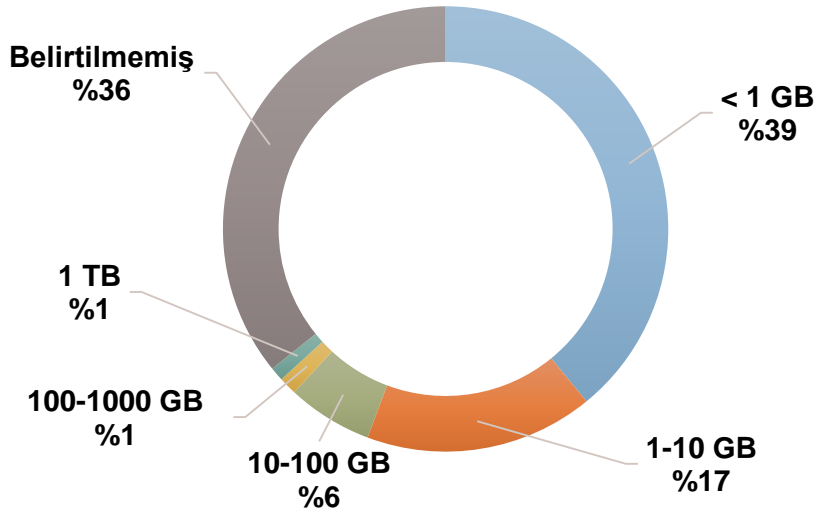
Öte yandan, dikkat çeken bir bulgu Tablo 1’de belirtilen 10 veri türünü de kullanmadığını belirten ve diğer seçeneğine de kullandığı veri türünün ne olduğunu belirtmeyen yaklaşık %36 oranındaki bir grubun varlığıdır. Benzer şekilde, katılımcıların yine yaklaşık %36’sı kullandığı/ürettiği veri formatı için Tablo 2’de belirtilen seçeneklerden herhangi birini işaretlememiş ve diğer seçeneğinde bunların dışında bir veri formatı belirtmemiştir.

Tablo 2. Kullanılan/üretilen veri formatlarının yüzdeleri

Veri türü	%
xls	49,5
txt	34,8
free text	16,4
sav	14,3
csv	13,5
spreadsheet	9,6
readme structured	2,1
readme unstructured	1,0

Not: Bu soru için birden çok seçenek işaretlenebilmiştir.

Üretilen/kullanılan veri formatlarına ilişkin yüzdelerin gösterildiği Tablo 2'ye göre xls (%49,5) ve txt (%34,8) öne çıkmaktadır. Katılımcıların %6'sı tarafından işaretlenen diğer seçeneğinde belirtilenlere bakıldığında en dikkat çeken 16 kişi tarafından (yaklaşık %1) bir veri formatı olarak düşünülen word/doc olmuştur. Veri formatı olarak seçenekler arasında istatistiksel bir yazılım olan SPSS programının veri formatı sav olmasına rağmen diğer seçeneğinde SPSS'i bir veri formatı olarak belirten iki kişi bulunmaktadır. Ayrıca, pdf'in de bir veri formatı olarak belirtildiği görülmektedir. Diğer seçeneğinde dikkat çeken bir diğer nokta resim/görüntü ve video formatları olmuştur (jpeg, tiff, png, mp4 gibi)



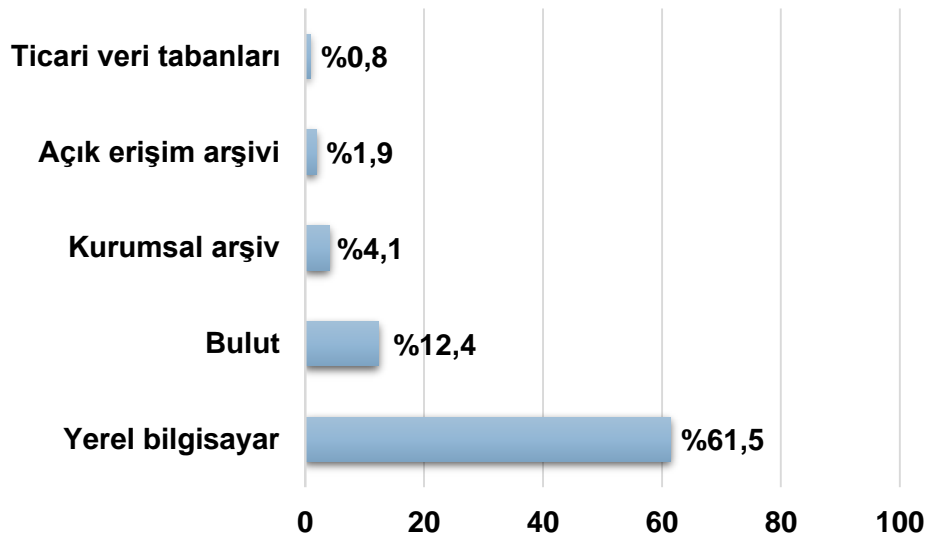
Şekil 3. Katılımcıların en son çalışmalarında kullandıkları/ürettikleri verinin ortalama büyüklüğüne göre dağılımı

Katılımcıların en son çalışmalarında kullandığı/ürettiği verinin ortalama büyüklüğü ile ilgili sorunun seçeneklerinden birini işaretlemeyen %36'lık kısmın konu ile ilgili fikir sahibi olmadıkları düşünülmektedir. 10 GB'ın üzerinde bir büyüklükte veri ürettiğini/kullandığını belirten kişilerin oranı %7 civarındadır. Genel olarak kullanılan/üretilen verilerin "büyük veri" niteliğinde olmadığı söylenebilir. Katılımcıların %39'u için kullanılan/üretilen verinin büyüklüğü 1 GB'ın altındadır.

Katılımcıların ürettikleri/kullandıkları verileri nerelerde sakladıklarına ilişkin bulgular Şekil 4 ile sunulmaktadır. Verilerin saklanması amacıyla en çok tercih edilen ortam yerel bilgisayarlardır

(%61,5). Seçenekler arasında “bulut” olmasına rağmen, diğer seçeneğini işaretleyen yaklaşık %5’lik grubun bu seçenek altında belirttiklerinin arasında drive ve dropbox olması ilginçtir. Diğer seçeneğini işaretleyenlerin ço büyük kısmı tarafından belirtilen veri saklama ortamı harici diskidir.

Kurumsal arşivler, açık erişim arşivleri gibi dijital arşivlerin veri depolama/saklama amacıyla kullanımının ülkemizde henüz çok yaygınlaşmadığını Şekil 4’ten de görmek mümkündür. Data Citation Index gibi ticari veri tabanlarını veri depolama amacıyla kullandığını belirten kişi sayısı 11 ile sınırlı kalmıştır. Katılımcılar yaklaşık %71 oranında araştırma verilerine atıf yaptıklarını belirtmelerine rağmen, kendi verilerini bu verilere atıf yapılabilmesini sağlayacak kurumsal arşiv, açık erişim arşivi ya da ticari veri tabanı gibi ortamlarda depolama konusunda çekingen davrandıkları düşünülmektedir.

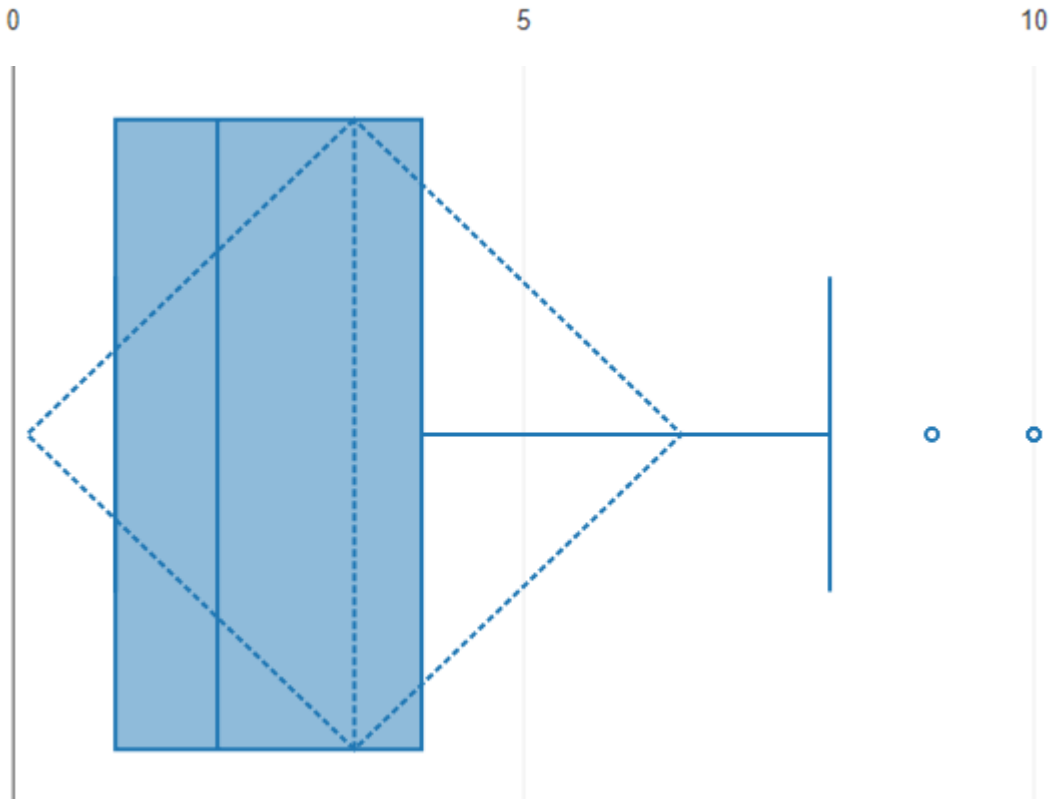


Şekil 4. Üretilen/kullanılan verilerin saklanma yerleri (Not: Bu soru için birden çok seçenek işaretlenebilmiştir)

Katılımcılara daha önce proje yürütüp yürütmedikleri sorulmuş ve bu soruya yanıt veren 1205 kişinin yaklaşık %76’sı daha önce proje yürüttüklerini belirtmiştir. Daha önce TÜBİTAK projesi yürütülüp yürütülmediği sorusuna cevap veren 913 kişinin ise yaklaşık yarısı (454 kişi) TÜBİTAK projesi yürüttükleri yönünde cevap vermiştir.

Herhangi bir rolde görev alınan toplam TÜBİTAK projesi sayısı ile ilgili olarak 423 kişi bilgi vermiştir (katılımcıların yaklaşık %32’si, daha TÜBİTAK projesi yürüttüğünü belirtenlerin %93’ü).

Söz konusu 423 kişinin yer aldığı toplam TÜBİTAK projesi sayısı 1413 iken, bireysel olarak görev alınan proje sayısı 1-24 arasında değişmektedir. Şekil 5'teki kutu grafiği 10'dan fazla (11-24) projede görev almış 14 kişi ihmal edilerek oluşturulmuştur. Şekilde kesikli çizgiler ortalama ve standart sapmayı temsil etmektedir. Görev alınan ortanca proje sayısı 2'dir. Diğer bir ifadeyle, bir TÜBİTAK projesinde çalışmış olanların yaklaşık yarısı için çalışılan proje sayısı 1 ya da 2'dir. Görev alınan proje sayısı 1-4 arasında yoğunlaşmaktadır. %31'lik bir kısım için görev alınan proje sayısı bir proje ile sınırlı kalmış, belirtilen 423 kişinin yalnızca %30'u üçten daha fazla projede yer almıştır.



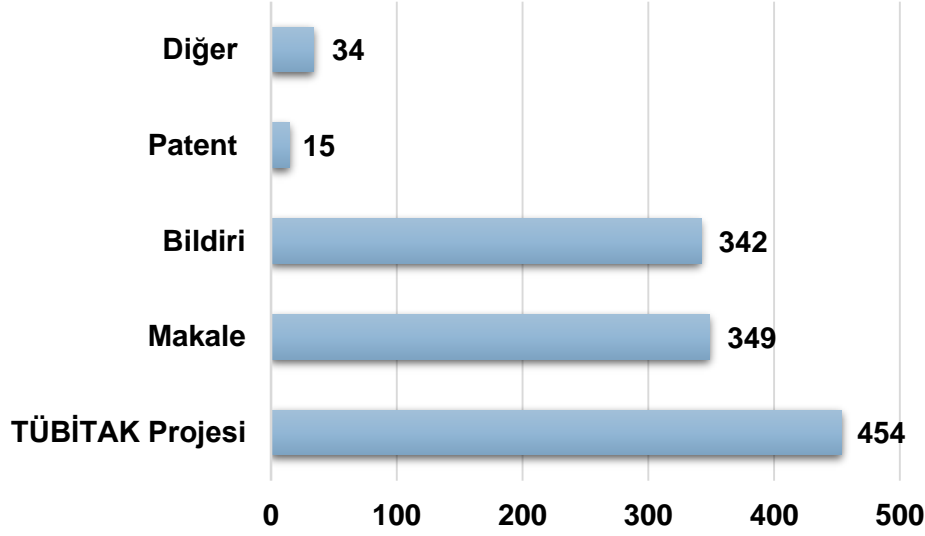
Şekil 5. Görev alınan toplam TÜBİTAK proje sayılarına ilişkin kutu grafiği

Görev aldığı TÜBİTAK projelerinin türünü belirten 422 kişiden alınan bilgiler incelendiğinde, en yaygın proje türünün 1001 kodlu *Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programı* kapsamındaki projeler olduğu görülmektedir. Katılımcıların %69'u en az bir 1001 projesinde görev aldığını yönünde beyanda bulunmuştur. 1001 projelerini %23 oranıyla 1002 kodlu *Hızlı Destek Programı* kapsamındaki projeler takip etmektedir. *Başlangıç Ar-Ge Projeleri*

Destekleme Programı (3001) projelerinin %13 oranında, *Öncelikli Alanlar Ar-Ge Projeleri Destekleme Programı* (1003) projelerinin %11 oranında ve *Kariyer Geliştirme Programı* (3501) projelerinin %9 oranında yapıldığı görülmektedir. Az sayıda kişi tarafından yapılan projeler aşağıda sıralanmıştır.

- 1005 – Ulusal Yeni Fikirler ve Ürünler Araştırma Destek Programı (7 kişi)
- 1007 – Kamu Kurumları Ar-Ge Projelerini Destekleme Programı (10 kişi)
- 1010 – Evrensel Araştırmacı Programı (EVRENA) (3 kişi)
- 1505 – Üniversite-Sanayi İşbirliği Destek Programı (7 kişi)
- 1507 – KOBİ Ar-Ge Başlangıç Destek Programı (7 kişi)
- 2211 – Yurt İçi Lisansüstü Burs Programı (1 kişi)
- 2237 – Proje Eğitimi Etkinliklerini Destekleme Programı (1 kişi)
- 4005 – Bilim ve Toplum Yenilikçi Eğitim Uygulamaları (4 kişi)
- 1301 – Bilimsel ve Teknolojik İşbirliği Ağları ve Platformları Kurma Girişimi Projeleri (İŞBAP) Destekleme Programı (1 kişi)
- 2209 – Üniversite Öğrencileri Yurt İçi / Yurt Dışı Araştırma Projeleri Destekleme Programı (7 kişi)
- 4004 – Doğa Eğitimi ve Bilim Okulları Projesi (10 kişi)
- 1501 – Sanayi Ar-Ge Projeleri Destekleme Programı (2 kişi)
- 2232 – Yurda Dönüş Araştırma Burs Programı (1 kişi)
- 2219 – Yurt Dışı doktora Sonrası Araştırma Burs Programı (10 kişi)
- 1503 – Proje Pazarları Destekleme Programı (1 kişi)
- 4007 – Bilim Şenlikleri Destekleme Programı (1 kişi)
- 1511 – Öncelikli Alanlar Ar-Ge Destek Programı (2 kişi)
- 1512 – Teknogirişim Sermayesi Desteği Programı (BiGG) (2 kişi)
- 1601 – Yenilik Girişimcilik alanlarında Kapasite Arttırılmasına Yönelik Destek Programı (2 kişi)
- 2217 – Lisansüstü Yaz Okulu ve Benzeri Diğer Etkinlikleri Destekleme Programı (1 kişi)
- 2218 – Yurt İçi Doktora Sonrası Araştırma Burs Programı (1 kişi)
- 2221 – Konuk veya Akademik İzinli (Sabbatical) Bilim İnsanı Destekleme Programı (3 kişi)
- 2229 – Bilimsel Eğitim Etkinliklerini Destekleme Programı (1 kişi)

Bunlar dışında, SAN-TEZ ve/veya TEYDEP projesi olarak belirten 8 kişi, ikili / çoklu işbirliği projesi olarak belirten 14 kişi, herhangi bir proje sınıfına dahil edilemeyip diğer olarak sınıflandırılan projelerde çalışmış 9 kişi bulunmaktadır.



Şekil 6. TÜBİTAK projesi yürüten ve bu projelerden makale, bildiri, patent vb. yayın üreten kişi sayıları

Şekil 6'dan da görülebileceği gibi daha önce TÜBİTAK destekli proje yürüttüğünü belirten 454 kişinin yaklaşık %77'si (349 kişi) bu projelere dayalı makale ürettiğini belirtmiştir. Üretilen makale sayısı ortancası 4, 1-3 makale ürettiğini belirtenlerin oranı %47'dir. TÜBİTAK projelerine dayalı bildiri türü yayın yapmak ta oldukça yaygındır. TÜBİTAK projesi yürüten kişilerin %75'i (342 kişi) projelerden bildiri üretmişlerdir. Üretilen bildiri sayısı ortancası 5, 1-3 arası bildiri üreten kişilerin oranı ise %42'dir. 15 kişi yaptığı TÜBİTAK projesinden patent aldığını belirtmiş, bunlardan üçü üçer patent almışlardır. Bunlara ek olarak diğer seçeneği altında TÜBİTAK projelerinden yüksek lisans ve doktora tezleri, belgesel, kitap, kitap bölümü türü yayınların çıkarıldığı, telif hakkı alınıp, faydalı model oluşturulduğu da anlaşılmaktadır.

Daha önce yürüttüğü projelerde veri yönetim planı hazırlanmasının fon sağlayıcılar tarafından zorunlu tutulduğunu belirten 121 kişi olmuştur (katılımcıların yaklaşık %9'u, ilgili soruyu cevaplayanların yaklaşık %14'ü). Bu soruyu cevaplamayan 427 kişinin (katılımcıların yaklaşık %33'ü) veri yönetim planı ile ilgili bir fikirlerinin olmadığı düşünülmektedir. Anketi yanıtlayanlardan yaklaşık %91'inin veri yönetim planının ne olduğunu bilmeyen ve/veya daha

önce hiç veri yönetim planı hazırlamamış kişiler olduğunu söylemek yanlış olmayacaktır. Katılımcıların yaklaşık yarısı (%49) Horizon 2020 projelerinde zorunlu olan açık veri ve veri yönetim planı uygulamalarının TÜBİTAK projeleri için de geçerli olması gerektiğini düşünmektedir. Olumsuz fikir belirten kişilerin %7 oranında kalmış olması sevindiricidir. Bu konuda fikir belirtmeyen %31'lik grubun konu ile ilgili bilgi sahibi olmadıkları düşünülmektedir. Görüş belirten 101 kişinin görüşlerini belli gruplar altında toplamak mümkündür. Bir grup araştırmacı konu ile ilgili hiç bilgi sahibi olmadığını söylerken, bir kısmı ise verilerin paylaşımı konusunda kaygılıdır. Araştırmayı yapan kişinin haklarının gözetilmesi, verinin izinsiz kullanımı, güvenlik ve etik sorunlar en çok kaygı duyulan konulardır. Öte yandan, araştırma verilerine açık erişimin daha çok kişinin bu veriden faydalanmasını sağlamayacağı, zaman kaybını ve kaynak israfını önleyeceği, bilimsel etkiyi artıracığı yönünde fikir bildirenler de olmuştur. Açık veri ve veri yönetimi konusundaki eğitim ihtiyacı da belirtilen konu ile ilgili belirtilen görüşler arasındadır. Ayrıca, yurtdışında verilere ulaşmanın kolay olduğu ancak ülkemizde devlet verilerine erişimin sıkıntılı olduğu belirtilmiştir.

Katılımcıların %59'u (bu soruya cevap verenlerin %76'sı) kamu kaynakları ile üretilen verilerin açık veri olarak halkın kullanımına açılması gerektiğini düşünmektedir. Bu konuda olumsuz görüş bildiren kişi sayısı oldukça azdır (katılımcıların %8,2'si, bu soruyu yanıtlayanların %10,6'sı). Konu ile ilgili olarak 121 kişi tarafından belirtilen görüşler incelendiğinde en öne çıkan, açık erişim felsefesinin de temelini oluşturan "kamu kaynakları ile üretilen tüm bilimsel çıktıların kamuya ait olduğudur". Veri paylaşımının kaynak israfının önüne geçeceği, daha kapsamlı araştırmaların yapılmasını ve yaygın etki sağlayacağını belirtenler olmuştur. Konu ile ilgili çekinceler bu verilerin kamu yararı yerine kişisel çıkar sağlamak amacıyla kullanılması, yüksek gizlilik derecesine sahip veriler, verilerin anonimleştirilmesi, verilerin denetim ve kontrolünün teminat altına alınması, verinin kullanımı ile ilgili şartların belirlenmesi, araştırmayı yapan kişinin haklarının korunması konularındadır.

Ankete katılanların %67'si, ilgili soruya cevap verenlerin ise %90,5'i TÜBİTAK tarafından veri yönetimi konusunda yapılacak eğitimlere katılma; oluşturulacak portal ve açık erişimli kaynaklardan yararlanma konusunda istekli olduğunu belirtmiştir. Konu ile ilgili olumlu ya da olumsuz görül bildirmeyen 371 kişilik bir araştırmacı grubunun varlığı dikkat çekmektedir (katılımcıların yaklaşık %28'i).

Arařtırmacılar, TÜBİTAK projeleri esnasında üretilen arařtırma verilerinin fonlayıcı kurum olan TÜBİTAK tarafından kontrol edilmesi konusundaki önerileri sorulmuş ve %40 oranında görüş bildirilmiştir. Belirtilen görüşler önemli ölçüde olumlu olmakla birlikte, yukarıda belirtilen görüşlere ek olarak verinin kontrol ve denetiminden sorumlu olması gereken kişinin veriyi üreten arařtırmacı olduđu düşüncesi yaygındır. TÜBİTAK'ın arařtırma verilerinin kontrolünü uzman kişiler denetiminde yapması gerektiđi belirtilmiş, verinin güvenliğinin sağlanması ve etik ihlaller konusunda çok sayıda arařtırmacının çekincesi olduđu anlaşılmıştır.