

Görüşler / *Opinion Papers*

Araştırma Verilerinin Yönetimi

Research Data Management

Yaşar Tonta* ve Umut Al**

Öz

Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü ile Goethe-Institut Ankara'nın Türk-Alman Araştırma Eğitim ve İnovasyon Yılı çerçevesinde ortaklaşa düzenlediği Uluslararası Araştırma Verilerinin Yönetimi Çalıştayı 21 Kasım 2014 tarihinde Hacettepe Üniversitesi Kültür Merkezinde gerçekleştirilmiştir. Tüm gün süren çalıştaya 100'ün üzerinde delege katılmıştır. Bu çalışmada çalıştay kapsamında araştırma verilerinin yönetimi üzerine gerçekleştirilen sunumlar ve tartışmalara yönelik bilgiler paylaşılmaktadır.

Anahtar Sözcükler: Araştırma verisi; araştırma verilerinin yönetimi; Türk-Alman Bilim Yılı.

Abstract

Co-organized by Hacettepe University Department of Information Management and The Goethe-Institut Ankara to commemorate the 2014 German-Turkish Year of Research, Education and Innovation, an International Workshop on Research Data Management was held at Hacettepe University's Culture Center on 21 November 2014. More than 100 people attended to the full-day workshop. Information about presentations and the discussions that took place on research data management during the workshop was presented in this study.

Keywords: Research data; research data management; Turkish-German Science Year.

Giriş

2014 Türk-Alman Araştırma Eğitim ve İnovasyon Yılı dolayısıyla Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü ile Goethe-Institut Ankara tarafından 21 Kasım 2014 tarihinde Ankara'da "Uluslararası Araştırma Verilerinin Yönetimi Çalıştayı" düzenlendi.¹ Hacettepe Üniversitesi Kültür Merkezinde yapılan ve TÜBİTAK tarafından da desteklenen çalıştaya Almanya'nın çeşitli üniversitelerinden beş, Türkiye'den (biri konuk öğretim üyesi olmak üzere) üç konuşmacı katıldı. Çalıştay 100'ün üzerinde delege tarafından izlendi ve sunular

* Prof. Dr., Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü, tonta@hacettepe.edu.tr

** Doç. Dr., Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü, umutal@hacettepe.edu.tr

¹ <http://rdm.bilgiyoneti.net/>

web üzerinden de canlı olarak yayımlandı. Çalıştay sırasında eş zamanlı çeviri yapıldı. Çalıştay Goethe-Institut Ankara Müdürü Dr. Thomas Lier ve Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü Başkanı Prof. Dr. Serap Kurbanoglu'nun açılış konuşmalarıyla başladı. Aşağıda bu çalıştayda sunulan bildiriler ve çalıştay sırasında yapılan tartışmalar kısaca özetlenmektedir.

Araştırma Verilerinin Yönetimi

Çalıştay Düzenleme Komitesi Eş Başkanı Yaşar Tonta “Araştırma Verilerinin Yönetimi: İyi Araştırma Uygulamasının Sınırlarını Genişletmek” başlıklı bir açılış konferansı verdi.² Tonta'nın konuşmasında “büyük veriler” in ve bu veriler üzerinde metin ve veri madenciliği yapmanın önemine, araştırma verilerinin yönetimine (AVY) ve veri yönetim politikalarına değinildi. Açılış konferansında üzerinde durulan konular aşağıdadır:

Dünyada her gün 2 exabyte veri yaratılmaktadır. Bu verilerin %90'ı son iki yılda üretilmiştir. Veri miktarı üç yılda bir ikiye katlanmaktadır. Finans alanında sadece ABD'de günde yedi milyar hisse el değiştirmektedir. Bu hisselerin işlemlerinin üçte ikisi bilgisayarlar tarafından gerçekleştirilmektedir. Facebook'a saatte 10 milyon yeni fotoğraf yüklenmektedir. Google her gün Amerikan Kongre Kütüphanesindeki tüm basılı kitapların içerdiği bilgilerin binlerce katından daha fazla (24 petabyte) veri işlemektedir. YouTube'a saniyede bir saatlik video yüklenmekte, Twitter'a her gün yüz milyonlarca tweet gönderilmektedir. Bu verilerin büyüklüğünü algılamak kolay değildir. Ama bu verilerin hepsi kitap olarak basılsaydı ABD'nin tüm yüzeyini 52 tabaka kitapla kaplamak mümkün olurdu.

Yaklaşık 1000 yıl önce bilim amprikti ve gözlem yaparak toplanan verilere dayanarak olgular tanımlanıyordu. Birkaç yüz yıl önce modeller kullanılarak kuramsal genellemeler yapmaya olanak veren kuramsal bilim yapılmaya başlandı. Son birkaç on yılda ise karmaşık olguları bilgisayarlı hesaplama (computation) yoluyla benzetmeye (simülasyon) dayanan bilim gelişti. Günümüzde veri keşfi yoluyla bilim (e-bilim) yapılmaktadır.

Mevcut araştırma verilerinin çok azı erişilebilir durumdadır. Araştırma verilerini bir buzdağına benzetecek olursak buzdağının suyun üstünde kalan kısmında sadece yayımlanmış makale ya da raporlarda yer alan kısıtlı sayıda tablo ve şekillerdeki veriler bulunur. Oysa buzdağının suyun altında kalan çok daha büyük kısmında sırasıyla işlenmiş veriler, veri derlemleri ve yapısal veri tabanları, ham veriler ve veri dizileri bulunmaktadır.

Okuma hakkının metin ve veri madenciliği (MVM) (text and data mining –TDM) hakkını da içermesi gerekir. Mevcut telif hakkı yasalarına istisnai hükümler eklenerek ya da lisans anlaşmaları yoluyla MVM'yle ilgili entellektüel mülkiyet hakkı sorunlarına çözüm bulma girişimleri ölçeklenebilir değildir. MVM'yle ilgili ticari ve ticari olmayan kullanım ayırımına son verilmeli, mevcut metin ve verilerden ticari amaçlarla da yararlanılabilmelidir.

Araştırmaya ABD'de yılda 60 milyar dolar, Avrupa Birliği (AB) Ufuk 2020 (Horizon

² Çalıştay programı ve çalıştayda yapılan sunumların slaytları için bkz. <http://rdm.bilgiyonetimi.net/program.html>. Sunuş slaytlarında genellikle yararlanılan kaynaklara da yer verilmektedir.

2020) Programında ise 80 milyar Avro bütçe ayrılmaktadır. Bu araştırmalar sırasında yaratılan/ üretilen veriler her yıl yüzde 30 artmaktadır. Örneğin, insan genomu projesinde yılda 3 terabyte (TB), İsviçre’deki Avrupa Nükleer Araştırma Kurumunda (CERN) ise yılda 30 petabyte (PB) veri üretilmektedir. Büyük verileri kullanıma açarak ve bu verileri paylaşarak Avrupa ekonomisine yılda 150 ile 300 milyar avro katkı sağlanabileceği tahmin edilmektedir. Büyük verilerin kullanıma açılmasıyla yeni bilimsel araştırmaların tetiklenmesi, bulguların doğrulanarak “kötü bilim” yapılmasının önüne geçilmesi ve bilimsel araştırmalara yapılan yatırımın getirisinin artırılması mümkündür.

ABD’de Ulusal Bilim Vakfı (NSF) araştırma önerilerine araştırma sırasında üretilen verilerin nasıl yönetileceğini açıklayan iki sayfalık bir veri yönetim planı (VYP) (data management plan –DMP) eklenmesini istemektedir. AB Ufuk 2020’de ise açık veri pilot çalışması yürütülmektedir. Pilot çalışmaya katılmak isteyen ve kazanan proje önerilerinin sahiplerinden proje başladıktan sonraki altı ay içinde bir veri yönetim planı istenmektedir. Birleşik Krallık’ta da kamu kaynaklarıyla yapılan araştırmalardan üretilen çıktılara erişime, verilerin yönetimine ve uzun dönemli saklanmasına önem verilmektedir. Ancak henüz veri yönetim planlarına uygun davranmayan proje sahiplerine herhangi bir yaptırım uygulanmamaktadır.

Bilimcilerin verileri uygun ve etkin bir biçimde yönetmelerine olanak veren bir alt yapı olmadan herhangi bir veri paylaşım kültürü geliştirilmesi mümkün değildir. Bu nedenle ortak veri hizmetlerini destekleyen, hem veri üretenlere hem de kullanıcılara yardımcı olan bir ortak veri alt yapısının kurulması ve işletilmesi gerekmektedir.

Araştırma Verilerinin Yönetimi: Almanya ve İsviçre Üniversitelerinden Örnekler

Almanya Göttingen Devlet ve Üniversitesi Kütüphanesinden Dr. Birgit Schmidt kütüphanelerde araştırma verilerinin yönetimi (AVY) uygulamaları, AB’de Alman Araştırma Vakfı (DFG) ve Göttingen Üniversitesindeki AVY politikaları, AVY destek ve alt yapıları hakkında bir sunuş yaptı. Dr. Birgit sunuşunda Avrupa Araştırma Kütüphaneleri Derneği LIBER’in Bilimsel İletişim ve Araştırma Alt Yapıları Komitesi³, Araştırma Verileri Birliği’ne (Research Data Alliance)⁴, Belmont Forum⁵ ve Alman Ağlaşmış Bilgi Girişimi’nin (DINI) araştırma verilerinin yönetimiyle ilgili çalışmalarına değindi. LIBER’in araştırma kütüphanelerinin araştırma verilerinin yönetimine el atmaları için 2012 yılında hazırladığı 10 öneriyi ve örnek olayları, bu örnek olaylardan destek hizmetleri, personel ve personel becerileri konularında çıkarılan sonuçları, Alman Araştırma Vakfının AVY politikasını, Avrupa Komisyonu’nun (AK) Ufuk 2020 Veri Yönetimi İlkelerini ve Açık Veri Pilot Çalışmasını⁶ özetledi.

Dr. Birgit daha sonra Göttingen Üniversitesi hakkında kısa bilgi verdikten sonra üniversitenin AVY politikasının Temmuz 2014’te yayımlandığını, 28 Ağustos 2014’te resmileştiğini, politika belgesinin Alman Rektörler Konferansının (Mayıs 2014) önerileriyle

³ <http://libereurope.eu/committee/scholarly-research/>

⁴ <https://rd-alliance.org/>

⁵ <http://www.bfe-inf.org/info/about>

⁶ http://ec.europa.eu/research/parPcipients/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

ilişkilendirildiğini belirtti.⁷ Göttingen Üniversitesi AVY politikası olan ilk üniversitelerden biridir. Politikada araştırma verileri, AVY ve amaçları, veri yönetim planları, destek ve eğitim hizmetleri, depolama yeri, ahlaki ve yasal standartlar ile açık erişim konularına değinilmektedir. Bu zamana kadar AVY bağlantı kişisi belirlenmiş, bazı prototip AVY hizmetleri geliştirilmiştir (örneğin, ortaklaşa projelerde gömülü AVY yöneticileri bulunmaktadır; seçilmiş disiplinler için alt yapı ve araçlar geliştirilmiştir; kütüphane ve bilimsel bilişim merkezi tarafından alt yapı sağlanmıştır). Ortaklaşa projeler arasında Max Planck Enstitü'sünün Dini ve Etnik Çeşitlilik Araştırması, Dijital İnsani Bilimler Araştırma Birliği, Biyolojik Veriler için ulusal alt yapı yaratılması (GFBio) da bulunmaktadır. Almanya'da Bielefeld, Hamburg, Berlin ve Heidelberg Üniversiteleri de AVY politikalarına sahiptir.

2014'te Göttingen eAraştırma Birliği (eRA) kurulmuştur. Amacı Göttingen yerleşkesindeki e-araştırma etkinliklerini eşgüdümlemek, gerekli bilişim ve bilgi teknolojileri ve bilgisayar desteğini sağlamak, dijital yöntem ve teknolojilerin kullanımı konusunda personeli desteklemek ve proje önerilerindeki veri yönetimiyle ilgili kısımları gözden geçirmek ve önerilerde bulunmaktır. Birlik personeli fen bilimleri, insani bilimler ve bilgisayar bilimleri eğitimi olan kişilerden oluşmaktadır. Personel her konuda uzman değildir ama e-araştırma konularıyla ilgili geniş bir uzmanlığa sahiptir. eRA araştırmacılara AVY, görselleştirme teknikleri, e-araştırma alt yapıları ve proje önerilerinde AVY ve VYP konusunda yardımcı olmaktadır. Ayrıca erişim denetimi, güvenli referanslar ve veri yayını ile ilgili alt yapı hizmetlerini kurmakta ya da mevcut olanları uyarlamaktadır. eRA personelinin yapmayı planladıkları arasında kurumsal AVY politikasının kurulmasını desteklemek ve izlemek için iş akışları yaratmak; kütüphane, bilimsel bilişim merkezi ve araştırma destek/yönetim arasındaki işbirliğini güçlendirmek; enstitülere AVY eğitimi vermek; ve hizmetlerin geliştirilmesi ve planlandığı şekilde başlatılması için bir yol haritası geliştirmek bulunmaktadır.

Alman Bilim ve Teknoloji Ulusal Kütüphanesinden (TIB) Dr. Janna Neumann Hanover Üniversitesinin araştırma verilerinin yönetimiyle ilgili etkinliklerini tartıştı. Sunuşuna araştırma verisi ve AVY'nin tanımlarıyla başlayan Dr. Neumann, AVY yaşam döngüsünü özetledi. Hanover Üniversitesini ve kütüphanesini, AVY Destek Hizmetleri Biriminin kütüphane ve proje etkinliklerini tanıttı. Üniversite kütüphanesinin aynı zamanda TIB olarak da hizmet verdiğini vurgulayan Dr. Neumann, AVY ile ilgili etkinliklerin 2005'te araştırma verilerine DOI verilmesiyle başladığını, 2009'da DataCite'in kurulduğunu, yer bilimleri ve alt yapı konusunda projeler yürütüldüğünü, üniversite ve kütüphanenin verilere DOI verilmesiyle ilgili ilkeler geliştirdiğini ve proje önerisi yazmada rehberlik hizmeti verdiğini vurguladı.

2015'ten başlayarak AVY projesi çerçevesinde gereksinimlerin değerlendirileceğini ve AVY destek hizmetlerinin verileceğini söyleyen Dr. Neumann AVY proje takımının kütüphane, Araştırma ve AB Bağlantı Ofisi ile Bilgi Teknolojileri hizmetlerinden oluştuğunu, AVY destek hizmetlerinin VYP'yi üst verileri, veri arşivlerini ve yayını ve atıfları kapsadığını, bu konularda

⁷ <http://www.uni-goettingen.de/en/28-august-2014-research-data-policy-of-the-georg-august-university-goettingen-incl-umg/488918.html>

araştırmacılar için Kasım 2014'te bir çalıştay düzenlediklerini belirtti ve çalıştayda üzerinde durulan konuları özetledi.

Alman Bilim ve Teknoloji Ulusal Kütüphanesinden (TIB) Bayan Frauke Ziedorn ise araştırma verilerine tekil kimlik vermeye yarayan DOI numaralarını ve DataCite projesini tanıttı. Araştırma verilerinin yayımlanmasının başlıca yararlarını sıraladı (verileri yayımlanan makalelerin daha yüksek atıf alması, daha kolay yeniden kullanılabilirlik, verilerin doğrulanabilmesi, veri toplama ve belgelemenin tanınması, araştırma fonu sağlayanların koşullarına uyulması, ikilemenin önlenmesi ve yeni araştırmaları güdülemesi).

TIB'de araştırma verileriyle ilgili çalışmalara 2003'te Alman Bilim Vakfı tarafından desteklenen bir projeyle başlanmıştır. 2005'te veri dizilerine DOI numarası atanmaya başlanmış, 2009'da yedi üyeye Londra'da DataCite kurulmuştur. 2014'te DataCite'in 18 ülkeden 22 üyesi vardır ve yaklaşık dört milyon DOI numarası verilmiştir. DataCite'in iki çalışma grubu (İşletme Uygulamaları ve Üst Veri) vardır. İlk çalışma grubu iş modelleriyle ilgili ilkeleri yayımlamıştır.⁸ Bu ilkeler DOI numarası atayan ajanslarla (ücretler, öne ekler, üyelik, sözleşmeler vd.) ve müşterilerle (verilere kalıcı erişim için adanmışlık, ana sayfalar, sözdizim, üst veri girme zorunluluğu vd.) ilgili bilgiler ile veri dizileri, veri makaleleri, gri yayın gibi kavramların tanımlarını vermektedir. Üst Veri çalışma grubu ise DataCite Üst Veri Şemasını geliştirmiştir.⁹ DataCite alt yapısı (üst veri deposu¹⁰ ve test sistemi¹¹), DataCite hizmetleri (arama,¹² istatistikler,¹³ Açık Arşiv Girişimi Üst Veri Harmanlama Protokolü (OAI-PMH) hizmeti¹⁴ ile ilgili ayrıntılı bilgilere ilgili web sayfalarından erişilebilir. DataCite'a üye olanlar yerel veri merkezleriyle çalışabilir, DOI numaralarını kaydettirebilir, çalışma gruplarında görev alabilir ve DataCite'in karar verme mekanizmasında yer alabilirler.

Regensburg Üniversitesi Kütüphane Müdürü Dr. Rafael Ball veri toplama aşamasından veri yönetim planı hazırlanmasına kadar kendi üniversitesindeki araştırma verilerinin yönetimi sürecini ayrıntılarıyla anlattı. Dr. Ball Regensburg Üniversitesini kısaca tanıttıktan sonra araştırma verisini tanımladı ve arkeoloji, kristalografi gibi farklı alanlarda verinin ne anlama geldiğini tartıştı. Üniversite kütüphanesinde verilen araştırma hizmetlerini ve açık erişimi açıkladı. Regensburg Üniversitesi Kütüphanesinin araştırma hizmetleri arasında aşağıdakiler yer almaktadır:

- Destek: Proje başvuruları için niyet beyanının hazırlanması; veri türleri, formatları, büyüklüğü ve sunumu ile ilgili danışma hizmeti verilmesi; erişilebilirlik (açık, bedava, kısıtlı); ve yasal çerçeve (veri gizliliğini koruma, entellektüel mülkiyet hakları);

⁸ <http://dx.doi.org/10.5438/0007>

⁹ <http://schema.datacite.org>

¹⁰ <http://mds.datacite.org>

¹¹ <http://test.datacite.org>

¹² <http://search.datacite.org>

¹³ <http://stats.datacite.org>

¹⁴ <http://data.datacite.org>

- Açık erişim dergileri;
- Yayın sunucusu: Yayınlar, araştırma verileri, yayınlar ve verileri ilişkilendirme; farklı veriler (görüşmeler, araştırmalar, veri ve ölçümler) için farklı üst veri formatları; hakların yönetimi (açık veri, kopya iste düğmesi);
- Sanal laboratuvar: Dijitalleştirilmiş yayınların verileri; veri işleme (bilgi çıkarma, dizinleme, ilişkilendirme); yayın yapma;
- Alt yapı: Üniversitedeki sunucu mimarisi; yedekleme ve veri güvenliği; uzun dönemli koruma; veri formatları; yeniden kullanım arayüzleri.

Regensburg Üniversitesi Kütüphanesi Araştırma Hizmetleri Biriminin gelecekte yapmayı planladığı etkinlikler arasında veri yönetimi politikası, veri yönetim planları, veri iş akışları sağlama ve yeni hizmetler yaratma bulunmaktadır.

Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü Öğretim Üyesi Dr. Fuat Akal daha önce çalıştığı Zürih İşlevsel Genomik Merkezinde (FGCZ) yaşam bilimleri veri yönetim sistemine neden gereksinim duyulduğunu ve bu amaçla kurdukları B-Fabric uygulamasını katılımcılarla paylaştı. Zürih Üniversitesi ve Zürih Federal Teknoloji Enstitüsü (ETHZ) Ortak Araştırma ve Teknoloji Merkezi analitik teknolojiler ve uzmanlık sağlayarak yaşam bilimleri projelerini desteklemektedir. FGCZ’de 40 moleküler biyoloji, biyokimya, analitik kimya, bilgisayar bilimleri, fizik ve matematik uzmanı çalışmakta, merkezde her yıl 200 yeni proje yapılmakta, yaklaşık 500 paralel proje yürütülmektedir. Yılda 6000 karmaşık biyolojik örnek analiz edilmekte, yüzlerce gelecek kuşak dizilim deneyleri yapılmakta, yılda 500 terabyte analitik veri üretilmektedir.

Merkezde tüm projeyi/sipariş yaşam döngüsünü izlemek; veri yakalamak, yönetmek ve sağlamak; hızlı analiz yapmak; gerektiğinde verileri yeniden üretmek; erişim denetimli veri paylaşımına izin vermek; kaynakları yönetmek, izlemek ve fiyatlandırmak; hizmetleri ve işlevselliği dinamik olarak eklemek ve gözden geçirmek; istatistikler ve raporlar yaratmak; proje/sipariş başvuru/yönetim zaman ve maliyetlerini azaltmak; ve hizmet (veri) kalitesini geliştirmek için yaşam bilimleri veri yönetim sistemine gereksinim duyulmuştur. Bu amaçla 2005 yılında B-Fabric projesi geliştirilmeye başlanmış; yazılımın ilk sürümü 2007’de tamamlanmış; takvim modülü 2008’de, sipariş modülü 2009’da işletmeye alınmış; modüller 2011’de birleştirilmiş ve sistem 2012’de tamamen yeniden tasarlanmıştır. Dr. Akal B-Fabric projesinin felsefesini (tüm ilgili verileri yakalayacak kadar jenerik olması), mimarisini, farklı modüllerin işlevselliklerini, veri ithal ve açıklama (annotation) iş akışlarını, veri analizini, dışsal uygulamaları, veri aramalarını ve yazılımın yararlarını ayrıntılı olarak açıkladı. Projeden çıkarılan dersler ise şunlardır: Esnek alt yapı; hızlı bütünleştirme; yeniden kullanım; Google-vari arama; kullanıcı özgürlüğü; güvenlik ve gizlilik; ve veri kalitesi.

Araştırma Verilerinin Yönetimi ve Entellektüel Mülkiyet Hakları

Konstanz Üniversitesi emekli öğretim üyelerinden Prof. Dr. Rainer Kuhlen madenleme hakkı diye bir hak var mıdır, yoksa bu hak lisans anlaşmalarına mı tabidir sorusunu tartıştı. Prof. Kuhlen

metin ve veri madenleme (MVM) açısından entellektüel mülkiyet/telif hakları sorunlarını, yayıncıların ve Avrupa Komisyonu'nun MVM meydan okumasıyla, ulusal mevzuatın lisanslarla nasıl başa çıktıklarını özetledi ve MVM açısından telif haklarında yeni bir istisnai hükme gerek olduğunu vurguladı.

MVM metin madenleme ve veri madenleme olarak ikiye ayırmaktadır. Metin madenleme yapısal olmayan büyük metin derlemlerinin ilgili metinleri toplamak; dinlemek; özetlemek; otomatik olarak çevirmek; konuşma, isim, örüntü ya da yüz tanımak; terimler, maddi bilgiler ve olaylar arasında ilişkiler kurmak amacıyla dilbilim ve makine öğrenme teknikleri ve bilgi erişim, doğal dil işleme ve istatistik yöntemler kullanılarak analizi anlamına gelmektedir. Veri madenlemesi ise her tür metin, veri, ses, görüntü ve diğer öğelerin ya da bunların birleşiminden meydana gelen dijital materyallerin yeni bilgiler ortaya çıkarmak amacıyla otomatik olarak işlenmesidir.

Metin ve veri analizi mevcut elektronik bilgilerden yeni bilgiler üretmek ve ilişkiler kurmak, eski araştırmalara dayanan yeni bilimsel bulgular inşa etmek için veri çıkarır (data extraction). Bu yeni yöntemler veri çıkarma sürecinin bir parçası olarak eski eserlerin kopyalanmasını gerektirdiğinden telif haklarına tabidir. Avrupa Birliği Veri Tabanı Yönergesinde MVM'yle ilgili hükümler sınırlıdır. Yayıncılar ise telif hakları mevzuatına MVM'ye izin veren istisnai hüküm eklenmesinin riskli olduğu, bu hususun lisans anlaşmalarıyla düzenlenmesi gerektiği görüşündedirler. Avrupa Komisyonu MVM'ye dayanan veri güdümlü yenilikçiliği (data-driven innovation) artırmak amacıyla telif hakkı da dâhil çözüm önerileri aramaktadır.¹⁵ Birleşik Krallık telif hakları yasasına Ekim 2014'te eklenen bir hükümlerle telif hakkına tabi bir belgeyi okuma hakkına sahip olan bir araştırmacının ticari olmayan amaçlarla MVM için eseri kopyalamasına izin verilmiştir. Bunu önlemek amacıyla yapılan lisans anlaşmaları yasal olarak geçersiz sayılacaktır (yok hükmündedir).

Prof. Kuhlen telif hakkı yasalarına veri madenlemesiyle ilgili muafiyet hakkı olmayan istisnai bir hüküm konulursa o zaman lisans anlaşmalarının MVM'yi engelleyemeyeceği görüşündedir. Çünkü o zaman lisans sahibi veri madenlemeyi yasaklayamaz.¹⁶ Ama bu yapılmazsa lisans sahibi istediği gibi MVM'yi yasaklayabilir. Bu bakımdan lisans maddeleri veri madenlemesinin önünde önemli bir engel olarak kalır, çünkü araştırmacılar bir MVM projesine başlayabilmek için lisans sahibinden izin almak zorunda kalırlar.

Avrupa Adalet Divanı'nın Unutulma Hakkı Kararı

Halen Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümünde konuk öğretim üyesi olarak çalışan Güney Florida Üniversitesi Bilgi Okulu Profesörü Dr. John N. Gathegi Avrupa Adalet Divanı (AAD) tarafından 13 Mayıs 2014 tarihinde yasal hale getirilen “unutulma hakkı”nın (right to be forgotten) ve kişilere bu hakkın tanınmasının araştırma verileri açısından yaratabileceği sorunlar hakkında çok ilginç bir sunum yaptı.

¹⁵ <http://edz.bib.uni-mannheim.de/edz/pdf/swd/2014/swd-2014-0214-en.pdf>

¹⁶ http://ec.europa.eu/internal_market/copyright/docs/studies/1403_study2_en.pdf

AAD unutulma hakkı kararını bir İspanyolun kendisiyle ilgili Google’da yapılan aramalarda 12 yıl önce sosyal güvenlik borçları ile ilgili bir açıklamanın arama sonuçları arasında yer alması, oysa bu sorunun çoktan çözülmüş olması ve dolayısıyla Google’ın kendisiyle ilgili kişisel bilgiler arasında yer alan bu açıklamaya verilen bağlantıları kaldırması gerektiği gerekçesiyle İspanyol Veri Koruma Ajansına (AEPD) ve Google’a karşı açtığı dava sonucu almıştır. Mahkeme arama motoru işleticilerinin veri koruma mevzuatına tabi olduklarına işaret ederek, AEPD’nin arama motoru işleticisinden (Google) davacıyla ilgili bu verileri kaldırmasını ya da belli verilere erişimi kısıtlamasını isteme gücü olduğuna karar vermiştir. Bunun sonucunda AEPD Google’dan söz konusu kişisel verileri dizinden çıkarmasını ve bu verilere erişimi önlemesini istemiştir. İlginçtir, AEPD özgün web sitesindeki verilerin silinmesini gerekli görmemiştir.

Google AEPD’nin kararı nedeniyle temyiz için İspanya Ulusal Yüksek Mahkemesine (UYM) başvurmuştur. UYM sorunun çerçevesini çizerken, kendileri hakkında üçüncü parti kuruluş ya da şirketler tarafından yayımlanan verilere sonsuza kadar erişim sağlanmasını istemeyen kişiler olduğunu not etmiştir. Arama motoru işleticilerinin o zaman kişilerin kişisel verilerini korumakla ilgili ne tür yükümlülükleri bulunmaktadır? UYM bu soruya verilecek yanıtın aynı zamanda İspanya’nın da uymak zorunda olduğu Avrupa Kişisel Verileri Koruma Yönergesinin (95/46/EC) nasıl yorumlandığına bağlı olduğunu söylemiştir. Fakat ilgili yönerge arama teknolojilerinden önce formüle edilmiştir. Bunun üzerine UYM herhangi bir karar vermemiş, dosyayı yönergenin yorumlanması için AAD’ye göndermiş ve AAD’nin özellikle şu sorular hakkında karar vermesini istemiştir:

- Avrupa kişisel verileri koruma yönergesinin AB üyesi ülkeleri bağlayıp bağlamadığı;
- Google’ın üçüncü parti kurumlar/şirketler tarafından yayımlanan kişisel verilerle ilgili arama etkinliklerinin (bulma, dinleme, depolama, erişim) yönergede yer alan “veri işleme” ye girip girmediği;
- Arama etkinlikleri sırasında yapılan işlemlerin, Google’ın, dinlediği kişisel verilerin “deneticisi” (controller) olması anlamına gelip gelmediği;
- Verileri içeren web sitesi sahibinin haberi olmadan Google’ın kendi dizininden bu verileri silmeye zorlanıp zorlanamayacağı;
- Verilerin silinmesi ya da bu verilere erişimin kısıtlanması ve “unutulma hakkı” yla ilgili maddelerin üçüncü partiler tarafından yasal olarak yayımlanmış bilgileri de kapsayıp kapsamadığı.

Bu sorulara AAD’nin verdiği cevaplar ise şöyledir: Kişisel verilerin işlenmesi AB üyesi bir devletin sınırları içinde gerçekleşmektedir. Ama arama motoru işleticisinin (Google) ve onun İspanya’daki branşının (Google Spain) etkinlikleri birbiriyle iç içe geçmiştir. Dolayısıyla yönerge AB üyesi ülkeleri de bağlamaktadır. Google’ın kişisel verileri içeren bilgileri dizinlemesi, bu verileri araması ve bu verilere erişim sağlaması “kişisel verilerin işlenmesi” ne girmektedir. Bu işlemle ilgili olarak arama motoru işleticisi verilerin deneticisi (controller) konumundadır.

Adli bir otorite Google gibi bir arama motoru işleticisinden arama sonuçları arasındaki kişisel verileri ve/ya da kişisel verilere erişimi sağlayan bağlantıları silmesini isteyebilir. Bu karar, verilerin eş zamanlı olarak üçüncü parti şirket ya da kurumun web sitesinden de silinmesi, web sitesi sahibinin verilerin silindiğinden haberdar edilmesi ya da web sitesi sahibinin de verileri silmesi gerektiği anlamına gelmemektedir.¹⁷

AAD'nin unutulma hakkıyla ilgili kararı bilginin yasal olarak yayımlanmış olup olmaması, kamunun temel haklara müdahale edip edemeyeceği, kamu çıkarları (örneğin, ifade özgürlüğü) ve oto sansür gibi konularda tartışma yaratmıştır. Bu karara uymamanın cezası ilgili şirketlerin dünya çapındaki gelirlerinin yüzde 2'sine karşılık gelmektedir.

Peki, unutulma hakkı kararı veri yönetimi açısından ne anlama gelmektedir? Bazı veriler yok olabilir ya da bu verilere erişim zorlaşabilir. Verilere erişim -özellikle verilere arama motorları yoluyla erişim- sonsuza kadar kısıtlanabilir. Bunun özellikle sosyal bilimlerdeki araştırmalar üzerinde etkisi olacaktır. Çoğu araştırma kişilerle doğrudan ilişkilendirilmeyen toplu verilere dayanmaktadır. Çevrimiçi ortamın özelliklerinden biri bireylerin hem çevrimiçi etkileşimler sırasında dijital ayak izlerini bırakmaları hem de başkalarının da kendileri hakkında veri gölgeleri (data shadows) yaratabilmeleridir. Bu ikisi kişisel olarak ayırt edilebilecek bir biçimde (örneğin, ad, fotoğraf) birbiriyle bağlantılıysa unutulma hakkı kararına göre bu bağlantılar silinebilir. Veri madenleme açısından verilere bağlantıların kaybolması muhtemelen bazı disiplinleri daha çok etkileyecektir. Veri simsarlığı işi yapan şirketler bu karardan etkilenecektir, çünkü madenlenen verilerin değeri bireylerle ilişkilendirildiği zaman daha da artmaktadır. Arama motorları aracılığıyla bilgi toplayan belli veri tabanları bu karardan olumsuz yönde etkilenecektir. Google uluslararası bir şirket olduğundan bu kararın etkileri sadece Avrupa'yla sınırlı kalmayacaktır. Dahası, başka ülkeler aynı hakkın kendi vatandaşlarına da uygulanmasını isteyerek Google'ı arama sonuçlarından daha fazla bağlantıyı silmeye zorlayabilirler (ki Arjantin'de ırkçı fotoğrafların kaldırılması için Google ve Yahoo'ya karşı açılmış davalar bulunmaktadır). Avrupa Birliği, ABD'yle yaptığı dış ticarete kişisel ve kurumsal verilerin korunmasının ilkelerini düzenleyen "Güvenli Liman" (Safe Harbour) anlaşmasından çekilebilir. Ya da AB (ifade özgürlüğüyle çelişen) unutulma hakkı da dâhil veri koruma yasalarının ABD'de de uygulanmasını isteyebilir.

Sonuç olarak; AB kararından bir yıl önce, Google'ın kişisel açıdan duyarlı bilgileri arama sonuçlarında daha alt sıralarda göstermesi (yani "verilerin gömülmesi") yönünde bazı istekler vardı. AAD'nin unutulma hakkı kararı bundan daha fazlasını yaptı: Verileri gömmek yerine verilere bağlantıyı kesti; unut, dedi; ya da en azından verilerin üzerine bir perde ört, dedi. Kısacası, veriler orada, ama bulabilirsen!

Yuvarlak Masa Tartışması

Sunuşların ardından Uluslararası Araştırma Verilerinin Yönetimi Çalıştayı yuvarlak masa tartışmasıyla sona erdi. Gerek konuşmacılar gerekse katılımcılar çalıştay sunumlarında öne

¹⁷ AAD'nin unutulma hakkıyla ilgili kararı için bkz. <http://curia.europa.eu/jcms/upload/docs/application/pdf/2014-05/cp140070en.pdf>

çıkan konuları ve sorunları görüştüler. Üzerinde tartışılan konuları ve ortaya çıkan sonuçları şu şekilde özetlemek mümkündür:

- Araştırma verisi nedir? Ticari ve akademik araştırma ayrımı nasıl yapılmaktadır ve söz konusu ayrımın araştırma verilerinin elde edilmesine etkisi var mıdır?
- Araştırma verilerinin yönetimi ile ilgili sorunlar nelerdir? Bu sorunlar arasında yasal düzenlemelerin olmaması, konuyla ilgili yaptırımların gerçekleştirilememesi, etik konular ve güdüleme eksikliği ön plana çıkmaktadır.
- Araştırma verilerinin kavramsallaştırılması kültürden kültüre, disiplinden disipline farklılık göstermektedir.
- Araştırmacılar veri paylaşmakta kıskanç davranmaktadırlar. Bunun önlenmesi için çalışmalar yapılmalıdır.
- Veriler hakkında bilgi toplanması (verilerin üst verisi) ve verilerin tanımlanması önemli bir konudur.
- Verilere atıf yapılması önümüzdeki dönemde popüler bir araştırma alanı olacaktır.
- Metin ve veri madenciliği (MVM) en sıcak araştırma konularından biridir. Gerek yeni bilimsel araştırmaların yapılması gerekse verilerin kullanımından elde edilecek ekonomik yararlar bu alanda bir an önce entellektüel mülkiyet haklarıyla ilgili yasal düzenlemeler yapılmasını gerektirmektedir.
- Araştırma verilerinin yönetimi konusuna bilginin ve bilgi yönetimi bölümlerinin ders programlarında yer verilmelidir.
- Kamu adına karar alıcıların verilerin erişime açılması konusunda duyarlılık göstermeleri gerekmektedir.
- Verilerin kullanılması konusunda telif hakları ile ilgili yasal düzenlemeler eksiktir.
- Açık toplum olmanın önemli bir aşaması verileri açmaktan ve bu verileri tutarlı bir şekilde hizmete sunmaktan geçmektedir.
- AAD'nin unutulma hakkı kararı ve bu hakkın araştırma verileri üzerindeki etkisi önümüzdeki günlerde çeşitli açılardan tartışılacaktır.

Çalıştay, Düzenleme Komitesi Eş Başkanı Yaşar Tonta'nın kapanış konuşmasıyla sona erdi. Tonta çalıştayın düzenlenmesinde emeği geçenlere teşekkür ederek, katılımcıları 24-26 Kasım 2014'te Antalya'da düzenlenen ve ana konusu "Araştırma Verilerinin Yönetimi ve Bilgi Keşfi" olan 5. Uluslararası Değişen Dünyada Bilgi Yönetimi Sempozyumuna¹⁸ davet ederek çalıştayı kapattı.

¹⁸ <http://imcw2014.bilgiyonnetimi.net/>