Öğrenci Verilerinin Korunması: Fatih Projesi Işığında Teknik Değerlendirme‡

Ali İNAN[1], Mehmet Ercan NERGİZ[2], Yücel SAYGIN[3]

[1]Bilgisayar Mühendisliği, Adana Bilim ve Teknoloji Üniversitesi, Adana, Türkiye
[2]Acadsoft Yazılım, Gaziantep, Türkiye
[3]Mühendislik ve Doğa Bilimleri Fakültesi, Sabancı Üniversitesi, İstanbul, Türkiye
ainan@adanabtu.edu.tr, nergiz@gmail.com, ysaygin@sabanciuniv.edu
(Geliş/Received:16.05.2016; Kabul/Accepted:03.11.2016)
DOI: 10.17671/btd.93204

***Özet*—** Mahremiyet temel bir insan hakkıdır ve 2010 tarihinde yapılan Anayasa değişikliğiyle Türkiye Cumhuriyeti vatandaşları için güvence altına alınmıştır. Mahremiyet, 7 Nisan 2016 tarihinde Resmi Gazete'de yayınlanan Kişisel Verilerin Korunması Kanunu ile korunmaktadır. Bu kanun ile beraber hem özel sektör hem de kamu kurumlarında veri koruması konusunun tartışmaya açılması beklenmektedir. Öte yandan ülkemizce eğitim alanında e-Okul ile başlayıp FATİH projesi ile çok daha kapsamlı hale gelen dijital dönüşüm projeleri halen devam etmektedir. Avrupa Birliği'nde bu tarz büyük çaplı projelere başlanmadan önce mahremiyet etki değerlendirmesi yapılması zorunludur. Projelerin tasarım ve uygulanma aşaması ise bu etki değerlendirmeleri dikkate alınarak gerçekleştirilmektedir. Bu makalenin amacı FATİH projesi göz önüne alınarak öğrenci verilerinin korunması ile ilgili teknik hususların değerlendirilmesidir. Bu bakımdan Türkiye'de yazarların bilgisi dahilinde ilk kez yapılan bu çalışmanın ileride farklı sektörlerde benzer çalışmalara vesile olup ışık tutması ümit edilmektedir.

***Anahtar Kelimeler*—** mahremiyet, veri analitiği, bulut bilişim, FATİH projesi

Student Data Protection: A Technical Assessment in the Context of the Fatih Project§

***Abstract*—** Privacy is a fundamental human right that has been coined in the Turkish Constitution since the 2010 amendment. Privacy is protected through Turkey's recent Personal Data Protection Law, published in the Official Gazette of the Turkish Republic on April 7th, 2016. With this new law, it is expected that discussions on data protection will increase in Turkey, in both private and public sectors. In the meanwhile, large scale educational e-transformation projects are in progress in Turkey such as the e-School project and the much more comprehensive FATİH project. In the European Union deployment of such large-scale IT projects is preceded by an obligatory privacy impact assessment. Furthermore, design and implementation phases of such projects are carried out in the light of these privacy impact assessment studies. The aim of this article is to investigate and report the data protection issues for students from a technical perspective in the context of the FATİH project. To the best of our knowledge, this is the first study in Turkey of its kind, and we hope that this work will initiate and shed light on further similar studies.

***Keywords*—** privacy, data analytics, cloud computing, FATİH project

‡ Bu araştırma TÜBİTAK tarafından desteklenen "Eğitimde Mahremiyeti Koruyan Veri Paylaşımı ve Analizi" başlıklı ve 114E261 numaralı proje kapsamında yapılmıştır.

§ This research was funded by The Scientific and Technological Research Council of Turkey (TUBITAK) under grant number 114E261.

1. GİRİŞ (INTRODUCTION)

Mahremiyet, Birleşmiş Milletler İnsan Hakları Bildirgesi tarafından tanınan temel insan haklarından birisidir. Türkiye'de 2010 yılında Anayasa'nın 20. maddesinde yapılan değişiklikle, her Türk vatandaşının kendi özel hayatına ve aile hayatına saygı gösterilmesini talep etme hakkının bulunduğu açıkça vurgulanmıştır. Bireylerin özel hayatlarının ve aile hayatlarının mahremiyetine müdahale edilemeyeceği de ayrıca belirtilmiştir. 2010 Anayasa değişikliğinden yaklaşık 6 yıl sonra bu anayasal hakkı korumaya yönelik kapsamlı bir veri koruma kanunu 7 Nisan 2016 tarihinde Resmi Gazete'de yayınlanarak yürürlüğe girmiştir [1]. Kanunun amacı "kişisel verilerin işlenmesinde başta özel hayatın gizliliği olmak üzere kişilerin temel hak ve özgürlüklerini korumak ve kişisel verileri işleyen gerçek ve tüzel kişilerin yükümlülükleri ile uyacakları usul ve esasları düzenlemek" olarak belirtilmiştir. Yürürlüğe giren çerçeve kanunun eğitim ve sağlık gibi sektörlere has yönetmeliklerle uygulanması beklenmektedir.

Eğitim alanında verilerin korunması, ABD ve Avrupa'da etkin olarak tartışılmaktadır. 15 Ocak 2015 tarihinde, ABD Başkanı Barack Obama, çevrimiçi veya mobil eğitim programları ile toplanan verinin sadece eğitimsel amaçlarla kullanılmasını garanti altına almayı amaçlayan Öğrenci Dijital Mahremiyet Yasası'nı [2] duyurdu. Öğrenci Dijital Mahremiyet Yasası'nın odak noktasından da anlaşılabileceği üzere, duyulan temel kaygı, öğrenci verilerinin özel sektör tarafından suistimal edilmesidir. Kitlesel Açık Çevrimiçi Dersler (KAÇD) de daha çok yüksek öğrenimle ilgili mahremiyet çerçevesinde tartışılmaktadır. Diğer yandan, Türkiye'de eğitim verilerinin korunması konusunun kamuoyu nezdinde yeterince tartışıldığı şüphelidir. Örneğin, Milli Eğitim Bakanlığı tarafından 2015 ve 2016 yıllarında Eğitim Teknolojileri Zirvesi düzenlenmiş, ancak, veri koruması hususu üzerine tartışmalar çok kısıtlı olmuştur [3]. FATİH (Fırsatları Arttırma ve Teknolojiyi İyileştirme Hareketi) Projesi [4] gibi devam etmekte olan geniş kapsamlı e-dönüşüm projeleri eğitim verileri ve eğitim amaçlı kullanılan mobil teknolojiler kapsamında veri koruması konularının Türkiye'de daha da detaylı değerlendirilmesini gerektirmektedir.

FATİH Projesi Türkiye'de 42.000 okulda, 570.000 sınıfı en güncel bilgi teknolojileriyle donatarak, tüm okullara bilgisayarlı ve akıllı sınıf olanağı sunmayı amaçlamaktadır. En güncel bilgi iletişim teknolojilerinin etkili kullanımıyla, eğitimde fırsat eşitliği sağlanacaktır. Tabletler ve yüksek hızlı internet erişimi olan LCD Etkileşimli Panolar, bu projenin temel yapı taşlarını oluşturacaktır. Geniş vizyonu ve büyük hedefleriyle FATİH Projesi, Türkiye'deki en önemli eğitim yatırımlarından birisidir. Bu kadar geniş kapsamlı bilgi teknolojileri projeleri, Avrupa ve ABD'de mahremiyet etki değerlendirmesi (MED) yapılmadan uygulamaya konmamaktadır. Fakat, FATİH Projesi çerçevesinde mahremiyet ve veri koruması açısından kamuoyu ile paylaşılmış herhangi bir değerlendirmeye rastlanmamaktadır. Bu yayının amacı, mobil teknolojilerin eğitim sistemine dahil edilmesine olanak sağlayacak olan FATİH Projesi'nden kaynaklanabilecek mahremiyet risklerinin altını çizmektir. Bu yayının, FATİH Projesi için detaylı bir MED çalışması için başlangıç noktası olabileceği umulmaktadır.

2. AKTÖRLER VE ROLLERİ (ACTORS AND ROLES)

Bu kısımda FATİH Projesi'nin temel aktörleri tanımlanmaktadır. Aşağıda ana hatlarıyla anlatıldığı gibi, her aktör sistemde iyi tanımlanmış ve sınırları belirlenmiş bir rol üstlenmektedir. Aktörler, Veri Koruma Topluluğu [5] içinde genel olarak kabul edilmiş tanımlara göre belirlenmiştir.

2.1. Temel Kavramlar ve Tanımlar (Key Concepts and Definitions)

- Veri Öznesi: Mahrem verisi işlenen (toplanan ve analiz edilen) kişi. FATİH Projesi'ndeki veri özneleri Öğrenciler, Öğretmenler ve Okul Yönetimi'dir. Bununla birlikte, veri koruması konusundaki temel odak olduklarından bu raporda öğrenciler üzerine yoğunlaşılacaktır.
- Veri Denetleyicisi: Kişisel verilerin, neden ve nasıl işleneceğini veya işlenmekte olduğunu belirleyen kişi ya da kişi grubu. FATİH Projesi'nde veri denetleyicisi Milli Eğitim Bakanlığı tarafından atanacaktır (atanmalıdır).
- Veri İşleyicisi: Veri denetleyicisinin çalışanı olmayan, fakat veri denetleyicisi adına kişisel verileri toplayan kişi ya da kişi grubu.

Yukarıdaki tanımlarda, *veri işleme* terimi kişisel veri kullanılarak yapılabilecekleri tanımlamakta ve şu eylemleri içermektedir: veri edinme, kaydetme veya depolama; veri üzerinde her türlü işlemi gerçekleştirme. Bu işlemler, aşağıdaki gibi, veriye her türlü fiili erişimi kapsamaktadır:

(a) Düzenleme, uyarlama veya değiştirme,
(b) Erişim, danışma veya başka kullanım,
(c) Açıklama, yayma veya başka şekilde kullanıma sunma,
(d) Sıralama, bağdaştırma, engelleme, silme veya yok etme.

2.2. FATİH Projesi'nin Aktörleri (Actors of the FATİH Project)

FATİH sistemi temel olarak öğrencilere hitap edecektir. Projede yer alması planlanan diğer meşru aktörler şunları kapsamaktadır:

- Milli Eğitim Bakanlığı'nın (MEB) çalışanları (öğretmenler, okul/ilçe yönetimleri)

- Sistemi tasarlayacak ve çalıştıracak olan mühendis ve teknikerler
- Kişisel veriye (muhtemel) erişimi sağlanacak olan üçüncü şahıs veri analistleri.

FATİH sisteminin her türlü izinsiz erişime karşı tamamen korunabilir olacağını varsaymak gerçekçi değildir. Daha kötüsü ise, meşru aktörlerin gerek istekleri dışında (acemi/bilgisiz olduklarından), gerekse bilinçli olarak sorumsuz davranışları olacaktır. Bu kapsamdaki her türlü davranış *saldırgan* kabul edilecektir.

FATİH sistemi için *saldırgan*, sisteme ya da aktörlerine - özellikle öznelerine, zarar vermeyi amaçlayan kişidir. Saldırgan davranışlar aşağıdakileri kapsamaktadır:

- Sistemin ya da bir hizmetin elverişliliğini bölgesel ya da geniş çaplı olarak düşürmek (hizmet engelleme saldırıları).
- Erişim kontrol mekanizmasını es geçerek erişim izinleri ele geçirmek veya mevcut izinleri yükseltmek (içerden saldırı senaryosu).
- Geçmiş ve/veya dış kaynaklı bilgiler kullanarak öznelerle ilgili hassas verileri açığa vurmak (üçüncü şahıs analistler için tipik saldırı durumu).

2.3. Aktör-Rol Eşleştirmeleri (Actor-Role Mappings)

Aktörlerle roller için eşleştirme Tablo 1 üzerinde sunulmuştur. Mimari kararlara bağlı olarak bu atamalarda ufak değişiklikler olabilir. Bunlar tabloda soru işareti '?' ile belirtilmiştir.

Tablo 1. Aktörlerin FATİH sistemindeki rolleri (Roles of the actors in the FATİH system)

Aktör-rol eşleştirmesi		Roller		
		Özne	*Denetleyici*	*İşleyici*
Aktörler	Öğrenciler	✓	?	×
	MEB çalışanları	?	✓	×
	Teknik çalışanlar	×	✓	×
	3. parti analistler	×	×	✓

3. VERİ TOPLAMANIN AMAÇLARI (PURPOSES FOR DATA COLLECTION)

FATİH Projesinin amaçlarından biri, öğrencilere e-öğrenme ile kendilerini geliştirecek imkanların sağlanması ve e-içerik üretilmesidir. Diğer yandan, öğretmenler de kendi öğretim yöntemlerini hem niteliksel hem de niceliksel olarak değerlendirebileceklerdir. Böylece öğretmenler kendi öğretim yöntemlerini geliştirebilme fırsatına sahip olacaklardır. Devlet, yetenek ve becerileri tespit ederek gelecek ihtiyaçlara yönelik önlemler alabilecektir. Bu amaçlara ulaşmak için veri, anahtar rol oynamaktadır. İrdelemelerimizi işletimsel amaçla ve veri analizi amacıyla toplanan veriler üzerinde yapacağız.

3.1. İşletimsel Amaçlar (Operational purposes)

İşletimsel amaçlar, öğrencilerin değerlendirilmesine olanak vererek okulun ve eğitim sisteminin işleyişini sağlamaktadır. Veri toplama amacı devamsızlık takibi, çalınan cihazların takibi veya güvenlik olabilir.

3.2. Veri Analitiği (Data Analytics)

Eğitim verisi üzerinde yapılan veri analitiği, bireysel olarak öğrenci için ve eğitim sistemini iyileştirmek için faydalı olabilir. Eğitim sistemi, milyonlarca öğrenci ve yüz binlerce öğretmenden elde edilen verilere dayanarak daha iyi hale getirilebilir. [6] Örneğin, eğer öğrencilerin çoğunluğu belirli bir konuda başarısız oluyorsa, bu ders içeriğinin veya dersin sunum şeklinin sorunlu olduğunun göstergesidir. Öğrenci için, yetenekleri ve ilgilerine göre kişiselleştirilmiş içerik ve öğretim yöntemleri sağlanabilir. Dersler ve ders içerikleri tablet aracılığıyla toplanan davranışsal verilere dayanarak değerlendirilebilir. En üst seviye karar verme, milyonlarca öğrenciden elde edilen büyük veri üzerinden yapılabilir. [7]

4. SİSTEM MİMARİSİ KONULARI (SYSTEM ARCHITECTURAL ISSUES)

Sistem mimarisi ile ilgili pek çok seçenek bulunmaktadır. Bu seçeneklerden, mahremiyet açısından potansiyel etkiler doğurması beklenenler aşağıda farklı başlıklar altında tartışılmaktadır.

4.1. Bulut Bilişim Kullanımı (The Use of Cloud Computing)

Mobil cihazlar ile sağlanacak eğitsel hizmetler için ideal ortamın mobil uygulamalar olması beklenmektedir. Bu bölümde, bu uygulamaların bulut bilişim teknolojilerinden ne oranda faydalanacağına dair seçenekler irdelenmektedir.

Uygulamaların ve/veya uygulama verisinin bulutta depolanmasına göre ortaya çıkan mimari çözümler şu şekilde sıralanabilir:

Tamamen-bulut mimari: Hem uygulamalar, hem de uygulamalar tarafından üretilen verinin buluta yerleştirildiği çözümdür. Uygulamalar ve ilgili veri MEB tarafından uzak sunucularda saklanır. Veri analizi, uygulama güncellemeleri merkezi olarak bulut üzerinde yapılır. Böyle bir mimaride, mobil cihazlar internet hizmetleri için sadece basit bir arayüz olarak işlev görmektedir.

Yerinde mimari: Uygulamalar ve uygulamalar tarafından üretilen veri, FATİH Projesi kapsamında öğrencilere verilen mobil cihazlarda yerel olarak tutulabilir.

Bulut-üzerinde-veri mimari: Uygulamalar mobil cihazlara yüklenebilir fakat uygulamalar tarafından üretilen veri bulut üzerinde saklanır. Bulut üzerindeki veri, şifrelenmiş olarak veya şifresiz olarak tutulabilir. Şifresiz metin yaklaşımı tamamen-bulut mimariye benzer artı ve eksilere sahiptir. Bu nedenle, verinin uygulama tarafından saklanması için buluta aktarılmadan önce şifrelenmesi üzerinde durulacaktır.

Çözüm yöntemlerinin faydaları aşağıda ana hatlarıyla açıklanmıştır. Tablo 2'de ise yöntemlerin karşılaştırması özet olarak sunulmuştur.

Tablo 2. Veri saklama ve uygulama mimarisi
(Architecture for data storage and applications)

Farklı mimariler seçeneklerin özellikleri		Mimari		
		Tamamen-bulut	*Yerinde*	*Bulut-üzerinde-veri /şifreleme*
Özellik	Her zaman çevrimiçi olma gereksinimi	evet	hayır	kısmi
	Cihazın kaynak gereksinimleri	düşük	yüksek	orta
	Mahremiyet üzerinde kullanıcı kontrolü	zayıf	güçlü	güçlü
	Veri koruma	güçlü	zayıf	güçlü
	Başarılı bir saldırının etkisi	yüksek	düşük	düşük

Tamamen-bulut çözümün şu faydaları sağlanması öngörülmektedir:

- Bulut tabanlı bir ekosistemde, mobil cihazlar daha az kaynağa ihtiyaç duyar ve bakım masrafları önemli ölçüde azalır.
- Hem veri, hem de uygulamalar bulut üzerinde her zaman günceldir.
- Bulut uygulamalarında veri, sunucular üzerinde depolandığından veri analitiği kolaylaşır.
- Kullanıcı tarafından yönetilen sistemlere kıyasla bulut, genellikle daha güçlü ve güvenilir veri koruması sunar.

Yerinde çözümü öne çıkaran faydalar ise şu şekildedir:

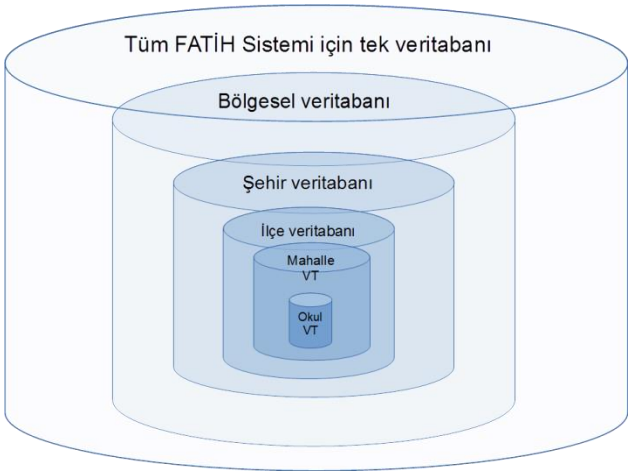
- Bulut tabanlı sistemler kaliteli ve güvenilir internet bağlantısı gerektirir. Buna karşılık yerinde sistemler çevrimdışı olarak da hizmet sunabilmektedir.
- Yerinde sistemlerde son kullanıcılar mobil cihazlarını kendi güvenlik tercihlerine göre özelleştirebilirler.
- Yerinde sistemlerde son kullanıcılar kendi verilerinin başkaları tarafından kullanımı üzerinde daha fazla kontrol sahibi olabilirler.
- Verinin gizliliği ve hizmetin elverişliliği üzerine yapılan başarılı bir saldırının etkisi, bulut tabanlı sistemlerde çok daha fazla olacaktır.

Bulut-üzerinde-veri çözümünün ise şu faydaları sağlaması beklenmektedir:

- Tamamen-bulut sistemlerde olduğu gibi, bulut-üzerinde-veri kaliteli ve güvenilir internet bağlantısı gerektirir. Bununla birlikte, eğer az miktarda verinin (örneğin, en son kullanılan veri) uygulama üzerinde yerel olarak saklanmasına izin verilirse, çevrimdışı durumdayken bile uygulamaların kısıtlı olarak işlev göstermesi mümkün olacaktır.
- Tamamen-bulut sistemlerde olduğu gibi, bulut üzerinde saklanan veri daha güvenilir şekilde korunacaktır. Başarılı bir saldırının etkisi de daha sınırlı olacaktır.
- Şifrelenmiş veri uygulama dışında başka kimse tarafından okunamayacağından, yerinde mimaride olduğu gibi, kişisel veri başka kimse ile paylaşılmayacaktır. Bulut-üzerinde-veri, bu şekilde kullanıcı gizliliği sunmaktadır. Buna karşılık olarak, bulut üzerinde veri analitiği kullanıcının izni ve işbirliği olmaksızın mümkün olmayacaktır.

4.2. Veri Yönetimi ile İlgili Konular (Issues Related to Management of Data)

Öncelikle farklı veritabanı dağıtıklık seviyelerinin etkisi incelenecektir. En dağıtık senaryoda, her okulda sadece o okulun öğrencilerinin verisinin saklanacağı bir veritabanı yönetilecektir. Bunun tam zıttı olarak, en merkezi senaryoda ise FATİH Projesi kapsamındaki bütün veriler tek bir veritabanında saklanacaktır. Ara çözümler arasında, her mahalle, ilçe, şehir ve bölge için bir veritabanı tutmak bulunmaktadır. Bu durum Şekil 1'de tarif edilmiştir.



Şekil 1. Veritabanlarının idaresi için farklı taneciklilik seçenekleri (Various granularities at which databases may be managed)

Merkeziyet seviyesi/verinin dağıtıklığının sistem üzerindeki etkileri aşağıdaki gibidir:

- **Veri elverişliliği:** Bir veritabanı sunucusunun hatası, hatalı sunucuda kendisiyle ilgili veri bulunan bütün veri öznelerini etkiler. Bu nedenle, tamamen merkezi bir sistemin kurulması veri elverişliliği açısından "tekil hata noktası" oluşturur.
- **Mahremiyet:** Bir veritabanı sunucusuna izinsiz erişim olması, o sunucu üzerinde kendisiyle ilgili veri bulunan bütün öznelerin verilerini açık etme tehlikesi taşır. Dolayısıyla, tamamen merkezi bir veritabanı sistemi, veri mahremiyeti açısından da "tekil hata noktası" oluşturmaktadır.
- **Geniş çaplı analitik:** Yüksek seviye merkeziyet, daha büyük veritabanlarının oluşmasına neden olur. Buna bağlı olarak, her türlü geniş çaplı analizin veya veri-odaklı işlemlerin merkezi çözümlerde daha hızlı yapılabilmesi beklenir.
- **Ağ gecikme süresi:** Dağıtık sistemlerde sağlaması en zor şeffaflık şekli "konum şeffaflığı" [8], yani kaynağın nerede bulunduğunu gizlemektir. Veritabanlarını veri öznelerinin coğrafi dağılımına göre dağıtmak olası çözümlerden biridir. Böylece veri, hizmet isteğine daha yakın olacak, istek/yanıt süreleri en aza indirgenecektir. Bu sayede, yerel işlemlerin dağıtık sistemlerde daha hızlı çalışması beklenmektedir.
- **Güvenlik:** Dağıtık sistemlerde güvenlik unsurlarının da veri ile birlikte dağıtılması gerekmektedir. En düşük seviye olan okullar seviyesinde veritabanı saklanması halinde, yeterli güvenlik uzmanlığının bu ölçekte sağlanamaması tehlikesi ortaya çıkmaktadır. Öte yandan merkezi sistemlerde ortaya çıkacak bir güvenlik açığı, tüm veritabanının tehdit altında olmasına sebep olacaktır.

Dağıtıklık seviyesi veri yönetiminin bir boyutudur. Bir diğer önemli boyut ise kullanılacak veritabanı yönetim sistemidir. Burada iki alternatif ön plana çıkmaktadır.

İlişkisel veritabanı yönetim sistemleri (İVTYS) verinin mutlak bir düzen içerisinde depolanmasına ve bu veri ile ilgili tüm işlevsel bağımlılıkların harfi harfine modellemesine ihtiyaç duyar. Bu gereksinimi ortadan kaldıran, güçlü ve yeni bir alternatif ise, NoSQL (Not-Only-SQL, SQL ve Daha Fazlası) veritabanlarıdır. NoSQL veritabanlarının veri modeli, yapılanmamış, çoğunlukla belge-tabanlı veya anahtar-değer bazlıdır.

Bu iki teknolojinin getiri ve götürüleri aşağıda kısaca tartışılmıştır.

Güvenlik ve erişim kontrolü: İlişkisel VTYS 1970'lerden beri, hareketsel sistemlerinden analitik işlemeye kadar çok çeşitli uygulamalarda kullanılmaktadır. Sonuç olarak, çoğu ilişkisel VTYS çağdaş erişim kontrol düzeneklerini (rol-bazlı erişim kontrolü [9] gibi) ve şifrelenmiş/korumalı depolama düzeneklerini tümüyle desteklemektedir.

Hareket desteği: Bir hareket, mutlaka bütün olarak çalıştırılması gereken bir dizi veritabanı işleminden oluşmaktadır. ATYS özellikli (atomiklik, tutarlılık, yalıtım, sağlamlık) veritabanları hareketsel işlemleri destekler.

Neredeyse bütün ilişkisel VTYS hareketseldir, ancak NoSQL VTYS biraz daha karışık bir durumdadır: çoğu NoSQL teknolojisi (MongoDB, BigTable, Cassandra gibi) hata durumunda elverişliliği arttırmak ve paralel yürütme ile performansı arttırmak amacıyla verinin kopyalanmasına dayanır. Bununla birlikte, kopyaların idaresinde "gecikmeli tutarlılık" [10] adı verilen bir yol izlenir. Bu kavrama göre, tutarlılık anlık olarak gerekli olmayan ancak belirli bir süre zarfında yerine getirilecek bir gereksinim olarak düşünülür. Bu nedenle, çoğu NoSQL VTYS, tümüyle hareket desteğine *henüz* sahip değildir.

Eğer FATİH sisteminin hareket işlemlerini tümüyle desteklemesi hedefleniyorsa, NoSQL VTYS yerine ilişkisel VTYS tercih edilmelidir.

Ölçeklenebilirlik: NoSQL VTYS kopyalama/verinin parçalanması üzerine dağıtılmış mimarileri sayesinde büyük veri ile doğal olarak baş edebilmektedir. Ölçeklenebilirlik bakımından ilişkisel VTYS'nin zayıf bir seçenek olduğu kabul edilmektedir.

Verinin çeşitliliği: İlişkisel VTYS veriyi bağıntılar kümesi olarak modeller. Her bağıntı bir veri grubu kümesi iken, her veri grubu ise kısıtlı alanlardan gelen sonlu bir listedir. Açık olarak, verinin bu şekilde modellenmesi, çizgesel veri, ağaç yapılı veri ve birden çok değerli

öznitelikler gibi karmaşık veri yapılarını yönetmek için etkili şekilde kullanılamaz.

NoSQL VTYS'nin veri modelleri ise daha çeşitlidir. Anahtar-değer depoları (Oracle NoSQL, Terracota gibi), belge depoları (MongoDB gibi) ve geniş sütun depoları (BigTable, Cassandra, HBase gibi) NoSQL VTYS'nin yeni veri modelleri arasında yer almaktadır. Bunların arasında, özellikle belge depoları yapılanmamış veya yarı-yapılanmış veri yapılarını desteklemektedir.

5. TOPLANABİLECEK VERİ ÇEŞİTLERİ (TYPES OF DATA THAT COULD BE COLLECTED)

Öğrencilerin mobil cihazları tarafından toplanabilecek çeşitli veri tipleri bulunmaktadır. Bu veri tipleri aşağıda ana hatlarıyla anlatılmıştır.

Belirli tipteki veriler, kaydı oluşturan öznenin kimliğini açık edebilir. Bu durum veri kaydının özgün bir belirleyici içermesinden kaynaklanır. Örneğin, bir e-posta trafik izlemesi, söz konusu konuşmada yer alan bütün öznelerin e-posta adreslerini listeleyebilir. Bir e-posta adresi tek bir kişiye aittir ve bu yüzden özgün olarak o kişiyi belirler.

Özgün bir belirleyicinin olmadığı durumlarda dahi, bazı kısmi-belirleyici verilerin (yarı-belirleyici) birlikte kullanılmasıyla bir kullanıcı belirlenebilir. Örneğin, posta kodu, doğum tarihi, ve cinsiyet kullanılarak ABD vatandaşlarının %87'si ayırt edilebilmektedir [11].

Listelenen her veri tipi için, olası özgün ve yarı-belirleyici kısımları da ayrıca detaylandırılacaktır.

5.1. Öğrenci Veri Kaydı (Student Data Record)

Öğrenci veri kaydı, okul veya Milli Eğitim Bakanlığı gibi eğitimle alakalı kurumlar tarafından öğrenci ile ilgili olarak saklanan bilgidir.

Özgün belirleyiciler, öğrencinin kimlik numarası (Türkiye için, TC Kimlik Numarası) veya açık olarak listelenmiş aile üyeleri, öğrenci numarası ve telefon numaraları gibi belirleyicileri içermektedir. Yarı-belirleyiciler etki alanına bağlı olmakla birlikte, çoğu zaman doğum tarihi, cinsiyet, ve posta kodunu içerir.

5.2. Öğrenci Hakkındaki Öğretmen Görüşleri (Teacher Opinions About the Student)

Öğretmenlerin öğrencileri hakkındaki düşüncelerinin veya öğrenciye yönelik olarak yapılan planlarının metin halinde sistemde depolanması mümkündür.

5.3. Kamera ve Mikrofon (Camera and Microphone)

FATİH Projesi kapsamına alınmış okullardaki öğrencilere tablet bilgisayarlar verilmiştir. Bu cihazlar üzerindeki donanımlar, özel aygıt yazılımları veya işletim sistemi dahil olmak üzere yazılım eklentileriyle, uzaktan kontrol edilebilmektedir. Bu kontrol, kamera ve mikrofon gibi donanımları da kapsamaktadır, böylece öğrencilerin video konferans yoluyla öğretmenlere soru sorabildikleri, 7/24 sağlanabilen bir eğitim desteği düşünülebilir.

Donanım üzerinde bu derece bir kontrol sağlamak, kişisel mahremiyete yönelik önemli bir tehdidi de beraberinde getirebilmektedir: bir saldırgan uzaktan kamera veya mikrofonu çalıştırarak, öğrenci ve sosyal ortamıyla ilgili görsel veya işitsel kayıt elde etmeye çalışabilir.

Bütün görsel ve işitsel kayıtlar, hassas kabul edilmelidir. Saldırganın teşhis yeteneklerine bağlı olarak (yüz fotoğrafı, ses gibi), bu tür veriler belirleyici olabilir.

5.4. İnternet Geçmişi (Web History)

Bir kişinin internet geçmişi, ziyaret edilen internet siteleri ve internet üzerinde yapılan aramalardan oluşmaktadır. Bu veriler okul içerisinden erişim sağlandığı durumlarda kolaylıkla toplanabilir, çünkü FATİH sistemi okullarda gerekli internet erişimini sağlayacak ve bu sayede internet trafiğini izlemek yeterli olacaktır. Erişilen web sayfaları belirlenebileceği gibi üretilen trafik şifreli olmadığı sürece paketlerin içeriğini okumak da mümkündür.

Okul dışında erişim durumlarında, internet geçmişi çerezler (dosya sisteminde yer alan basit dosyalar) aracılığıyla veya tablet bilgisayara (belki tarayıcı üzerine) bir raporlama modülü eklenerek toplanabilir.

İnternet geçmişi verisi genellikle hassas veri kategorisindedir: nadir rastlanan ve ölümcül bir hastalıkla ilgili semptomların aratılması gibi. İnternet geçmişi aynı zamanda yarı-belirleyici özelliği taşır: arama dizelerinden bazı örnekler ve ziyaret edilen sayfalar bir kullanıcıya özgün olabilir ve bu örneklere sahip herhangi bir kişi kullanıcının kimliğini tespit edebilir. Bu durumun bir benzeri Amerikan AOL firması tarafından kullanıcıların arama geçmişlerinin paylaşılması ile ortaya çıkan gizlilik ihlalidir [12].

5.5. Çokluortam Erişimi (Multimedia Access)

Çokluortam erişimi verisi toplama yöntemleri, internet geçmişi verisi toplama yöntemlerine benzemektedir. Çokluortam kullanıcı tarafından okunan kitap ve gazeteleri; kullanıcının izlediği filmleri ve de kullanıcının dinlediği müziği içermektedir. Kullanıcı ile ilgili 6. Bölümde bahsedilen bilgileri yansıtabileceğinden dolayı bütün bu veriler hassas verilerdir.

5.6. Konum Verisi (Location Data)

Tablet bilgisayarlar çeşitli kanallardan konum verisinin toplanmasını sağlayabilirler: eğer KKS (küresel konumlandırma sistemi) teçhizatı varsa donanım yoluyla, eğer mobil internet bağlantısı (GPRS, 3G, 4.5G gibi) varsa hizmet sağlayıcı üzerinden konum bilgisi edinilebilmektedir. Bunlara alternatif diğer bir kanal ise kullanıcılar tarafından konum bildirmekte kullanılan çevrimiçi uygulamalardır (ör. FourSquare, Facebook gibi).

Konum verisi belirleyicidir. Çoğu kişi zamanının ciddi bir kısmını (özellikle akşam saatlerinde) evinde geçirmektedir. Bu bilginin açık adres rehberleriyle birlikte değerlendirmesi, kişinin olası kimliğini bir evin sakinlerine kadar indirgeyebilir.

Konum verisi aynı zamanda hassastır, çünkü bir kişi tarafından ziyaret edilen mekanlar politik ve dini görüş, ekonomik durum (sağlık, ruh hali, bağış yapma/alma gibi) ve cinsel yönelimler hakkında belirtici olabilir. Çoklu yörüngeler, bir etkinlikte birlikte yer alan insanlardan oluşan faaliyet gruplarını açığa vurabilir.

5.7. Çevrimiçi Sosyal Ağ Verileri (Online Social Network Data)

Bir öğrencinin çevrimiçi sosyal ağ verilerinin içinde arkadaşlık bağlantıları, özel mesajlaşmalar, çokluortam paylaşımları, beğenme ve favorileme gibi sosyal seçimler bulunmaktadır. Çevrimiçi sosyal ağ verileri, tablet bilgisayarın internet trafiğinin takip edilmesiyle elde edilebilir.

Bir kullanıcının çevrimiçi profili kesinlikle hassastır. Çevrimiçi ağlar üzerindeki paylaşımlar, konum, sosyal bağlantılar ile doğum tarihi, politik görüşler gibi yarı-belirleyiciler de dahil olmak üzere pek çok mahrem veriyi açık edebilir. Ayrıca profil bilgileri pek çok kullanıcı için belirleyici özelliktedir, çünkü bu bilgiler isim (takma isim yerine), kullanıcının arkadaşlarının isimleri ve etiketlenmiş fotoğraflarla, mesajları açık eder.

5.8. Kişi Listesi/Rehber (Contact List)

Kullanıcının kişi listesindeki herkes potansiyel bir sosyal bağlantı olduğundan, kişi listesi de çevrimiçi sosyal ağ verisine benzemektedir. Bu veri, işletim sistemi üzerinde yapılacak bir değişiklikle kolaylıkla toplanabilir. Çoğu mobil uygulama kullanıcının kişi listesine, izin almak kaydı ile, erişim sağlayabilmektedir.

Kişi listesi hassas ve belirleyici bilgidir. Kişi listesi; isimler, telefon numaraları, e-posta adresleri, ve doğum günü gibi yarı-belirleyiciler içerdiğinden belirleyici olma riski çok daha yüksektir.

5.9. Uygulama Verisi (Application Data)

Eğer öğrencilere, tablet bilgisayarlarına uygulama yükleme izni verilirse, bu uygulamalara ait veriler işletim sistemi veya ağ hizmetleri kullanılarak toplanabilir.

Bu tarz uygulama verileri, uygulamanın özelliklerine bağlı olarak hassas ve belirleyici olabilir. Örneğin, mesajlaşma ve/veya e-posta trafiğine olanak sağlayan her uygulama, kişi listesi özelliği taşımaktadır. Kişi listelerinin mahremiyet konusundaki olası tehlikelerinden yukarıda bahsedilmiştir.

5.10. Diğer Dijital İzler (Other Digital Traces)

Yukarıda açıklanandan çok daha fazla dijital iz bulunmaktadır. Örneğin, tablet bilgisayarın kullanım verisi, cihazın etkin olup olmadığını tespit edebilecek bir dijital izdir. Bu, kullanıcı mahremiyetine karşı kullanılabilecek bir yan kanaldır: bir öğrencinin tableti üzerinde çevrimiçi bankacılık uygulamalarının tespit edilmesi, tabletin ebeveynler tarafından da kullanılmakta olduğunu ortaya koyacaktır.

6. ÇIKARIMI YAPILABİLECEK HASSAS VERİLER (SENSITIVE DATA THAT COULD BE INFERRED)

Hassas kişisel veri, kişiye karşı ayrım yapmak amacıyla kullanılabilecek bilgiler içeren kişisel veri anlamına gelmektedir.

FATİH Projesi kapsamında bulunabilecek hassas kişisel verilerin sınıflandırması aşağıda yer almaktadır. Bu sınıflandırma bazı iyi bilinen hassas kişisel veri türlerini (sendika üyeliği gibi) kapsam dışı bırakıp, FATİH Projesi'ne özgü sınıfları içermektedir.

6.1. Irksal Konular (Racial Issues)

Irksal konularla ilgili hassas veri, öğrencinin ırksal ya da etnik kökenini açık edecek her türlü bilgiyi kapsamaktadır. Aşağıdaki senaryolarda bu bilginin çıkarımı yapılabilir.

- Bir öznenin kimlik numarası (TC Kimlik Numarası gibi), ebeveyn bilgileri, doğum yeri, veya nüfusa kayıtlı olunan yer gibi öznenin kimliğini içeren verileri paylaşmak.
- Takip edilen gazete köşeleri, okunan kitaplar, izlenen filmler - özellikle anadilde olmayan dijital içerik gibi çokluortam erişimleri.
- Menşe yeri, akrabaların menşe yerleri, belirli olaylar / sayfalar / kişiler ile ilgili beğenme / beğenmeme bilgilerinin açık olarak paylaşılması gibi çevrimiçi sosyal medya faaliyetleri.
- Ziyaret edilen kültürel yerler, memlekette geçirilen tatiller gibi konum verilerinin paylaşılması.

6.2. Politik Görüşler (Political Views)

FATİH Projesi'nin veri öznelerinin yaş grubu hesaba katılırsa, çoğu öğrencinin henüz herhangi bir politik görüşe sahip olmaması muhtemeldir. Ancak bu öğrenciler, yine de ebeveynlerinin politik görüşlerinden dolayı ayrımcılığa maruz kalabilirler. Böyle bir bilgiye, aşağıdaki senaryolarda ulaşılabilir.

- Mobil cihazlardan erişilen çokluortam ögeleri (gazete, kitap, film gibi).
- Paylaşılan içerik, belirli olaylar / sayfalar / kişiler ile ilgili beğenme / beğenmeme gibi çevrimiçi sosyal medya faaliyetleri.
- Kullanıcı açık faaliyetlerde bulunmasa dahi, kullanıcıya ait her türlü etiketin (siyasi görüş, inanç ve cinsel yönelim) arkadaş, aile, akrabalar gibi sosyal medya komşularının etiketlerinden çıkarılabileceği gösterilmiştir [13]. Bu yüzden, sosyal medya, kişi listesi, e-posta ve konum verisi gibi her türlü sosyal ilişkiyi açık eden her türlü bilgi, politik görüşlerin çıkarımı için kullanılabilir.
- Ziyaret edilen bazı internet siteleri ve benzer şekilde internet günlükleri, forumlar, takip edilen veya oluşturulan Wikipedia sayfaları politik görüşleri açık edebilir.

6.3. Dini inançlar ve Diğer İnanışlar (Religious and Other Beliefs)

Bazı durumlarda, ırksal köken ve dini inanç birbiriyle yakından ilişkilidir. Bu nedenle, böyle ikili gruplar için, ırk verisini çıkarmada kullanılabilecek her türlü kanal, aynı şekilde dini inançlar için de geçerlidir.

6.4. Sağlık Konuları (Health Issues)

Kişinin sağlık bilgileri hem fiziksel hem de ruhsal sağlık durumunu içerir. Geçmişinde belirli sağlık problemlerinin (ruhsal problemler, hayati tehlike oluşturan hastalıklar, cinsel yolla bulaşan hastalıklar ve hamilelik gibi diğer sağlıkla ilgili durumlar) olması, kişinin sosyal ayrımcılığa maruz kalmasına sebep olabilir. Bu nedenle, sağlık bilgileri hassas kabul edilmekte ve pek çok ülkede yasalar tarafından korunmaktadır (ör., A.B.D.'de yürürlükte olan sağlık verisi koruma yönergesi HIPAA [14]).

Sağlık bilgilerinin çıkarılabileceği olası senaryolar aşağıda listelenmiştir.

- Konum verisi ziyaret edilen sağlık kuruluşlarını (hastaneler, klinikler vs.) açığa çıkarabilir. Özellikle belirli hastalıklar (kanser ya da hamilelik gibi) üzerinde yoğunlaşmış kuruluşlarla ilgili edinilen bilgiler, kişinin mahremiyetini tehdit edebilir.
- Çeşitli sağlık durumları ile ilgili erişilen her türlü bilgi, kullanıcının o durumla bağlantısı olduğunu ima edebilir. Sağlıkla ilgili indirilen çokluortam verileri ve ziyaret edilen internet sayfaları, kullanıcının sağlık sorunlarını açık edebilir.

6.5. Cinsel Yönelimler (Sexual Preferences)

Gençler büyüdükçe ve cinselliğin farkına varmaya başladıkça, 'onaylanmayan' ve 'geleneksel olmayan' cinsel yönelimleri olduğu takdirde, özellikle kendi yaşıtları tarafından sosyal ayrımcılığa uğrama tehlikesiyle karşı karşıya kalabilirler. Bu tehlikeden dolayı, kişinin cinsel yönelimleri hassas kabul edilir. Cinsel yönelimleri açık edebilecek bilgi kaynakları aşağıda listelenmiştir.

- Bazı cinsel yönelimler, kimi aileler tarafından sağlık sorunu olarak görülür. Bu durumlarda, sağlık durumuyla ilgili her türlü bilgi cinsel yönelim çıkarımını yapmak için de kullanılabilir.
- Yukarıda belirtildiği gibi, kullanıcının sosyal komşuları (arkadaşlar vs.) eğer kendileri bir şekilde etiketlenmişlerse veya kendileri ilgili bir açıklama yapmışlarsa, kullanıcı ile ilgili gizli etiketlerin ortaya çıkmasına da neden olabilirler. Sosyal komşulara yönelik bilgiler (kişi listesi, e-postalar, konum verileri ve sosyal medya) bu bağlamda mahremiyete yönelik bir tehdit olarak kabul edilebilir. A.B.D.'de MIT tarafından yapılan GayDar adlı proje bu durumun bir örneğidir [15].

6.6. İşlenmiş veya İddia Edilmiş Suçlar / Disiplin Suçları (Committed or Alleged Crimes / Disciplinary Acts)

Kullanıcının *"işlediği ya da işlediği iddia edilen suçlar, veya işlediği ya da işlediği iddia edilen suçlarla ilgili devam eden yargılamalar, bu yargılamalardan beraat etmiş veya ceza almış olması"* [5] ile ilgili her türlü bilgi hassas kabul edilmektedir. Bu bilgileri açık etmek sosyal ve profesyonel bağlamda ayrımcılığa sebep olabilir. Bu nedenle kullanıcının izni olmadan açıklanmamalıdır. Bu tarz geçmişi olan öğrenciler çoğunlukla profesyonellerden yardım almaktadır. Dolayısıyla, öğrencinin bu profesyonellerle herhangi bir iletişiminin olduğunun açığa çıkması, öğrencinin geçmişinin de açığa çıkmasına neden olur.

6.7. Öğrenci ile İlgili Görüşler / Değerlendirmeler / Yorumlar (Opinions / Evaluations / Comments on a Student)

Öğretmenler ve okul yönetimleri öğrenci ile ilgili görüş, değerlendirme ve yorumları, öğrencinin veri kaydının bir parçası olarak saklayabilirler. Bu veriler kesinlikle hassas olma niteliği taşımaktadır.

7. VERİNİN ÜÇÜNCÜ PARTİLERLE PAYLAŞILMASI (SHARING DATA WITH THIRD PARTIES)

FATİH sisteminin verisi üzerindeki bazı analizler üçüncü partiler (dış aktörler) tarafından yürütülebilir. Dışarıdan,

üçüncü partiler tarafından yapılacak bu tarz erişimler, bir çeşit veri paylaşımıdır ve öznelerle ilgili hassas verilerin uygun şekilde korunmasını gerektirir. Hassas verilerin analiz amacıyla paylaşılması sorunu, genel olarak "mahremiyeti koruyan veri analitiği" şeklinde tanımlanmaktadır.

Koruma mekanizması, üçüncü partinin veriyi nasıl (hangi amaçla) kullanacağına bağlıdır. FATİH veri kümesinin boyutu ve çeşitliliği / zenginliği göz önüne alındığında çok çeşitli uygulamalar mümkündür. Veri kümesi, araştırma (pedagoji / öğretim teknikleri ve alıştırmaları), öğrenciler için önerilerde bulunma (seçmeli dersler, bölüm seçimi), sistemin genel kalitesini yükseltme (veritabanına ayar yapma, ağ yönlendirmesi vs.) gibi amaçlarla kullanılabilir.

Üçüncü partilerin FATİH verisine erişimi için iki model bulunmaktadır: (1) Veri kümesinin isimsizleştirilmiş hali paylaşılabilir, (2) Korumalı bir veritabanı erişimi sağlanabilir. İki çözümün de getirileri ve götürüleri aşağıda anlatılmıştır.

7.1. Anonimleştirme Yoluyla Koruma (Protection through Anonymization)

Anonimleştirmenin amacı, bir veri kaydı ile ait olduğu veri öznesinin ilişkisini ortadan kaldırmaktır. En basit haliyle, eğer veri kaydının sahibi ile ilgili bir çıkarım yapmak mümkün değil veya sınırlandırılmış ise, o kaydın anonim halde paylaşılmasının bir zararı yoktur.

Anonimlik çeşitli şekillerde tanımlanabilir. Bu tanımlar, saldırı çeşitlerinin, saldırganın ne kadar bilgiyi ele geçirmesine izin verdiğine göre değişiklik göstermektedir.

***k*-anonimlik [9]:** Bir kişinin yarı-belirleyici alanları, en az (k-1) farklı kişiden ayırt edilemez olmalıdır. Dolayısıyla kişinin paylaşılan herhangi bir kayıta eşleştirilme ihtimali azaltılmış olur.

***l*-çeşitlilik [16]:** Aynı (anonimleştirilmiş) yarı-belirleyici alanları olan bütün kayıtlar için en az *l* farklı, iyi açıklanmış hassas özellik değeri olmalıdır. Böylece kişinin hassas herhangi bir etiket ile ilişkilendirilmesi ihtimali sınırlandırılmış olur.

***t*-yakınlık [17]:** Aynı yarı-belirleyici değerlerine sahip her (anonim) kayıt grubu için, hassas değerlerin dağılımı, asıl veri kümesi üzerindeki dağılımdan çok farklı olmamalıdır.

Anonimleştirmenin temel faydası, ortaya konan verinin doğruluğudur. Anonimleştirme sırasında yarı-belirleyici alanlar genellenmektedir. Mersin doğumlu bir öğrenci kaydının Akdeniz'de doğan öğrenci olarak değiştirilmesi genellemeye örnektir. Böylece daha fazla kişi aynı şekilde gösterildiğinden, saldırganın bir kişinin kaydından kimlik çıkarımı yapması çok daha zor hale gelmektedir. Genelleme bir kayıt için olası değer aralığını genişletmekle birlikte, bu belirsizlik haricinde hiç bir yanlış bilgi vermemektedir.

Bir diğer fayda ise, anonimleşmiş veri üzerinde her türlü analizi yapmanın mümkün olmasıdır. Anonim veri, tablo halinde sunulmaktadır ve üzerinde her türlü analiz işlemini kolaylıkla yapmak mümkündür.

7.2. Diferansiyel Mahremiyet Yoluyla Koruma (Protection through Differential Privacy)

Diferansiyel mahremiyet [18], anonimleştirmeye alternatif yeni bir koruma yöntemidir. Bu modelde, mikro-veriler (kayıtlar) veritabanı sunucularından çıkmaz. Sadece istatistiksel sorguların yapılmasına izin verilir ve istatistiksel sorgular rastgele gürültü ile karıştırılır. Bu kısıtlı erişim arayüzü haricinde veriyi analiz etmek mümkün değildir.

Diferansiyel mahremiyetin kişiye sunduğu garanti paylaşılan her türlü bilginin, kişi veritabanında yer almamayı tercih etmiş olsa bile oluşturulabilmesinin benzer ihtimalde olmasıdır. Bu garantiden ötürü eklenen gürültü miktarı, istatistiksel sorguların ne kadar detaylı olduğuna göre değişmektedir. Bu kavram *sorgu betiğinin hassasiyeti* olarak tanımlanmıştır.

Saldırı gücüne göre çeşitli mahremiyet tanımlarının mümkün olduğu anonimleştirmenin aksine, diferansiyel mahremiyet düzeneği tüm geçmiş bilgilerine ulaşılabilecek bütün kanalları göz önüne alır. Sonuç olarak, bu model ile çok daha sağlam bir mahremiyet koruması sağlanır. Bu durum, diferansiyel mahremiyetin temel faydasıdır.

Bu çok önemli faydaya rağmen, diferansiyel mahremiyetin getirdiği bazı zorluklar da vardır. Pek çok sebepten dolayı diferansiyel mahremiyetin uygulanması çok kolay değildir. İlk olarak, sorgu betiğinin hassasiyetinin hesaplanması NP-zordur [19]. İkinci olarak, analizlerin istatistiksel veritabanı sorguları şeklinde ifade edilmeleri gerekmektedir. Üçüncü olarak da, eklenen gürültüden dolayı elde edilen sonuçlar doğruluktan uzaklaşmaktadır. Son olarak arayüz üzerinden sistemi kullanan analistlerin birbirleri ile sonuç paylaşımı yapabilecekleri riski arzu edilen mahremiyet standardını sağlamayı zorlaştırmaktadır.

8. ÖNERİLER (SUGGESTIONS)

Yapılan bu incelemeler ışığında, FATİH Projesi'nin tasarımında ve uygulamasında dikkate alınması dileği ile, Milli Eğitim Bakanlığı tarafından değerlendirilmek üzere aşağıda özetlenen öneriler derlenmiştir.

- Milli Eğitim Bakanlığı tarafından FATİH projesine özel yönerge ve yönetmelikler oluşturulmalıdır. Veri toplama ve işleme süreçleri için uygunsuz davranışlar

net bir şekilde belirlenmeli ve bu davranışların nasıl cezalandırılacağı tayin edilmelidir.

- Veri koruması ile ilgili gereklilikler "tasarımda mahremiyet" prensibi çerçevesinde henüz sistem mimarisi ile ilgili temel kararlar verilirken dahi sürecin bir parçası olmalıdır (bknz. Bölüm 4). Mahremiyet koruması bir eklenti şeklinde sonradan projeye dahil edilmemelidir.
- Kişisel bir verinin toplanması kararı verilirken, verinin yapılacak analizlerdeki yararlılığı ile birlikte, açığa çıkmasının doğuracağı mahremiyet riskleri de değerlendirilmelidir (bknz. Bölüm 6).
- Hangi verinin ne süreyle depolandığı şeffaflık ilkesi çerçevesinde kamuoyu ile paylaşılmalıdır.
- Toplanan veriye denetleyici ve işleyiciler tarafından erişim sıkı bir şekilde kontrol edilmeli, her erişim denetleme mekanizmaları tarafından kayıt altına alınmalıdır.
- Veri sızması vakaları yaşandığı takdirde durum ivedilikle kamuoyu ile paylaşılmalı, etkilenmesi beklenen bireylere gerekli uyarılar iletilmelidir.
- Her veri öznesine kendi verisini yönetme ve nasıl kullanılacağı konusunda sınırlar belirleme imkanı (mümkün olduğu nispette) tanınmalıdır.
- 3. şahıslarla doğrudan veri paylaşımına müsade edilmemeli, zorunlu hallerde ise, anonimleştirme veya diferansiyel mahremiyet gibi veri koruma yöntemlerine başvurulmalıdır (bknz. Bölüm 7).

FATİH projesinin kaydettiği aşama itibariyle bazı önerilerin hayata geçirilmesinde geç kalınmış olunabilir. Ancak proje kapsamında kişisel verisi toplanacak bireyler Türkiye'nin geleceğidir ve unutulmamalıdır ki, dışarı sızan verinin geri döndürülmesi mümkün olmamaktadır. Bu sebeple, tam veri koruması sağlamak adına, gerekli hallerde FATİH projesinin belirli bileşenlerinin yeniden tasarlanması dahi değerlendirilmelidir.

9. SONUÇ (CONCLUSION)

Mahremiyet, Türkiye Cumhuriyeti Anayasası tarafından da güvence altına alınmış olan temel bir insan hakkıdır. Yakın zamanda yürürlüğe giren Veri Koruma Yasası ile bu konudaki usül ve esaslar düzenlenmiştir. Öte yandan öğrenciler göz önüne alındığında veri koruması çok daha hassas bir konu haline gelir, çünkü gelişme aşamasında olan çocuklarımız korunmaya çok daha fazla muhtaçtır.

Türkiye'de gerçekleşen en önemli inisiyatiflerden olan FATİH projesi amaçlandığı gibi eğitimde fırsat eşitliğini sağlama potansiyeline sahiptir. Ancak bu kadar büyük ve önemli bir eğitim ve bilişim projesinin veri koruması açısından yaratabileceği sorunlar da göz ardı edilmemelidir. Bu makale ile FATİH projesi kapsamında bir öncü çalışma yapılıp bilişim dünyasında bu konunun daha derin irdelenmesi için bir zemin hazırlanması amaçlanmıştır. Bu yapılırken sistem mimarisi, veri tabanları, mobil cihazlar, sosyal ağlar, bulut bilişim gibi çok farklı alan bir arada değerlendirilmiş ve mahremiyeti koruyan veri paylaşımı ve analizi kapsamında kullanılması gereken standartlar ve teknolojilere değinilmiştir.

Mahremiyeti koruyan veri paylaşımı ve analizi teknikleri üzerinde uzun süredir çalışılmaktadır. Bu tekniklerin Türkiye'de devam eden büyük çaplı bilişim projelerine uyarlanması ve bu projeler kapsamında toplanması öngörülen büyük veriler üzerinde denenmesi gerekmektedir. Türkiye'de bu alanda önemli bir potansiyel olduğunu düşünülmekte ve bu çalışma ile araştırmacıların ilgisinin bu konuya çekilmesi ümit edilmektedir.

KAYNAKLAR (REFERENCES)

[1] Internet: Kişisel Verilerin Korunması Kanunu, Resmi Gazete, http://www.resmigazete.gov.tr/eskiler/2016/04/20160407-8.pdf, 15.05.2016

[2] Internet: Student Digital Privacy and Parental Rights Act of 2015, US Congress, https://www.congress.gov/bill/114th-congress/house-bill/2092, 15.05.2016.

[3] Internet: Eğitimde FATİH Projesi | Eğitim Teknolojileri Zirvesi 2016, T.C. Milli Eğitim Bakanlığı, http://fatihprojesietz.meb.gov.tr/, 15.11.2016.

[4] Internet: FATİH Projesi, T.C. Milli Eğitim Bakanlığı, http://fatihprojesi.meb.gov.tr/, 15.05.2016.

[5] Internet: Data protection principles, Information Comissioner's Office (ICO), UK https://ico.org.uk/for-organisations/guide-to-data-protection/key-definitions/, 15.05.2016.

[6] C. Romero, S. Ventura. "Educational data mining: a review of the state of the art." *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 40(6), 601-618, 2010.

[7] P. Ice, S. Diaz, K. Swan, M. Burgess, M. Sharkey, J. Sherrill, D. Huston, H. Okimoto, "The PAR Framework Proof of Concept: Initial Findings from a Multi-Institutional Analysis of Federated Postsecondary Data." *Journal of Asynchronous Learning Networks*, 16(3), 63-86, 2012.

[8] A. S. Tanenbaum, M. V. Steen, **Distributed Systems: Principles and Paradigms (2nd Edition)**, Pearson, A.B.D, 2006.

[9] Ferraiolo, D.F. and Kuhn, D.R.. "Role-Based Access Control", **15th National Computer Security Conference,** Maryland, A.B.D., 554–563, Kasım, 1992.

[10] W. Vogels. "Eventually consistent". *Communications of the ACM* 52(1), 40-44, 2009.

[11] L. Sweeney. "k-anonymity: a model for protecting privacy". *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10 (5), 557-570, 2002.

[12] Internet: M. Arrington, TechCrunch, AOL proudly releases massive amounts of user search data. http://techcrunch.com/2006/08/06/aol-proudly-releases-massive-amounts-of-user-search-data/, 15.05.2016.

[13] J. Lindamood, R. Heatherly, M. Kantarcioglu, B. Thuraisingham, "Inferring private information using social network data", **18th**

international conference on World wide web (WWW '09), Madrid, Spain, 1145-1146, Nisan, 2009.

[14] Internet: HHS, Health Information Privacy http://www.hhs.gov/hipaa/, 15.05.2016.

[15] C. Jernigan, B.F. Mistree, "Gaydar: Facebook friendships expose sexual orientation". *First Monday*, 14(10), 1-10, 2009.

[16] A. Machanavajjhala, D. Kifer, J. Gehrke, M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity". *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 1-12, 2007.

[17] N. Li, T. Li, S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity", **IEEE 23rd International Conference on Data Engineering, (ICDE'07)**, İstanbul, Türkiye, 106-115, Nisan, 2007.

[18] C. Dwork, "Differential privacy", **33rd International Colloquium on automata, languages and programming, (ICALP'06)**, Venedik, İtalya, 1-12, Temmuz, 2006.

[19] X. Xiao, T. Yufei, "Output perturbation with query relaxation" *Proceedings of the VLDB Endowment* 1(1): 857-869, 2008.