



Utrecht University



Veri işbirlikleri ve cep telefonu verisi ile göç ve mülteci çalışmaları

Albert Ali Salah^{1,2}, PhD

¹ Utrecht University, Department of Information and Computing Sciences

² Boğaziçi Üniversitesi, Bilgisayar Mühendisliği Bölümü

Göç çalışmaları ve büyük veri

Göç ve mülteci çalışmalarında klasik veri kaynakları sınırlıdır:

- Resmi istatistikler düzenli toplanmamış olabilir
- Düzensiz göç ile ilgili veri olmayabilir
- Ülkeden ülkeye “mülteci,” “göçmen” gibi temel tanımlar farklılık gösterebilir

Büyük veri kaynakları (ör. cep telefonu verisi, uydu görüntüleri, sosyal medya verisi) bazı durumlarda eksik bilgileri tamamlamak için kullanılabilir.

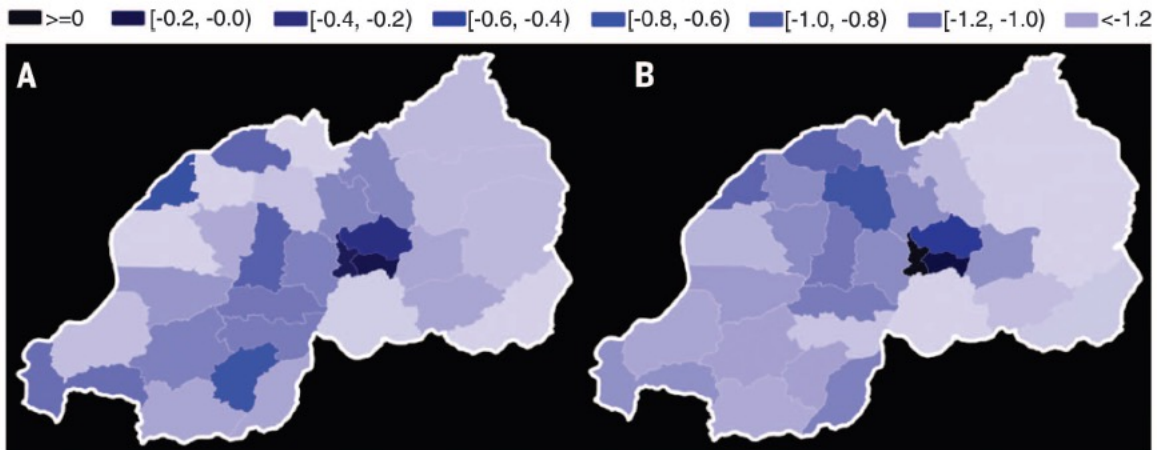
Özel sektör ve büyük veri

- Cep telefonu verisi,
- Uydu görüntüleri,
- Sosyal medya içeriği,
- Giyilebilir algılayıcılar, akıllı saatler,
- Altyapı kullanımı (ör. demiryolları),
- Emlak sitelerinin verileri,
- Çok oyunculu büyük bilgisayar oyunları,
- ... binlerce insanın kullandığı herhangi bir sistemde
biriken veri

Cep telefonu verisi ile refah düzeyi kestirimi

Table 1. Summary statistics for primary data sets. Phone survey data were collected by the authors in Kigali, in collaboration with the Kigali Institute of Science and Technology. Call detail records were collected by the primary mobile phone operator in Rwanda at the time of the phone survey. Demographic and Health Survey (DHS) data were collected by the Rwandan National Institute of Statistics. N/A, not applicable.

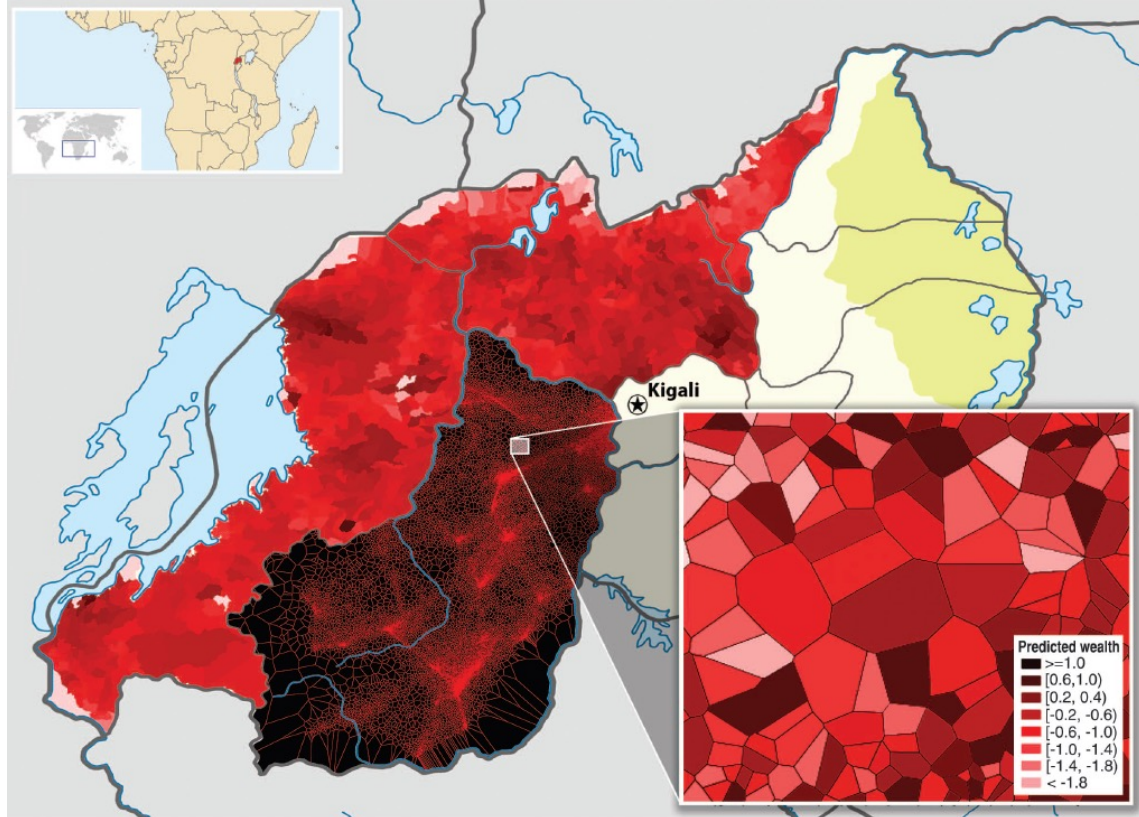
Summary statistic	Phone survey	Call detail records	DHS (2007)	DHS (2010)
Number of unique individuals	856	1.5 million	7377	12,792
Data collection period	July 2009	May 2008–May 2009	Dec. 2007–Apr. 2008	Sept. 2010–Mar. 2011
Number of questions in survey	75	N/A	1615	3396
Primary geographic units	30 districts	30 districts	30 districts	30 districts
Secondary geographic units	300 cell towers	300 cell towers	247 clusters	492 clusters



Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata.

Science, 350(6264), 1073-1076.

Cep telefonu verisi ile refah düzeyi kestirimi



Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata.

Science, 350(6264), 1073-1076.

Veri işbirlikleri

'Veri işbirlikleri özel sektör ve kamu arasında veri paylaşımı ve istatistik hesaplamaya dayanan işbirlikleridir.'

'Sektörler arası veri paylaşımının riskleri ve getirileri iyi anlaşılmazsa, özel sektör riskten kaçacak ve verinin ikincil kullanımı ile sosyal etkisi azalacaktır.'

Cep telefonu verisini nasıl kullanabiliriz?

01

Telefona kurulu bir uygulama ile (veri toplayan) başka bir sistemin etkileşimi ile

02

Doğrudan telefona kurulu bir uygulama ile

03

Operatörün cep telefonu konuşma ve veri transferi kayıtları (CDR/xDR) ile

Cep telefonu verisi

Veritabanı	Veri	Yıl	# kişi	Notlar
MIT Reality Mining	Nokia telefonlar: Bluetooth cihazlar, yer bilgisi, arama ve SMS logları	2004-2005	100	MIT öğrencileri ve hocaları
Friends and Family	Akıllı telefon: yer bilgisi, arama/SMS, baz istasyonu uzaklık, anketler	2010-2011	140	Mobile Territorial Lab
Mobile Data Challenge	Nokia telefonlar: Bluetooth cihazlar, yer bilgisi, arama, SMS, uygulama ve medya kullanımı, pil durumu, akustik veri	2009-2011	185	Gençler
LiveLab	iPhone verisi: arama, SMS, tarayıcı geçmişi, ivme ölçer, pil, ekran, uygulama kullanımı, baz istasyonu/kablosuz ağ ID	2010	25	Lise öğrencileri
Device Analyzer	Android telefonlar: uygulama, kablosuz ağ, pil ve arama kayıtları	2014	20,000	175+ ülkeden
PhoneLab	Android telefonlar: yer, pil, kablosuz ağ, baz istasyonu verisi	2015	199+ 288	
Sensible DTU	FB verisi, okul performansı, cep telefonu Android telefonlar: kablosuz ağ, arama, SMS, baz istasyonu	2012-2013	1000	TU Denmark öğrencileri
Copenhagen Networks Study	Android telefonlar: baz istasyonuna uzaklık, arama, SMS, FB arkadaş bağlantıları	4 hafta, 2019	700+	TU Denmark öğrencileri

Cep telefonu arama kayıtları (CDR)

Avantajları:

- Çok geniş kapsam
- Belli senaryolar için kullanıcılardan ayrıca izin almadan çalışılabilir
- Zamansal ve uzamsal detayı yüksek

Dezavantajları:

- Zor elde edilen bir veri
- Hassas veri
- Gerekli olduğu durumda bilgilendirilmiş onam alınması zor
- İzin alınmayacaksa anonim ve gruplanmış şekilde çalışılması gerekli

CDR (ses)

HR_ORG	TLFN_A	TLFN_B	CD_GEO_A	CD_GEO_B	DT_ORG	CD_SNTD	CD_ERB	CD_CCC	QT_DUR
20:05:31	XXX	YYY	3	11	20140519	2	1562	568	33
...

CDR (SMS)

HR_ORG	TLFN_A	TLFN_B	CD_GEO_A	CD_GEO_B	DT_ORG	CD_SNTD	QT_TRFG
15:53:54	XXX	ZZZ	3	25	20140506	2	1
...

Kullanım	Sosyal Ağ	Hareketlilik
Konuşma süresi	Düğüm dereceleri	Hareketlilik çapı
Konuşma sayısı	Günlük konuşulan kişi sayısı	Toplam hareket mesafesi
Konuşmalar arasındaki süre	Günlük yazışılan kişi sayısı	Antenlerin trafik miktarı
		

CDR/xDR verisi

Veritabanı	Veri	Yıl	# kişi	Notlar
D4D Ivory Coast	CDR	2011-2012 (1 year)	5 milyon	Antenler arası iletişim (saatlik), 50k anten seviyesinde hareketlilik; 500k ilçe seviyesinde hareketlilik; 5k iletişim ağı
D4D Senegal	CDR	2013 (1 year)	9+milyon	Antenler arası iletişim, 300k kullanıcı hareketlilik
Telecom Italia	CDR, elektrik kullanımı, hava durumu, yağmur miktarı, haberler, Tweet'ler	2014	300+k	Milano ve Trento
Telefonica	Adım sayısı, hastane yatışları, ulaşım, Twitter, suç verisi	2013 (3 hafta)	500+k	Londra
D4R Türk Telekom	CDR, "mülteci" işaretleri	2018	1 milyon	Türkiye'de Suriyeli mültecilerin hareketliliği
HummingBird Turkcell	xDR, CDR	2020-2021	-	Türkiye'den Avrupa'ya göç

Salah, A.A., "Mobile data challenges for human mobility analysis and humanitarian response," in M. McAuliffe (ed.) *Research Handbook of International Migration and Digital Technology*, Edward Elgar, 2021.

Data for Development Yarışması (D4D)

- Data for Development Challenge, France Telecom & Orange Cote d'Ivoire ve Fildişi Sahilleri Hükümeti tarafından organize edildi,
- 90 araştırma grubu katıldı,
- 267 başvuru geldi, 166 grup veritabanına erişti,
- **5 milyon kullanıcı, 16 aylık veri,**
- **Katılımcılar yarışma sonunda veriyi silmek zorundaydı,**
- İkinci D4D Yarışması Senegal ve Katar verisi ile yapıldı.

Data for Development Yarışması (D4D)

- Saat bazında antenler arası trafik,
- 50,000 müşteri için ikişer haftalık periyotlarla anten seviyesinde konuşma kaydı
- 500,000 müşteri için bütün veri toplama aralığında ilçe seviyesinde konuşma kaydı
- 5,000 müşteri için iletişim ağı

Kişisel bilginin ortadan kaldırılması

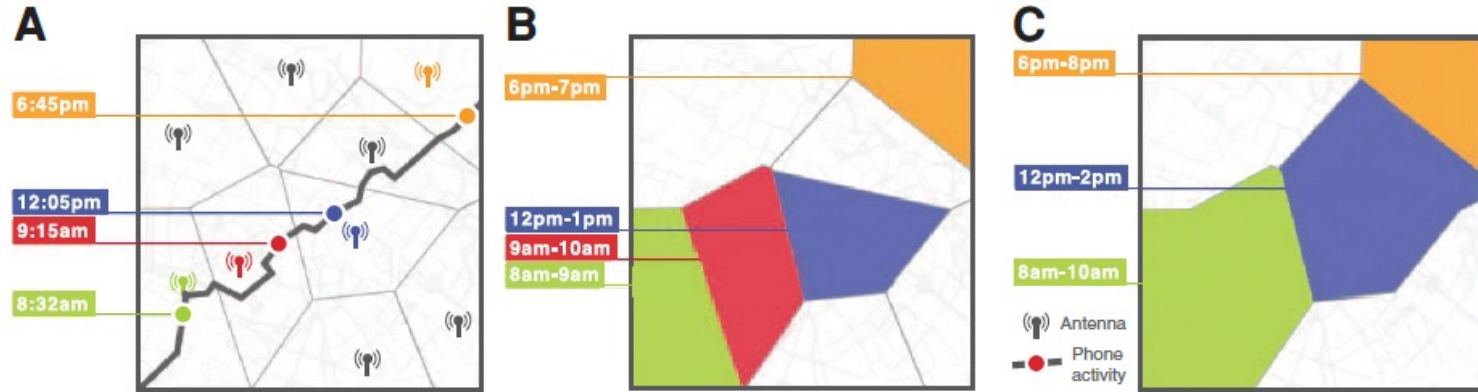


Figure 1 | (A) Trace of an anonymized mobile phone user during a day. The dots represent the times and locations where the user made or received a call. Every time the user has such an interaction, the closest antenna that routes the call is recorded. (B) The same user's trace as recorded in a mobility database. The Voronoi lattice, represented by the grey lines, are an approximation of the antennas reception areas, the most precise location information available to us. The user's interaction times are here recorded with a precision of one hour. (C) The same individual's trace when we lower the resolution of our dataset through spatial and temporal aggregation. Antennas are aggregated in clusters of size two and their associated regions are merged. The user's interaction are recorded with a precision of two hours. Such spatial and temporal aggregation render the 8:32 am and 9:15 am interactions indistinguishable.

Mülteciler için Büyük Veri (D4R) Yarışması

200K mülteci ve 800K diğer müşterilerden, bir yıllık veri.

1- Baz istasyonu (anten) pozisyonları

2- Antenler arası trafik, saat bazında.

Toplam konuşma sayısı ve dakika cinsinden toplam, “mültecilerden” ve “mülteci olmayanlardan” olarak iki ayrı grupta.

3- Detaylı hareketlilik verisi, küçük bir kullanıcı kümesi için, ikişer haftalık periyotlarda, anten seviyesinde.

4- Detaysız hareketlilik verisi, küçük bir kullanıcı kümesi için, bütün sene, ilçe seviyesinde.

D4R Konuları

Güvenlik

Şiddet, hırsızlık, insan kaçakçılığı

Sağlık

Hastanelere erişim, salgın hastalık ve aşılama modelleri

Eğitim

Okullara erişim, dil eğitimi

İşsizlik

İşsizliğe veya mevsimsel işlere bağlı hareketlilik, kaynak yönetimi

Sosyal entegrasyon

Olaylara ve kurumlara bağlı entegrasyon endikatörleri, ayrışma

Proje kořulları

- Veri korunumu ile ilgili detaylı anlaşma imzalanması
- Kurumda kalıcı kadroda bir arařtırmacının yöneticilięi
- Projeyi anlatan teknik bir dokümanla başvurulması
- Projenin mültecilere ne gibi katkıları olacaęının detaylandırılması

Temel etik konular

Onam

Veri sahiplerinin veri paylaşımına izin vermiş olması

Koruma

Bireylerin izlenemiyor olması

Saklama, arşivleme

Verinin sınırlı kullanımı

Enformasyon

'Önce zarar verme' prensibi

Düzeltilme ve itiraz

Kişisel bilgi tutulmaması

Sorumluluk ve hesap verilebilirlik

Açık yasal doküman

Tasarım aşamasında veri korunması

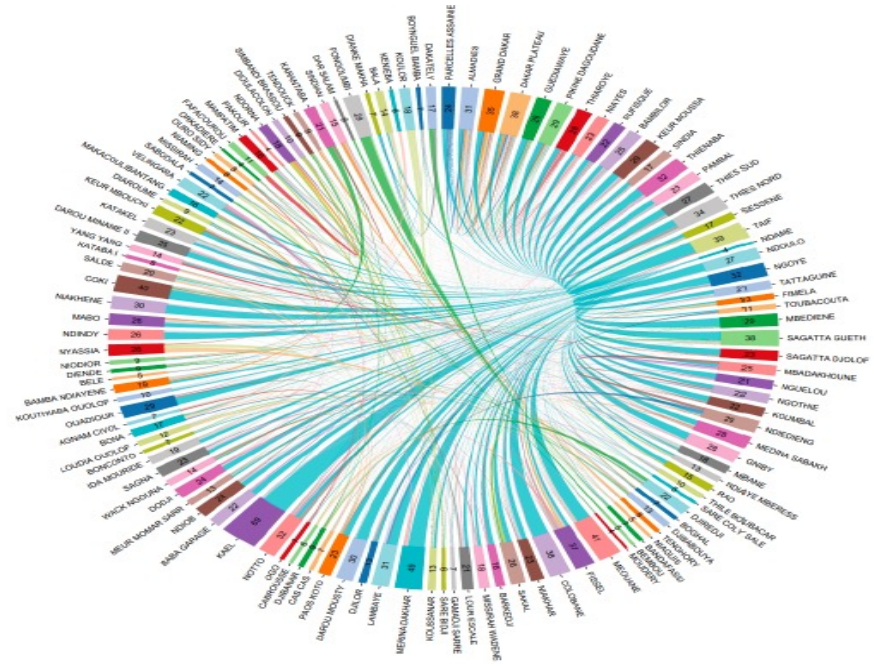
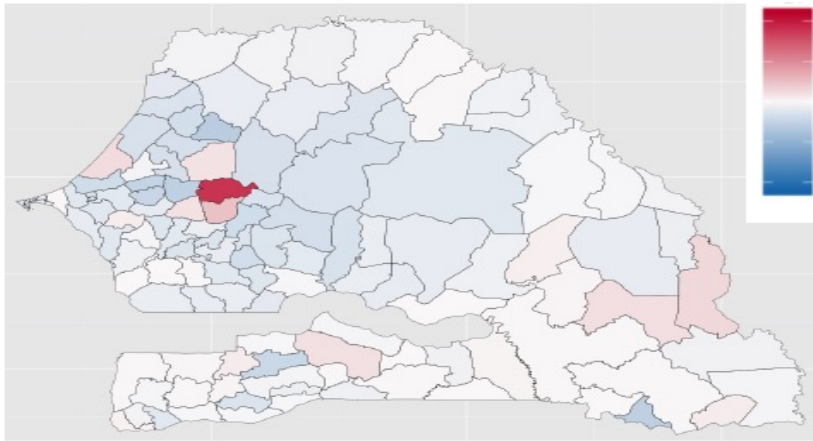
Verinin toplanması aşamasında anonimleştirme

Sınırlama

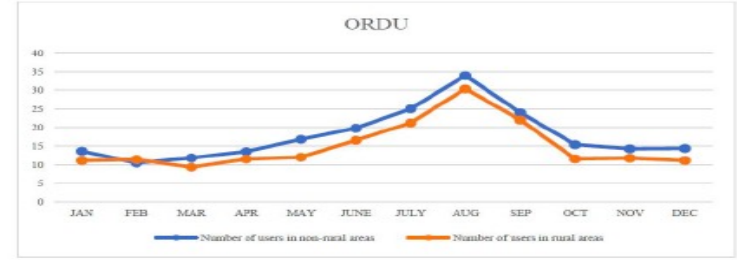
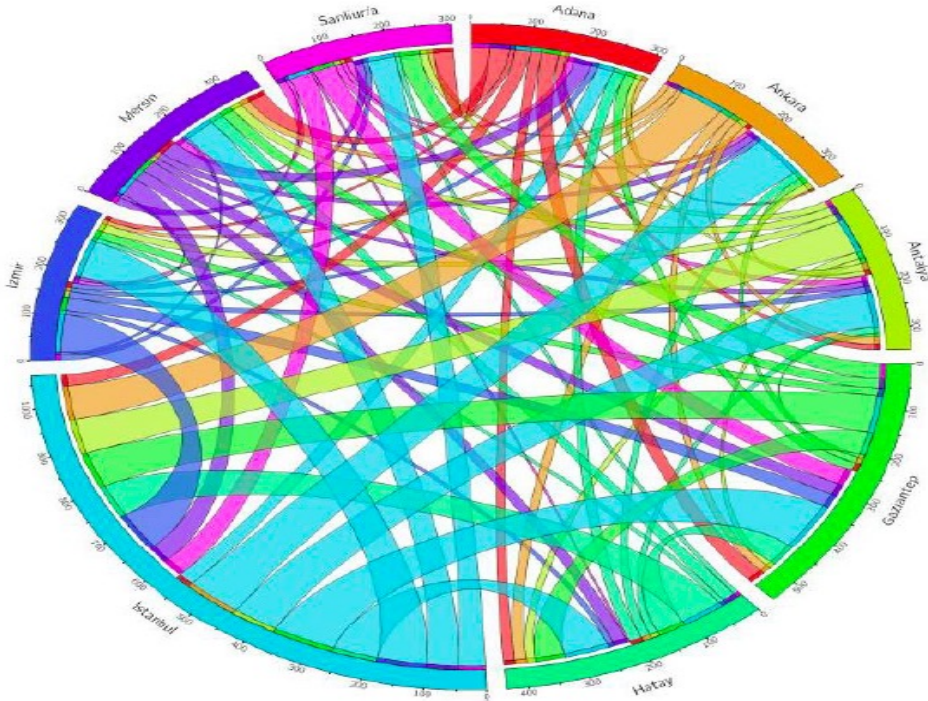
Tüm kurumdaşları kapsamayan, bireysel sözleşme

Vinck, P., Pham, P. N., & Salah, A. A. (2019). "Do No Harm" in the Age of Big Data: Data, Ethics, and the Refugees. In Guide to Mobile Data Analytics in Refugee Scenarios (pp. 87-99). Springer, Cham.

Hareketliliğin modellenmesi

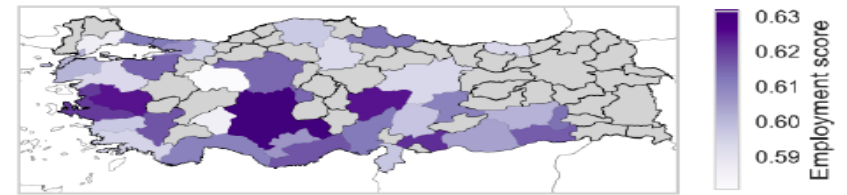
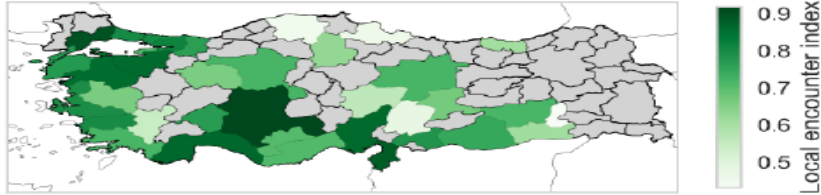
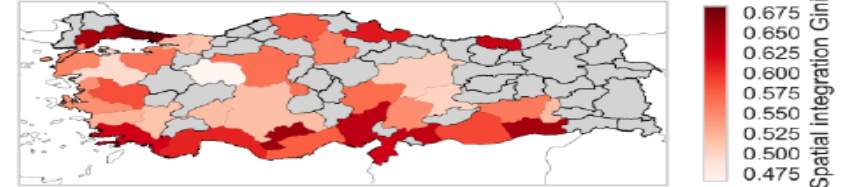
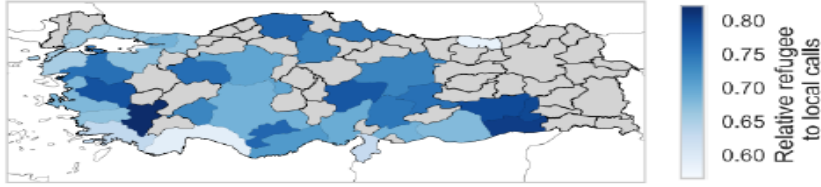


Mevsimsel işçilerin hareketliliği



Alısık ST, Aksel DB, Yantaç AE, Baruh L, Salman S, Kayı I, İçduygu A, Bensason I (2019). UDMIT: an urban deep map for integration in Turkey. In Data for Refugees Challenge Workshop

Sosyal entegrasyon endikatörleri



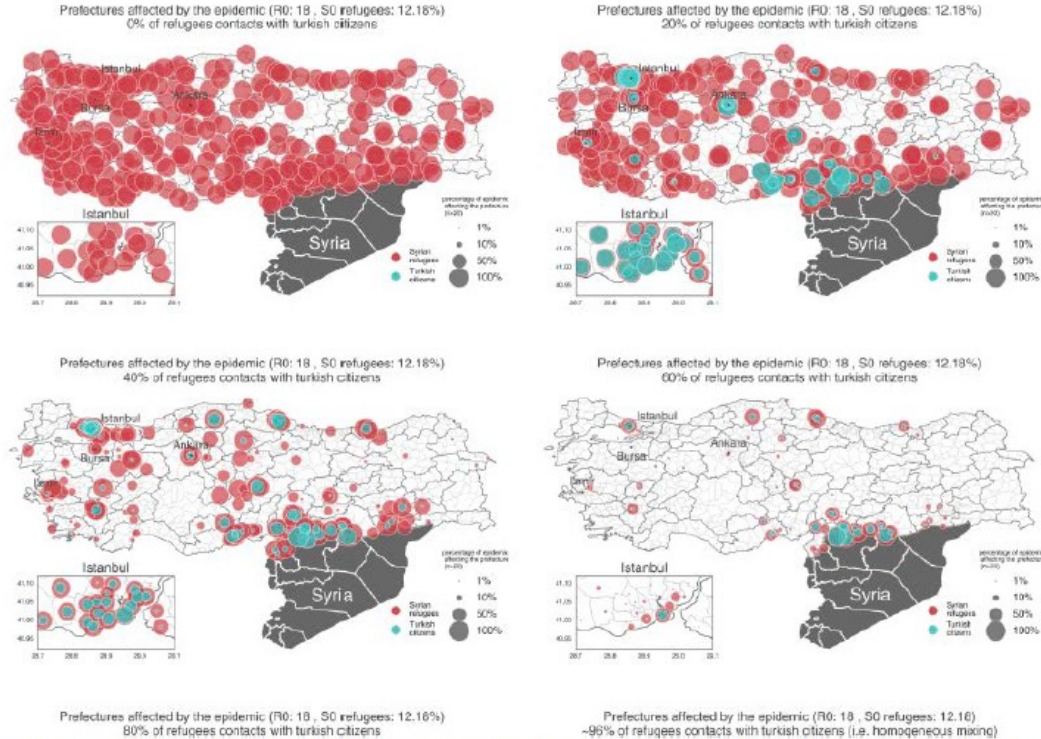
Bakker M, Piracha D, Lu P, Bejgo K, Bahrami M, Leng Y, Balsa-Barreiro J, Ricard J, Morales A, Singh V, Bozkaya B, Balcisoy S, Pentland A (2019) Measuring fine-grained multidimensional integration using mobile phone metadata: the case of Syrian refugees in Turkey. In: Data for Refugees Challenge Workshop

Sağlığa erişim



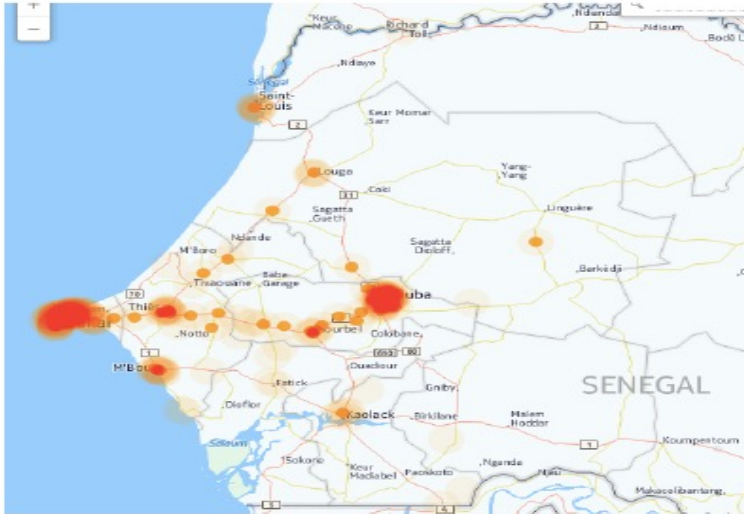
Fig. 2. Map with red pins showing the locations of current MHCs over the choropleth which illustrates the density of refugee residents at different districts in Istanbul.

Salgınların yayılımı

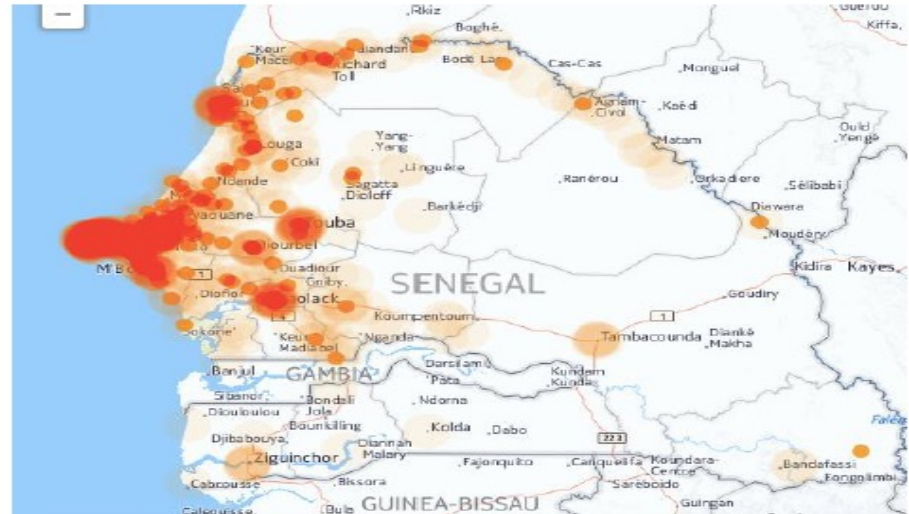


Bosetti, P., Poletti, P., Stella, M., Lepri, B., Merler, S., & De Domenico, M. (2019). Reducing measles risk in Turkey through social integration of Syrian refugees. In Data for Refugees Challenge Workshop.

Özel günlerde hareketlilik

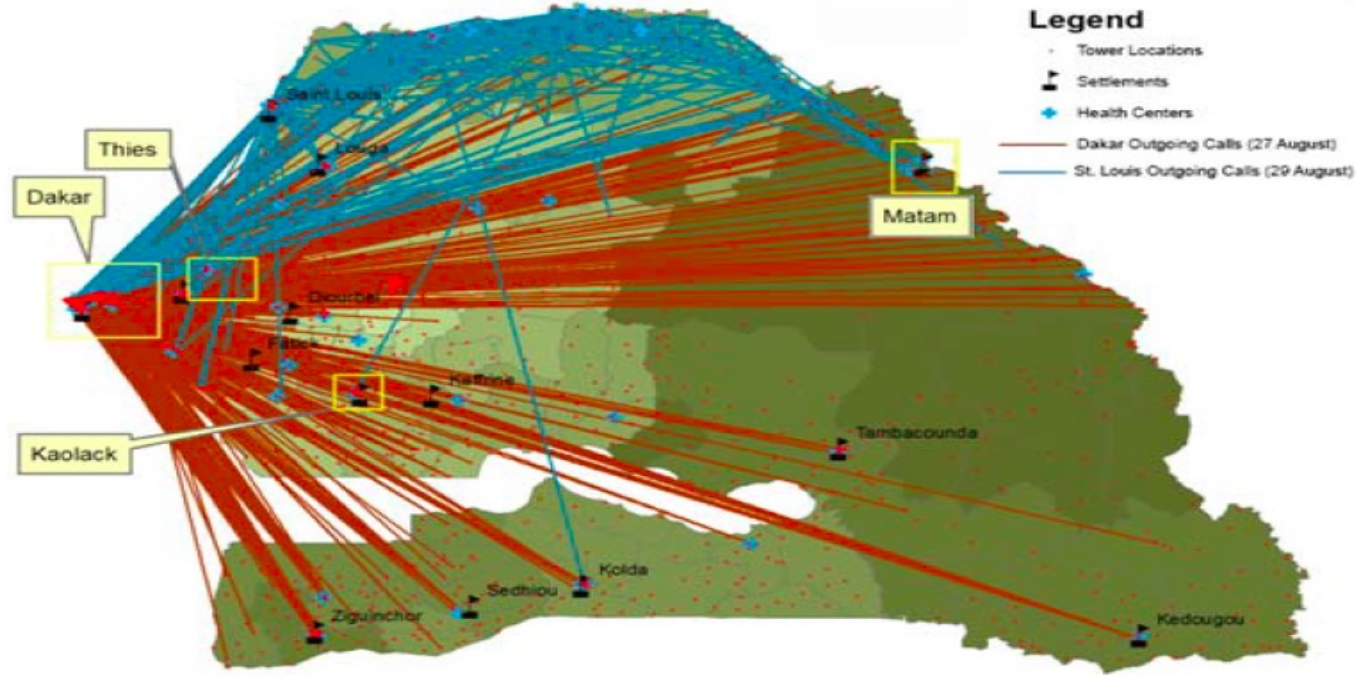


Human Mobility during Magal of Touba (12/22/2013)

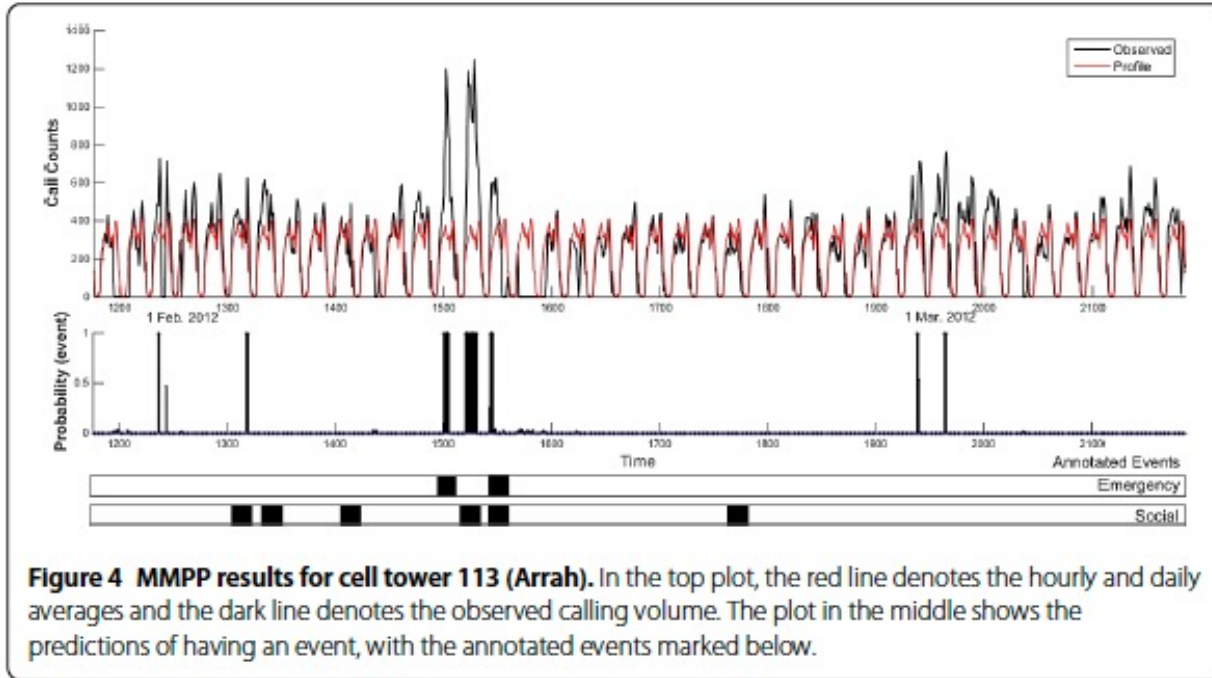


Human Mobility during Gamou of Tivaouane (01/23/2013)

Acil durumlar ve felaketlerde hareketlilik

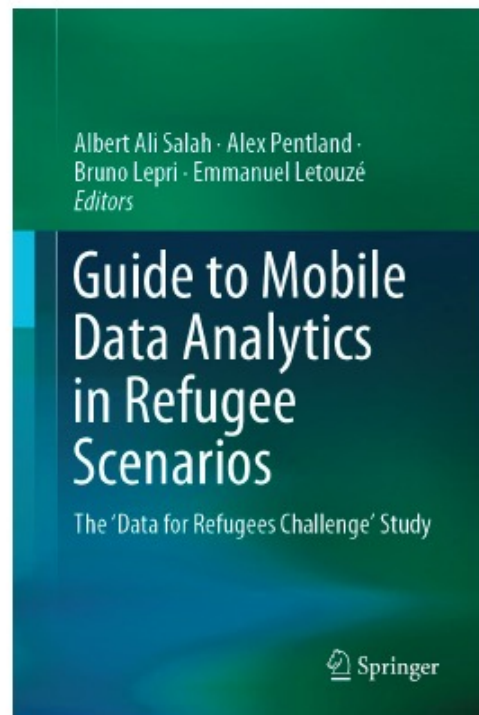


Anormal durumların tespiti

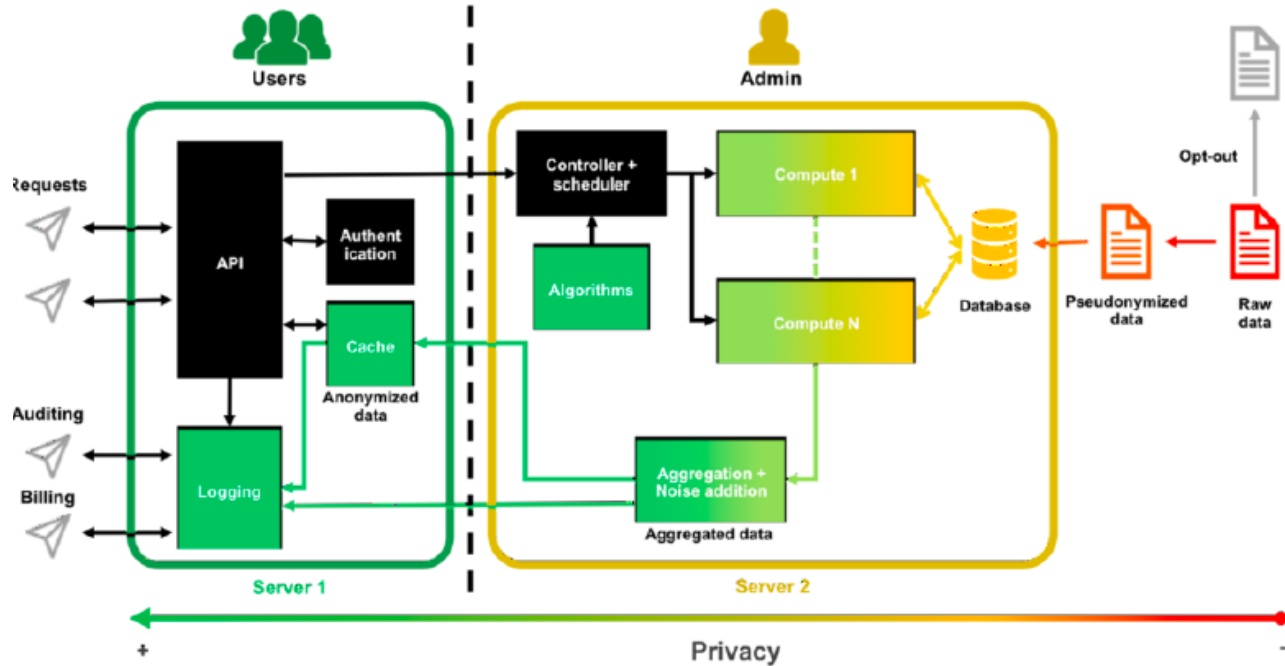


The 'Data for Refugees Challenge' Study

- Provides evidence-based insights into issues of refugee health, education, unemployment, social integration, safety, and security
- Serves as a sourcebook for refugee policy interventions based on big data analysis
- Describes best practices for ethically processing sensitive data on refugee mobility
- Presents results from the first big data challenge on refugees, offering insights into the dynamics of the Syrian refugee population in Turkey, currently the world's largest refugee population

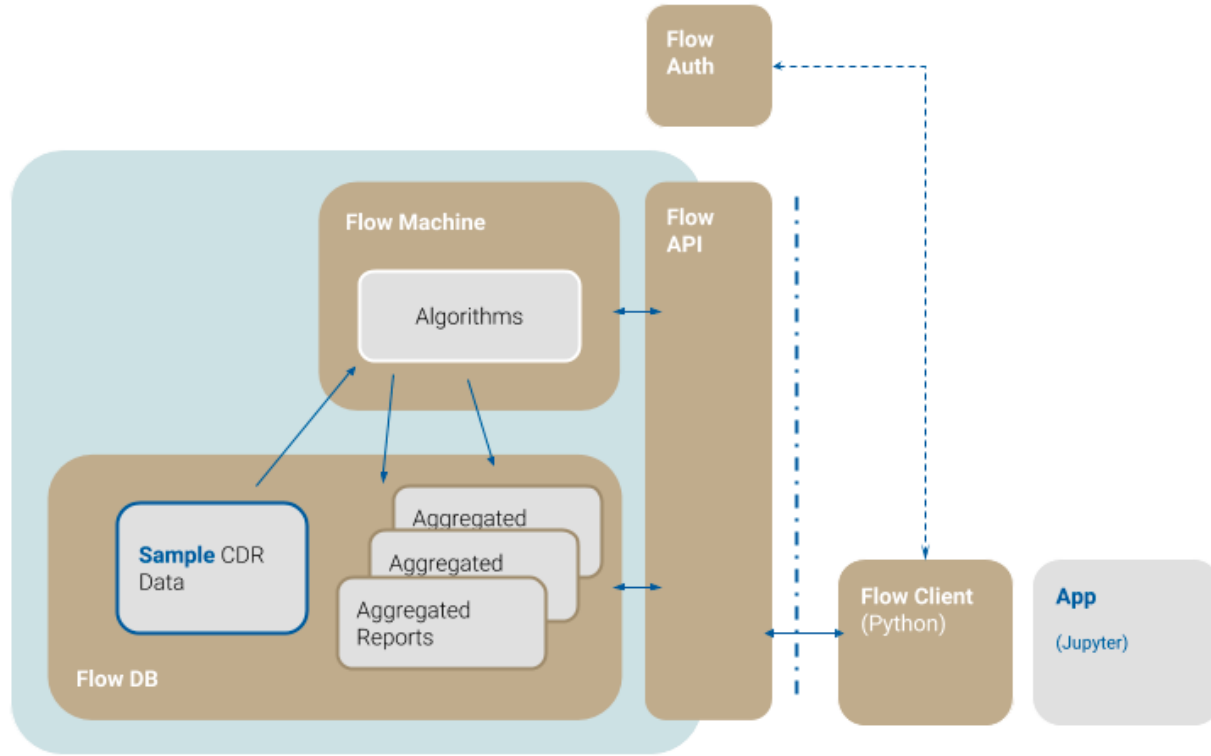


Open Algorithms (OPAL) sistemi



Letouzé, E. (2019). Leveraging Open Algorithms (OPAL) for the Safe, Ethical, and Scalable Use of Private Sector Data in Crisis Contexts. In Guide to Mobile Data Analytics in Refugee Scenarios (pp. 453-464). Springer, Cham.

FlowMinder – FlowKit Sistemi





Hummingbird Projesi

- Temel hedef: Göç konusunda veri eksiklerini gidermek
- 16 kurum, 10 ülke, H2020 RIA (2019-2023).
- Hedefler
 - Göç konusundaki temel belirsizlikleri tespit etmek, kavramları tekrar değerlendirmek
 - Göç örüntülerini bulmak, projeksiyonlar yapmak,
 - AB'nin göç politikalarına bilgi sağlamak,
 - Nitel senaryolar hazırlamak,
 - Büyük veri kaynakları ile göç ve göçmenler hakkında bilgi edinmek

HumMingBird alıřtayı

Expert workshop, Istanbul

'Making the most of mobile phone data to map migration'

17-18 Kasım 2022, Boęaziçi Üniversitesi

Davetli konuşmacılar:

Joshua Blumenstock, (*University of California, Berkeley*)

Stefano Maria Iacus, (*European Commission*)

Linus Bengtsson, (*FlowMinder*)

Özet yollamak için son tarih: 15 Haziran

<https://easychair.org/conferences/?conf=hmb2022>

<https://hummingbird-h2020.eu/news/event-items/EW17112022>

MENU ▾

nature
International journal of science

Subscribe



Search



Login

NEWS FEATURE • 29 MAY 2019

Can tracking people through phone-call data improve lives?

Researchers have analysed anonymized phone records of tens of millions of people in low-income countries. Critics question whether the benefits outweigh the risks.

Amy Maxmen





Utrecht University

Sorular:

a.a.salah@uu.nl

Twitter:

@SzassTam

Hesaplmalı sosyal bilimler ağı:

soc-comp-tr@googlegroups.com