Can Google Trends Forecast Forced Migration Flows? Perhaps, but Under Certain Conditions

Phillip Connor
Pew Research Center

Presented at Population Association of America meetings
April 2017
Chicago, Illinois

Big data have been used to forecast a variety of human behaviors, including migration. This paper uses the case of recent Syrian and Iraqi refugees entering Europe to explore whether online search data can be used to forecast forced migration. Comparisons between Google Trends and arrival/asylum seeker data demonstrate that online search activity might be a useful method for forecasting forced migration when the online search data can be specified for a migratory population with high technology access. Additionally, there must also be few barriers to migration (migrant resources, border controls) for online search data to predict migration.

Estimating international migration flows is a challenging science, fraught with data inadequacies. This is particularly true for forced migration since rarely are administrative data available to estimate these migration flows, let alone provide models for forecasting future movement.

Big data research, however, has demonstrated how the digital exhaust left by prospective and actual migrants on the internet and social media can be used to estimate international migration flows. The United Nations, for example, has explored how online search data can forecast international movement (UNFPA 2014). Other social media applications like Facebook (Hofleitner, Chiraphadhanakul and State 2013) and Yahoo email accounts (Zagheni and Weber 2012) have also been used to show global movement. However, very little research has examined the use of digital data in forecasting forced migration, a phenomenon that is estimated to involve more than 60 million people in the world or nearly 1% of the globe's population (UNHCR 2016a)

The unprecedented migration of refugees from war-torn countries like Syria and Iraq into Europe is one such migratory movement worthy of examination. Many of these migrants traveled through Turkey in 2015 and 2016, crossing into Greece via the Mediterranean and ultimately migrating to final destinations within continental Europe. Migrants were counted as they landed on Europe's shores (Frontex 2016; UNHCR 2016b) and again counted when claiming asylum in countries across Europe (Eurostat 2016).

The volume of people heading towards Europe in 2015 was of historic proportions, but so too was technology use in guiding their journeys. Recent reports (AFAD 2013; Sebti 2016) describe how the overwhelming majority of migrants carried their cellular telephones with them, using Google maps and social media to guide their way to the Greek islands and eventually through continental Europe (Kozlowska 2015).

Using Google trends data with administrative data on migration, this paper seeks to explore the potential use of online search data in estimating and predicting forced migration flows. This evaluation of online search data for forecasting forced migration uses the case of refugee movement from Syria and Iraq, through Turkey, into Greece and finally into Germany during 2015 and 2016. Not only do available data allow for such an exploration, events such as changing policies by EU governments over these two years permits for a changing environment for migration. These governmental policies, or period events, allow for a changing set of conditions whereby the link between online search activity and actual migration can be studied at different time intervals. In addition to presenting findings for this particular case, the greater objective is to describe the necessary conditions for search activity to be used in forecasting forced migration.

The study first provides some background information on how digital data has been used to estimate migration more broadly in recent years followed by a description of the Syrian/Iraqi refugee case under examination in this paper. Data and methods are explained followed by findings linking online search activity with actual refugee flows. Finally, based on the empirical evidence, some proposals are made of best-case conditions for using digital data like Google Trends to forecast forced migration. The paper is largely exploratory in understanding the relationship between online search activity and forced migration and the possible factors underlying its proper use. The paper does not aim to compute a hard set of numbers that predicts the rise and fall of forced migration based on online search activity. Such analysis requires additional cases of comparison that over time may become available.

Digital data and estimating migration

Digital data (sometimes referred to as big data) have been increasingly used to track and forecast human behavior. In fact, it has become a central part of sociological inquiry in recent years (Boyd and Crawford 2012). Sometimes considered digital footprints or digital exhaust of human activity, these data can include credit card purchases to shipping inventory to tweets. The analyses of these data help measure cultural shifts (Bail 2014), political engagement (Street, Murray, Blitzer and Patel 2015) and flu pandemics (Carneiro and Mylonakis 2009).

Big data is becoming an innovative data source for estimating migration as well. Researchers used Facebook data on the hometown city by Facebook users and how that correlates with their stated current city of residence (Hofleitner, Chiraphadhamakul and State 2013). Such correlations enabled researchers to view the most traveled migration corridors between cities among Facebook users, of which most were internal within countries from smaller towns to larger cities.

Researchers have also used Yahoo email accounts combined with IP addresses to determine emigration rates of Europeans within Europe (Zagheni and Weber 2012) as well as inferring international and internal migration patterns from Twitter data (Zagheni, Garimella, Weber and State 2014). In both research projects, changes in emigration rates of users of these internet sites indicate shifts in migration patterns as well as the gender and age breakdown of migrants.[1]

The United Nations Population Fund also has been exploring uses of digital data in estimating migration flows. In a recent report, they compared actual migration to Australia with job searches in Australia from internet users in other countries (UNFPA 2014). Internet searches in regions of countries with particularly

[1] The researchers aimed to also correct for selection bias for portions of national populations who don't use Yahoo email or Twitter.

high unemployment rates were highly correlated with actual migration to Australia, indicating that where migration motivation is high (presumably to find employment), digital data might be a useful tool for forecasting migration.

These empirical studies using digital data to estimate migration in general, however, have focused far less on forced migration. But, digital data is a likely appropriate form of data to use in forecasting forced migration. Often, forced migration is suddenly motivated. Also, the developing world – where most armed conflict that leads to forced migration is occurring – has become technologically mobile and has increased their use of the internet using mobile devices (Poushter 2016).

Europe's recent refugee surge

In the summer of 2015, chancellor of Germany, Angela Merkel, welcomed refugees crossing the Eastern Mediterranean from Turkey into Greece to continue making their way through other European countries and into Germany. In all, well over 1 million people crossed the Eastern Mediterranean during 2015 and 2016. Most continued their journey through Balkan states, into Hungary and Austria before applying for asylum in Germany and other countries further north and west. As the migrant flow continued through the latter months of 2015, countries such as Serbia and Hungary began to close their borders to migrants, often using physical barriers and security personnel to keep migrants out of their countries.

The flow of migrants continued in 2016. Faced with mounting pressure by European governments to stop the flow, the European Union negotiated a deal with Turkey in March 2016 that permitted Europe to return any future arrivals to Turkey. In exchange, Europe would consider a number of country integration

agreements with Turkey including easing visa restrictions for its citizens and billions of Euros to assist refugees living within Turkey's borders.

About four-in-ten (38%) asylum applications in 2015 and 2016 were from Syrians and Iraqis. Both groups mostly speak Arabic, an uncommon language in Turkey where many passed through from Syria and Iraq into Europe. Arabic is also a rarely used language in Germany where more than half (57%) of Syrian and Iraqi refugees entering Europe in 2015 applied for asylum. The relatively small number of Arabic speakers in Turkey and Germany prior to the refugee surge offers a unique opportunity to analyze trends of internet search terms in Arabic as mostly belonging to recent Syrian and Iraqi refugees.

Consequently, the movement of Syrians and Iraqis into Europe during 2015 and 2016 provides a suitable case for connecting big data with forced migration. With a restless, waiting population of refugees within Turkey combined with Germany's temporary withdrawal from the Dublin Regulation that forced asylum seekers to apply for asylum in the first EU country they enter, the migration gates flung open for a potential migrant population. And as Syrian and Iraqi refugees used their smart phones to chronicle and guide their way to Greece, they used Google search for assistance. Their digital fingerprints are easily isolated from the majority-population in Turkey and Germany because of their use of Arabic – a language spoken by less than 1% of the resident population in Turkey and Germany. Finally, the timing of various search terms like "Greece" which peaked during pre-dawn hours shows the high reliability for such travel terms.

Conditions for tracking migration from online search data

In this paper, I propose two broad necessary conditions for tracking migration using online search data: 1) digital conditions – adequate technology coverage and internet penetration of the population, analytic isolation of the population and construct validity of search terms under examination, 2) migration conditions – few border controls that would impede potential migration, sufficient motivation for migration and necessary resources to migrate.

Previous work employing big data to estimate and forecast human activity is predicated on the notion that the data is representative of the population. However, rarely do big data represent the entire population under study. For example, not everyone has a cell phone or a data plan. Migration studies relying on internet data in assessing movement have either stated potential biases based on internet penetration rates or have sought to correct the estimates using weights. Whatever the solution or set of caveats, it is critical that any analysis using internet search data to estimate migration is drawing on a population with near complete internet penetration (regular access to broadband at a low price) and adequate resources to obtain technological devices to access data plans or wifi.

It is also important that internet search data analysis indicating migration isolates the studied population. This is challenging since publicly available data using internet search data can only be limited by language selection, broad geographic location and time. However, whenever possible, it is important to remove any competing populations that may influence correlations in the findings. Otherwise, trends appear less defined or may even indicate a reverse trend for a larger population that shares the same time, language and physical space as the one under study.

Lastly, the right search terms must be discovered and available for analysis. These terms must show correlations to other search terms that are indicative of the concept being measured. If at all possible, it is helpful to know that the terms are highly searched for terms in the population. Otherwise, trends presented may not be externally valid.

Would-be migrants may search the internet for possible migration options, but ample motivation to move can be impeded by the practicalities on the ground to make the move. Thus, the relationship between internet search data and actual movement, even when necessary digital conditions are in order, can be weak. For example, migration policy is known to highly influence the flow of people across international borders (Zolberg 1989). For forced migration, this practically involves border controls between nation states as well as legal frameworks permitting the successful acceptance of forced migrants.

Migrants searching online for opportunities to move must also have sufficient motivation to make the move. Migration theorists have pointed to a number of push and pull factors that lead a person to migrate, including economic motivators, both short-term and long-term, as well as influences from family and conditions in the local community (Massey et al 1999). Among forced migrants, security is the number one priority. Therefore, using digital data as an indicator of forced migration requires dire conflict situations before large masses of people will move when searching for migration pathways on the internet. Internet trends may not be a one to one match of an online search leading to a migration episode every time, but the trend over time can show the increases and decreases of forced migration flows.

Even with ample motivation to move based on security conditions and adequate migration policies whereby forced migrants can cross international borders and have a high degree of success in staying in the new

country, there remains the issue of resources to make the journey. This not only involves sufficient funds to pay for the journey but also networks – friends and family or human traffickers – to assist with movement.

This paper uses a case study of Syrian and Iraqi refugees entering Europe as an example of a population that meets the aforementioned criteria. For digital conditions, several research and media reports indicate both the high internet penetration rate among Syrian and Iraqis in Turkey – the transit country for many seeking later refuge in Europe and an area of the world where internet penetration has continued to climb by a record pace (Poushter 2016). Many also relied on technological devices like cell phones for communication with family and friends and assisting their movement (AFAD 2013).

Since the language of Syrian and Iraqi refugees is Arabic, this language's relatively infrequent use of Arabic in countries under examination (Turkey and Germany) provide a natural experiment to track the relationship between internet online searches and migration as largely representing this new, suddenly mobile, population. Also, news reports (Al Jazeera English 2015) and correlation data indicate which search terms have been used the most and for migration purposes. This offers a relatively high degree of reliability in the selection of the best, and most widely used, search terms for analysis.

The Syrian/Iraqi refugee case during 2015 and 2016 also offers a rare opportunity when necessary migration conditions were also met. During the height of the crisis, Germany withdrew from the Dublin regulation directing asylum seekers to apply for refugee status in the first European Union country they enter (for most Syrian/Iraqi refugees, this meant Greece). By opening Germany's door, policy restrictions were eased, at least temporarily, for refugees to make the move to Europe. Europe also has an extensive liberal framework that includes full legal process of would-be asylum seekers, another important aspect for those weighing the decision to migrate.

The multi-year civil war in Syria has been one of the world's most bloody and deadly in recent memory. As the number of refugees entering Turkey – the main transit country into Europe – increased, resources for refugees remaining in Turkey became tight. Lack of security in both Syria and Turkey became unbearable for most in 2015, providing ample motivation to make the move to Europe. Moreover, many Syrian and Iraqi refugees had resources from their home countries to make the journey (AFAD 2013).

With most necessary conditions for using internet search data to estimate forced migration in place, this paper uses the recent surge of Syrian and Iraqi refugees into Europe as a case example. Without additional cases, the paper does not aim to test these hypotheses, but seeks to provide a beginning point of exploration in how internet search data can be used to estimate and possibly forecast migration, especially among forced migrants.

Data and Methods

This paper relies on data from Google Trends, an analytical tool that permits data users to explore search volume by language and location over a specified time period. Online searches are standardized on a 0 to 100 scale, meaning that lower numbers for a certain day, week or month are the lowest time when the search term was searched while higher numbers represent dates search terms were searched the most frequently.

Google is the most widely used global search engine and is the top search engine used in Turkey and Germany, countries used in the analysis of this paper. Google also is the top search engine of choice in Syria and Iraq and assumed to be the top search engine used by refugees coming from these countries.
It appears that Arabic speaking refugees had near universal access to technological devices and mobile networks to search the internet while living in Turkey. A 2013 report sponsored by the Turkish government

(AFAD 2013) indicates that nearly 90% of Syrian refugees in Turkey had a mobile phone. At the same time, internet usage rates in Turkey has risen considerably during the past decade, with more than two-thirds (68%) of the Turkish adult population in 2015 indicating they use the internet at least occasionally (Poushter 2016).

The analysis in this paper assumes that most Syrian and Iraqi refugees use Arabic as their primary search language on the internet. This assumption is important since Arabic was not commonly spoken in Turkey (less than 1%) or Germany (less than 1%) before thousands of Syrian and Iraqi refugees entered these countries.[2] Thus, any major shifts in trends for search terms in Arabic are assumed to represent new arrivals of Syrian and Iraqi refugees. Thousands of Syrians and Iraqis also settled in other European countries, including Austria and Hungary. However, Google Trends data were either unavailable or provided no identifiable trend for these and other European countries.

Consequently, sharp shifts in online searches are assumed to represent new Syrian and Iraqi arrivals and not other Arabic-speaking minorities within Turkey of Germany. To support this assumption, the Arabic term for "Greece" searched in Turkey occurred mainly along coastal regions and along the Syrian border. Syrian refugees were largely concentrated in these regions in 2015 and 2016, waiting along the coast to go to Europe via Greece or taking refuge from the Syrian civil war on the Turkish side of the Turkey-Syria border.

It is possible that some searches for "Greece" in Arabic could come from Turkey's large Kurdish population. However, searches for "Greece" in Arabic were not prominent in regions of Turkey where

[2] Arabic is the majority language of Syrians and Iraqis in their home countries, but less than 1% of the non-refugee population in Turkey speaks the language. Similarly, an estimated 0.5% (or about half a million) of Germany's population spoke Arabic prior the arrival of most Syrian and Iraqi refugees in 2015 and 2016, based on the number of foreign-born population from Arabic-speaking countries from United Nations migrant estimates for mid-year 2015.

Kurdish is widely spoken. Nor does it appear that the small, non-refugee, Arabic-speaking population in Turkey was searching for "Greece" as a potential tourist location. If this had been the case, there most likely would have been upticks in online searches for "Greece" in the summer of 2016, something that did not occur.

Arabic is also a minority language in Germany. This analysis assumes that fluctuations in Arabic searches largely represent the online searches of new refugee arrivals in Germany. Thousands of Syrians and Iraqis also settled in other European countries, including Austria and Hungary. However, Google Trends data in Arabic were either unavailable for other European countries or when available provided no consistent direction in the trend.

Selection of search terms presented in this report does not rely on top search terms in Arabic in Turkey or Germany. Google does not release top search terms for minority languages. Instead, several terms thought to show migration were explored such as "smuggling", "asylum" and geographic destinations in Europe such as "Germany" and "Sweden".

In many cases, geographic and migration terms were linked together. For example, the Arabic word – يونان الـ – for "Greece" among internet searches in Turkey is often searched with the Arabic word for "smuggling", according to Google Correlate, a tool that presents statistics on relationships between search terms. Within Turkey, "smuggling" in Arabic was also searched alongside "seek refuge", further supporting the notion that "Greece" was used alongside other search terms of potential migrants. Searching for map terms like "Greece" for the purposes of migration is also consistent with a report on Google search activity for Syrians within Syria in 2015 when "immigration to Germany", "asylum in Germany" "map of Europe", "Izmir" (primary coastal town in Turkey for departure to Greece) and "Greece" were listed as some of top searches as the migration crisis in Europe began.

The term "Greece" in Turkey was selected above others because of its closer proximity for migration to Europe. It is assumed that "Greece" is less of an aspirational search, but a practical one for those searching assistance with actual travel. At the same time, trends of the Turkish word for "Greece" in Turkey were flat, suggesting the trend identified in the Arabic use of the term is unique to refugees in the country and not the Turkish population as a whole. [3]

At the same time, the Arabic term for "Greece" within Turkey appears to be intended for migration purposes and not part of broader Arabic searches for the same term worldwide. The Arabic term for "Greece" worldwide, for example, peaks for a single week in July 2015 and subsides within a few weeks.

In Germany, the Arabic word – ألماني ة – for "German" is meant to indicate interest in learning the German language or performing online translation. Other terms in Arabic such as "integration course", "school", "jobs" were also examined, but showed no notable trends with the number of asylum applications. At the same time, there was no noticeable trend for the search term "German" in the German language in either Germany or worldwide. Also, there was no identifiable trend in Arabic searches for the Arabic word for "German" worldwide. This suggests that the trend observed in Arabic searchers for the word "German" in Germany is uniquely representative of Arabic speakers in Germany. [4]

Google Trends do not release total search volume. Instead, it standardizes search volume on a 0 to 100 scale across the time period under examination, with higher values indicating times when search volume was

[3] A more robust selection of search terms would involve a series of terms in Arabic, determining all possible searches for this series of terms with additional words (possibly using Google's autocorrect feature) and then examining trends for each of these terms. This paper, however, is an exploratory pursuit. Further research comparing like instances to this one could conduct more robust testing.

[4] Analysis of similar language terms in other leading asylum countries for Arabic-speaking refugees (for example, "Swedish" in Sweden, "Dutch" in the Netherlands) were conducted. These analyses did not show any correlation with monthly asylum applications or provided no results when entered into Google Trends.

highest during the time period. With this metric, search volume is less important than the trend over time and how this correlates with shifts in migration during the same time period. Google Trend data from Jan. 1, 2015 through Dec. 31, 2016 are in weekly estimates. These figures change slightly between downloads.[5] To smooth out the variation, most charts in this report show monthly averages taken from weekly search volume.

Administrative data for comparisons between Google Trends and recorded migration is derived from UNHCR arrival data and asylum application data from Eurostat. Both data sources offer monthly counts of arriving refugees or asylum seekers. Because it can take time for travel (into and throughout Europe) and migrant registration to occur, monthly migrant counts are compared with Google Trends with a two month lag.[6] Since monthly trends from Google are unavailable, the average weekly trend is calculated for each month under examination.

The presentation of the findings involves visual comparisons of Google Trends with migrant records from administrative data. (Charts were unavailable for the submission deadline, but will be presented at the conference). Comparisons are first made for all monthly arrivals (based on data from UNHCR) into Greece and Google Trends data of searches for "Greece" in Arabic within Turkey.[7] The same set of Google Trends from Turkey are then compared with combined monthly total of asylum seekers from Syria and Iraq data for all of Europe (European Union countries plus Norway and Switzerland, for a total of 30 countries). A further examination of hourly timing of Google Trends and the likely movement of refugees is also

[5] The variation between downloads in any given week's standardized search volume, rarely exceeds +/- 10 on the 0 to 100 scale. A more robust approach would be to download a series of samples from Google Trends and taking the average weekly set of scores and adjusting them to monthly averages for the purposes of matching monthly migration data.

[6] Various lags were examined, including no lag, 1,2 and 3 months. Correlations between Google Trends and recorded migration was highest with a 2 month lag, but correlations for all examined lags remained high, well above a Pearson's R of 0.4.

[7] Monthly arrival data from UNHCR does not allow differentiation for Syrian and Iraqi refugees alone for all months under examination, thus counts for all refugees, regardless of nationality, are used.

presented. Finally, searches for the word "German" in Arabic within Germany are compared with asylum seeker data from Germany for Syrian and Iraqi asylees.

Metrics examining the correlations between Google Trends and administrative data largely rely on Pearson's R. This measure is not ideal for time-based analysis, but provides a rough metric to show how the two measures of migration (Google Trends and administrative data) compare. Pearson's R also provides a sense of the relationship strength at different points of the two year time period of 2015 and 2016, particularly following the March 2016 agreement between the European Union and Turkey to halt migration. This paper is exploratory, thus its aim is not to create a metric that with x-change in Google Trend, a y-change in migration will occur. It is the relationship found in the trend over time that is important for this analysis. Thus, results are not focused on numbers, but more on the visual relationship between the two variables.

Findings

Note: Charts were unavailable for the paper submission deadline, but will be presented at the conference.

Online search activity is highly correlated with arrivals of migrants on Greece's shores for most of 2015 and 2016. Google Trend for the Arabic search of "Greece" in Turkey is largely in step with the flow of migrants registered with UNHCR two months later.[8] This correlation continues for all of 2015 and for the first few months of 2016. The relationship starts to deteriorate come February and March 2016.

The correlation for the remaining months of 2016 (April through November) is not strong. Because of the scale, the fluctuations of arrivals alongside Google Trends are actually not that correlated at all. In other words, it appears that the change in migration policy following the EU-Turkey deal may have interrupted the capacity to predict migration based on Google Trends. One of the defining conditions hypothesized in using online search data to predict forced migration – unfettered capacity for movement – was removed, breaking the correlation between online searches and actual movement.

A similar pattern of migration correlated with online searches occurs when using asylum application data throughout Europe. The rise of Arabic online searches for "Greece" in Turkey is largely correlated month to month with the rise of asylum applications specifically for Syrians and Iraqis two months later throughout the European Union's 28 member countries, Norway and Switzerland. This correlation is maintained through all of 2015 and into much of 2016. Only in the latter months, following the EU-Turkey deal, does the correlation fall. However, unlike arrivals, asylum applications continue to be correlated with Google Trend activity among Arabic speakers in Turkey for several months into 2016.

[8] Unfortunately, UNHCR monthly arrival data for 2015 and 2016 are not available by origin country. Thus, all arrivals of which the majority in 2015 and 2016 were Syrian and Iraqis (both Arabic speaking) are included in the analysis.

An interesting sidebar to this discussion is the actual peak timing of searches within Turkey. A capture of hourly search activity during the final seven days prior to the implementation of the EU-Turkey deal indicates that the highest peak time for Arabic searches of "Greece" in Turkey was in the middle of the night, the likely time when boats were being cast into the Mediterranean. Pre-dawn hours are the typical time when most migrants make the journey across the Eastern Mediterranean. Such movement under the cover of darkness lowers the changes of detection by marine authorities while also allowing for daylight toward the end of the journey (Stockmans 2016). Unfortunately, Google Trends only lets users capture hourly data for the previous seven days. Therefore, it is unknown whether this hourly trend was consistently the case during the refugee crisis for most of 2015 and the beginning months of 2016.

As further evidence of the relationship between online searches and forced migration, an analysis of online search activity and movement *after* actual migration confirms that the relationship between online activity and forced migration is worthwhile for examination. The Google Trend for the term "German" in Arabic within Germany is tightly correlated with asylum applications specifically for Syrians and Iraqis in the same month. Thus, as the number of asylum seeker applicants rose, so did searches for the word "German" in Arabic, a common term that could be used by new Arabic-speaking residents for translation or other informational purposes. Only in the latter months of 2016 does the relationship start to break as the number of new asylum applications drops quickly.

Discussion and Conclusion

This paper sought to explore the relationship between online search activity and forced migration, using the Syrian and Iraqi case into Europe during 2015 and 2016. Results indicate there was a strong relationship between online search activity and migration trends. This was found to be true both in recent arrivals from

Turkey to Greece and asylum applications throughout Europe two months later. From the destination point of view, rising number of asylum applicants was associated with a rising Arabic searches of "German" in Germany.

Correlations between online search activity and migration were weaker following the EU-Turkey deal in March 2016, possibly indicating that this change in policy had a bearing on the effectiveness of using online searches for forecasting forced migration. The paper hypothesized necessary digital and migration conditions in using online search activity to predict forced migration. It appears the necessary migration condition of open borders altered the potential effectiveness of using online search activity in forecasting the migration of refugees in Turkey into Europe. It does not appear the necessary digital conditions changed, although some other aspects of migration conditions (adequate resources, networks) could have been trending downward after a prolonged period of mass migration. In other words, the number of refugees with the necessary resources (money and networks) to migrate into Europe may have dwindled and coincided with the timing of the EU-Turkey deal.

This Arabic-speaking migrant case into Europe may be unique. For example, it does not appear that migrants from sub-Saharan Africa traveling through Libya, across the Mediterranean into Italy have the same access to technology and internet as Syrians and Iraqis in Turkey. Also, in locations where sudden movement of forced migration occurs due to conflict, necessary migration data to correlate search activity may not always be available. Moreover, isolating the population under examination by language, like that done by Arabic speakers in non-Arabic speaking locations, may not always be possible.

Without additional cases of forced migration, however, it is not possible to calculate an approximate level of migration that can be associated with changes in internet searches. Also, actual search volume would assist

in making such calculations for the future, something Google Trends does not make publicly available. As new situations of sudden, forced migration arise, researchers should test correlations between online searches and reliable migration data so that our knowledge of both the necessary conditions for using online search data to forecast migration can be understood.

References

AFAD (2013). Syrian Refugees in Turkey 2013, Field Survey Results. Resource document, *Republic of Turkey Prime Ministry Disaster and Emergency Management Presidency:* Ankara, Turkey. https://data.unhcr.org/syrianrefugees/download.php?id=4085

Al Jazeera English. (2015). The top Google searches in Syria. *Al Jazeera English:* Doha, Qatar. http://www.aljazeera.com/news/2015/09/top-google-searches-syria-150929151549994.html

Bail, Christopher. (2014). The cultural environment: measuring culture with big data. *Theory and Society* 43:465-482.

Boyd, Danah and Kate Crawford. (2012). Critical Questions for Big Data. *Information, Communication and Society* 5:662-679.

Carneiro and Mylonakis. (2009). Google Trends: A Wed-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases* 49:1557-1564.

Frontex. (2016). Fran Quarterly, Quarter 4, October through December 2015. *Frontex:* Warsaw, Poland. http://frontex.europa.eu/assets/Publications/Risk_Analysis/FRAN_Q4_2015.pdf

Hofleitner, Aude, Ta Virot Chiraphadhanakul and Bogdan State. (2013). Coordinated Migration. *Facebook data science papers*: Menlo Park, California. https://www.facebook.com/notes/facebook-data-science/coordinated-migration/10151930946453859

Massey, Douglas, Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino and J. Edward Taylor. (1999). *Worlds in Motion: Understanding International Migration at the End of the Millennium.* Clarendon Press, Oxford.

Kozlowska, Hanna. 2015. The most crucial item that migrants and refugees carry is a smartphone. *Quartz*, New York, NY. http://qz.com/500062/the-most-crucial-item-that-migrants-and-refugees-carry-is-a-smartphone/

Poushter, Jacob. (2016). Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. *Pew Research Center*: Washington, D.C. http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/

Sebti, Bassem. (2016). 4 smartphone tools Syrian refugees use to arrive in Europe safely. *Voices, Perspectives on Development.* The World Bank: Washington, D.C. http://blogs.worldbank.org/voices/4-smartphone-tools-Syrian-refugees-use-to-arrive-in-Europe-safely

Street, Alex, Tomas A. Murray, John Blitzer and Rajan S. Patel. (2015). Estimating Voter Registration Deadline Effects with Web Search Data. *Political Science,* doi: 10.1093/pan/mpv002.

Stockmans, Pieter. (2016). Diary of a Syrian refugee family trying to reach Greece. *Al Jazeera English*: Doha, Qatar. http://www.aljazeera.com/indepth/features/2016/02/diary-syrian-refugee-family-reach-greece-160202112221725.html

UNFPA. (2014). Estimating Migration Flows Using Online Search Data. Global Pulse Project Series no. 4. United Nations Population Fund. http://www.unglobalpulse.org/sites/default/files/UNGP_ProjectSeries_Search_Migration_2014_0.pdf

UNHCR. (2016a). Global Trends: Forced Displacement in 2015. Geneva, Switzerland. http://www.unhcr.org/576408cd7

UNHCR. (2016b). Refugees/Migrants Emergency Response – Mediterranean. Data on refugee arrivals for Mediterranean countries in Europe. http://data.unhcr.org/mediterranean/regional.php

Zagheni and Weber. (2012). You are where you email: using email data to estimate international migration rates. *ACM Web Science Conference,* 348-351.

Zagheni, Kiran, Weber and State. (2014). Inferring International and Internal Migration Patterns from Twitter Data. *ACM Web Science Conference*, Seoul, Korea.

Zolberg, Aristide R. (1989). The Next Waves: Migration Theory for a Changing World. *International Migration Review* 23:403-30.