

Predicting Refugee Flows from Ukraine with an Approach to Big (Crisis) Data: A New Opportunity for Refugee and Humanitarian Studies

By Tado Juric*

This study was created due to the need to predict the migration flows of refugees from Ukraine to the EU in the absence of official data. We present a descriptive analysis of Big Data sources, which are helpful in determining, as well as for estimating and forecasting refugee emigration flows from Ukraine and help crisis managers. The objective of this study was to test the usefulness of Big Data and Google Trends (GT) indexes to predict further forced migration from Ukraine to the EU (mainly to Germany). The primary methodological concept of our approach is to monitor the digital trace of Internet searches in Ukrainian, Russian and English with the GT analytical tool. The control mechanism for testing this sort of Big Data was performed by comparing those insights with the official databases from UNHCR and national governments, which were available two months later. All tested migration-related search queries (20) about emigration planning from Ukraine show a positive linear association between the Google index and data from official UNHCR statistics; $R^2 = 0.1211$ for searches in Russian and $R^2 = 0.1831$ for searches in Ukrainian. Increase in migration-related search activities in Ukraine, such as “граница” (Rus. border), кордону (Ukr. border); “Польща” (Poland); “Германия” (Rus. Germany), “Німеччина” (Ukr. Germany) and “Угорщина” and “Венгрия” (Hungary) correlate strongly with officially UNHCR data for externally displaced persons from Ukraine. The results show that one-fourth of all refugees will cross into Germany. According to Big Data insights, the estimated number of expected refugees until July 2022 is 5.9 Million refugees and mid-2023 Germany can expect 1.5 million Ukrainian refugees.

Keywords: *refugee, forecasting refugee flows, Ukraine, big data, Google trends, forced migration, UNHCR*

Introduction¹

The Ukraine war that started on 24 February 2022 destroyed civilian infrastructure and forced people to flee their homes to seek safety, protection, and assistance. According to UNHCR (2022c), more than a million refugees from Ukraine crossed borders into neighbouring countries in the first week since the war outbreak (2.5 Mil. in the second week), and many more are on the move both inside and outside the country. As the situation unfolds, an estimated 4 million people may flee Ukraine (UNHCR 2022a). The UNHCR estimated on February 27 that there would be 7.5 million internally displaced people in Ukraine in two months. According to the UNHCR, 18 million people are affected by the conflict, and 12 million people will require health care (UNHCR 2022a).

* Associate Professor, Catholic University of Croatia, Croatia.

¹The pre-print version of this study is published on 16 March 2022 on pre-print server *MedrXiv*, <https://doi.org/10.1101/2022.03.15.22272428>.

Nobody knows how long the war will last, what kind of conflict this will develop into, and the civilian toll. This crisis is just one of many during which having reliable data would have assisted UNHCR and UN agencies in preparing high-quality projections for emergencies. But such data are often either unavailable or only available with a considerable time lag, which renders them useless for operations and emergency preparedness (UNHCR Global Data Service 2021).

This study was created due to the need to predict the migration flows of refugees from Ukraine to the EU. We present a descriptive analysis of Big Data sources, which could be helpful in determining, as well as for estimating and forecasting refugee emigration flows from Ukraine.

The expression “Big Data” has been spreading since 2011; the term is used in academia, industry and the media, but it is not even today precisely clear what it means. In this study, the term “Big Data” refers to the interdisciplinary method for deriving new understanding from massive aggregations of the information sampled through Google analytical tools Google Trends. We refer to the term Big (Crisis) Data as the subset of Big Data sources that have been applied in the humanitarian work of UNHCR (UNHCR 2021).

Since there has been a substantial increase in internet access compared to creating credible migration monitoring registration systems (Syratos et al. 2018), the development of statistical tools that combine traditional and new sources of information is likely to become an accepted approach for monitoring migration and refugee flows. In demography, researchers have begun to use non-traditional or alternative data (mobile phone records, social media use, satellite maps, and internet-based platforms) to understand migration and mobility in light of new methodological approaches (Wanner 2020). This study will show that the analytical tool Google Trends (GT) can give valuable complementary data in the field of migration and refugee studies and be useful in the current refugee crisis.

The structure of the paper is as follows: after briefly showing the results of relevant studies, we explain this study’s methods and show the limitations of the method. We then discuss Ukraine refugee flows and reveal the results achieved via this approach. In the section on results, we show the correlation between the Google search index and the official UNHCR statistics and discuss how to forecast refugee flows with Google Trends.

Forced Migration from Ukraine

Ukraine was one of 15 republics of the Soviet Union and became an independent state in 1991² (Britannica.com); it had a population of approximately 44 million. Since the end of the Soviet Union, the country had been a country of origin, transit and arrival for forced migration, as well as a provider of “durable solutions” for ethnic Ukrainians returning to the country from which they had been

²Britannica. *Ukraine: History*. Britannica Online Encyclopedia. Retrieved from: Britannica.com [Accessed 6 March 2022]

exiled during the Soviet era (Bergtora and Garnier 2022). In the 1990s, the ethnic composition of Ukraine changed significantly. Many ethnic Russians left, and many ethnic Ukrainians returned, including Tatars originally from Crimea, whose returned population increased fivefold between 1989 and 2001. The suspension of negotiations of the EU association agreement by a pro-Russia government in 2013 led to the Maidan protests in Kyiv and a pro-European regime change the following year, rapidly followed by Russia's occupation of Crimea (2014) (Bergtora and Garnier 2022). 5.2 million people were affected by the conflict, and 1.6 million were displaced within and outside Ukrainian borders (Bergtora and Garnier 2022). As of 2019, there were 1.4 million internally displaced people in Ukraine. Most Ukrainians sought refuge in Russia, but many also went to EU countries.

In refugee, migration and humanitarian studies, little attention has been given to the continued refugee resettlement of Ukrainians (c.f. UNHCR 2022b). A larger body of work can be found in area studies (cf. Bergtora and Garnier 2022), including research on the Ukrainian internally displaced people (IDP) (Kuznetsova and Mikheieva 2020) in Crimea (Charron 2020) and Donbas (Sereda 2020) and the humanitarian crisis (Scrinic 2014).

In the recent exodus from Ukraine, UNHCR estimates that over 4 million people could flee from Ukraine and seek protection and support across the region (UNHCR, Ukraine 2022) (Table 1). Our estimates obtained with Big Data show that there will be at least 5.4 million refugees. According to data from UNHCR, the speed of the exodus is already more extensive than the migration crisis of 2015, when 1.3 million asylum seekers from Syria, Iraq, Afghanistan and Africa, fleeing poverty and wars, entered Europe (UNHCR, 2016). "This is the fastest-moving refugee crisis in Europe since the end of the second world war" (The Guardian 2022).

Most Ukrainians (53%) fled in the first week to Poland, which welcomed at 04 March 2022 about 756,303 people, followed by Hungary, with 157,004. The number of refugees is also high in Hungary, Romania, Slovakia and Moldova. But not everyone decides to stay in the countries they first arrive in from Ukraine. For example, 140,000 people who came to Romania during the first eight days of the war travelled to other countries, leaving about 60,000 in Romania, according to the UNHCR (2022a) (Figure 1).

Figure 1. *Refugee from Ukraine (UNHCR)*



Source: UNHCR, <https://data2.unhcr.org/en/situations/Ukraine> (08.03.2022).

Table 1. *A Total Projected Refugee from Ukraine by July 2022 (UNHCR)*

Hosting Countries	Total projected refugee population by July 2022*
Hungary	250,000
Moldova	100,000
Poland	1,500,000
Romania	250,000
Slovakia	60,000
Other Countries	1,840,000
Total	4,000,000

Source: UNHCR. *Ukraine situation regional refugee response plan*. Retrieved from: <https://data2.unhcr.org/en/situations/Ukraine>. [Accessed 8 March 2022].

Refuge Welcome Policy in the EU

In just one week, into the EU came so many refugees as in the entire so-called Yugoslav war (cf. Radelić et al. 2006). In the light of the current crisis, we believe that it is useful here to look back at the breakup of Yugoslavia and try to predict the course of the refugee crisis from lessons from the past. During the Homeland War in Croatia, 150,000 Croats, or 3% of Croatia's population, emigrated to Germany, while about 250,000 arrived from BiH. Of that, 1/3 remained there forever, 1/3 returned because their temporary stay (so-called *Duldung*) expired, while 1/3 returned of their own free will (cf. Jurić 2021a).

If we draw a parallel between the war in Ukraine and the Homeland War and assume that the same number of Ukrainians as former Croats under the pressure of war will escape to Germany, we expect that at least 1.5 Million Ukrainians will emigrate to Germany. We assume that further migration waves will occur in the years to come when men are allowed to leave the country and emigrate to

Germany due to family reunification and other reasons (for now, men between the ages of 18 and 60 cannot leave the country).

There are numerous differences in the welcome policy of Germany today and Germany in the 1990s. For entry to Germany during the Homeland war in Croatia, Croats asked the so-called letter of guarantee, i.e., proof that one of the relatives, etc., would guarantee the care of refugees. At the supranational level in the current crisis, the EU has triggered the adoption of the EU temporary protection directive for the first time. The directive grants immediate protection to Ukrainians fleeing to war for three years (Bergtora and Garnier 2022). Ukrainians do not require Schengen visas to enter the EU, and there is no other restriction. The EU admits Ukrainians without valid passports if they pass an individual assessment conducted by border officials. Moreover, Germany openly encourages the immigration of Ukrainians. Namely, Ukrainian nationals will be given the right to live and work in the European Union for up to three years without claiming asylum (Visitukraine 2022).

What has changed from the 1990s to today in Germany? In short, the demographic picture. Today, due to the demographic crisis in its own country, the lack of labour force of all profiles and young people in general, Germany needs to import about 400,000 employees every year to maintain its pension and health care system, according to a study by Bertelsmann Stiftung (2015). According to projections, the potential labour force in Germany will decrease by 16.2 million workers between 2012 and 2050 for purely demographic reasons (Arbeitsagentur 2020). The German generations with the highest birth rates will have left working life around 2035. According to model calculations, the net migration with countries of the EU will soon drop significantly from the current number to slightly below 300,000 (Arbeitsagentur 2020). For the next 36 years, an average of between 276,000 and 491,000 people would have to immigrate from third countries every year to Germany for the labour force potential to remain constant (Arbeitsagentur 2020).

Our primary hypothesis is that up to one-third of all refugees will cross into Germany. Surely, no refugee from Ukraine will be forced to stay in Germany, as neither are Croats during the Homeland War, but the combination of fear, a better standard and open hospitality will influence this decision.

The German Trade Union Confederation (DGB) and the German Employers' Union (BDA) issued a joint statement calling on the German government to urgently remove "legal and bureaucratic" obstacles to the rapid integration of Ukrainian refugees into German labour market. "Companies and works councils are ready to take their share of responsibility in order to take over these people as soon as possible, train or retrain them and thus integrate them into the labour market," the joint statement said (Deutsche Welle 2022).

Furthermore, the German Railways (Deutsche Bahn), three days after the start of the war, already on 27 February, put at its disposal trains for free transport of all Ukrainians from Poland to Germany (Deutsche Bahn 2022).

Some authors state that refugees from Syria and Africa have received worse treatment than refugees from Ukraine (Abcnews 2022). In the 2015 European migrant crisis (Almustafa 2021), when 1.3 million people came to the continent to

request asylum (BMI 2015), the welcome policy was completely different - for example, many EU members have lifted the razor wire on their borders. In March 2016, Turkey agreed to close its border to the EU in exchange for money and diplomatic favours, which effectively stopped the further passage of refugees through Eastern and South-eastern Europe (Berkay 2020).

Innovative Data Sources in the Refuge and Migration Studies

Human migration within Europe is difficult to measure due to the lack of efficient statistical systems in numerous EU countries (cf. Jurić 2022a). Numerous researches have shown (Böhme et al. 2020) that obtaining timely, evidence-based information on potential migratory movements is of the utmost importance to develop early-warning strategies and set up the necessary reception mechanisms in transit and receiving countries.

Traditional data sources, based either on surveys or registers, generally fail to provide statistical information on refugee flows quickly and do not facilitate the accurate short-term anticipation of these flows (Wladyka 2017). This limitation is one reason underlying the development of new methods based on alternative sources, so-called Big Data (Jurić 2022a).

The potential of Big Data to facilitate early warning of crises lies in its ability to provide granular, almost real-time information in locations where there are few other data sources. According to UNHCR, some Big Data sources can have the power to fill some of the information gaps left by conventional data acquisition channels, especially during crises (UNHCR Blog et al. 2022). According to Ibáñez Sales (2021), there are three main types of Big Data sources that are highly relevant to complement traditional migration and refugee data sources: communication services (calls records and texts messages), geo-localised activity as well as Internet-based exchanges (social media activity, online searching preferences, and online money transfers).

The main advantages of this approach for refugee studies are that those data are easily collected and generated in real-time, they are incredibly robust, and they provide a profound insight into migrants' or refugees' opinions (cf. Jurić 2021c). This data can be used to gather insights into what was going on in the user's mind in a non-invasive manner (see section Restrictions about ethical concerns) (Wang et al. 2018). Moreover, digital traces provide documentation of both movement and activities, which can help researchers bypass possible sources of error in survey data, such as inability to recall, bias, and the like (Jurić 2021a). Finally, digital traces can provide access to groups that are difficult to reach or are generally underrepresented by traditional research techniques (Cesare et al. 2018).

According to Jurić (2022c) and Choi and Varian (2012), Google Trends data have been used in various types of research: US unemployment, flu outbreak, predicting consumer behaviour, predicting inflation rates, predicting the housing market, predicting stock market changes, modelling tourism demand, etc. All the research results showed that GT analytical tools could reveal valuable insights about intentions (Önder 2017).

Several case studies have demonstrated that the inclusion of Big Data sources significantly improves the power and accuracy of predictive models of refugee flows (Agrawal et al. 2016). A study by Singh et al. (2019) on internally displaced persons (IDPs) movements between provinces in Iraq shows that a mix of social media data and traditional register data improves the predictive quality compared to predictions based on register data alone. Therefore, Big Data hold huge potential for humanitarian organisations to improve their early warning systems and generate better contingency planning figures (UNHCR Global Data Service 2021, p. 11). Although previous research has established that digital data can be employed to study migration, there are still significant methodological issues and scepticism regarding the feasibility of using alternative data sources (Jurić 2022a).

In 2014 the UN conducted the first research on the use of Big Data for demographic research, with its report released in 2018 (United Nations 2014). These data were shown to provide useful insights into the quantitative and qualitative characteristics of international and other migrations (United Nations 2019). Since the UN confirmed the relevance of these data, explorations have been carried out on social networks (Zagheni et al. 2014, Zagheni et al. 2017, Zagheni and Weber 2015), and several studies have used Big Data sources to analyse migration-related phenomena directly (e.g., Dubois et al. 2018, Hawelka et al. 2014, State et al. 2014, Jurić 2021a, 2022a,b).

The European Commission concluded similarly to the UN that a) Big Data sources do not replace traditional data sources but can complement them, b) they can still be used to assess trends (Jurić 2021c). Furthermore, it has been established that both kinds of data sources can complement each other (Spyratos et al. 2018). “In addition to being extremely robust, these data are easily collected, are generated in real-time, and provide significant insights into the opinions of individuals” (Jurić 2022a).

According to Gabrilovich (2020), would-be migrants often use online searching to get answers about the country they plan to emigrate to, and according to Wanner (2020), Google is the first source of information for most users planning relocate. Several studies have used immense data sources to analyse migration-related phenomena directly. The first successful analysis of this type of data was in 2009, and the first study in the field of migration examined during the 2015 Migration Crisis searches for particular terms in Arabic in Turkey and Germany according to selected terms such as “Greece” or “Germany” (Connor 2017). A study by the PEW - research centre showed that digital prints left by internet searches could provide insight into the movement of migrants. Namely, during their travels in 2015 and 2016, many migrants used smartphones that provided access to information and maps and travel tips via social media. It was then unequivocally proven that these indicators could be used to predict migration (Jurić 2022b). Compared to approaches using social media, the advantage of Google Trends is that limitations related to fake accounts are not prevalent (Jurić 2022a).

Undoubtedly, the Internet penetration rate is for this method significant. When it comes to the use of Internet services, Ukrainians are generally comparable to the EU average (with a slightly lower share) (Internet World Stats 2022).

Migrants and refugees are nowadays increasingly using mobile devices and digital interactions on social media to communicate with relatives and send them news and pictures. Migrants use smartphones to facilitate their movements across borders, stay up to date with weather predictions, get last-minute information in transit and destination countries, or send and receive international remittances (Ibáñez Sales 2021). Moreover, it is proven that the refugees are more interested in information from the Internet than the average and that they generally use smartphones during migration (cf. Connor 2017).

Methodological Concept

In this study, we present a descriptive analysis of Big Data sources, which could be useful for predicting refugee flows from Ukraine to the EU and especially to Germany. The primary methodological concept of our approach is to monitor migration-related searches with the analytical tool Google Trends (GT)³ during the Ukraine war crisis. This tool shows the popularity of a specific term and shows if a trend is rising or falling. GT does not provide information on the actual number of keyword searches - instead, it standardises search volume on a scale of 0 to 100 over the period being examined, with higher values indicating the time when the search volume was greatest, allowing for verifiable metrics (Jurić 2021b). In previous studies, it was a significant limitation that each of these searches was conducted for its reason and did not answer researchers' questions, so "googling" the term "Germany" was not necessarily an implication that someone wants to move to Germany, but may be interested in tourist information or just looking for the German Bundesliga (Jurić 2022a). However, in this crisis, it is quite certain that no one is interested in the Bundesliga when googling the word Germany. However, it is essential to pay attention to the overall context by interpreting the results.

The Google search Index cannot estimate the exact number of searches, so with the help of this tool, the exact number of emigrants cannot be calculated, but the increase in the trend can be noticed very precisely (Jurić 2021a). This method we have already tested in our previous studies with quite good predictive indicators (Jurić 2021c, Jurić 2022a). The justification for using the GT method for the assessment of refugee flows is that mismatch between intention and actual behaviour, which is one of the most significant restrictions of this approach, is, in the case of refugees from Ukraine, not to such an extent represented as in the case of predicting migrations in peacetime (cf. Jurić 2021a).

For this method, it is essential to establish the time flow from the expressed intention to the realisation of migration. We hypothesise that an increase in the number of Google searches will, following a delay of about one to three days, translate to a rise in the number of refugees. According to Curry et al., trigger events happen immediately before migration, usually within a time frame of 1-2

³Google Trends. <https://trends.google.com/trends/?geo=HR>.

days before emigration occurs and are events that the individual perceives as threatening their integrity (Curry et al. 2018).

To standardise the data for Google Trends, we extracted data from 24 February 2021 to 24 February 2022. We then divided the keyword frequency for each migration-related word, which gave us a search frequency index that we then compared with official statistics using a linear regression method (cf. Wanner 2020, Wilde et al. 2020). In order to estimate the model, linear regression was used to measure the correlation between the number of searches (x) and the number of moves (y) evidenced by the official UNHCR statistics, which were available one to two months later, to prove the significance of our results.

Initially, keywords were chosen by brainstorming possible words that we believed to be predictive, specific, and common enough to forecast refugee flows (Table 2). After the significance test, we selected followed keywords and topics.

Table 2. List of Keywords

English	Ukrainian	Russian
<i>asylum</i>	<i>притулок</i>	<i>убежище</i>
<i>border</i>	<i>кордону</i>	<i>граница</i>
<i>border control</i>	<i>прикордонний контроль</i>	<i>пограничный контроль</i>
<i>migrant</i>	<i>мігрант</i>	<i>мигрант</i>
<i>refugee</i>	<i>біженець</i>	<i>беженец</i>
<i>interpreter</i>	<i>перекладач</i>	<i>переводчик</i>
<i>weather forecast</i>	<i>прогноз погоди</i>	<i>прогноз погоды</i>
<i>Schengen</i>	<i>Шенген</i>	<i>Шенген</i>
<i>visa</i>	<i>віза</i>	<i>виза</i>
<i>job</i>	<i>робота</i>	<i>работа</i>
<i>PCR</i>	<i>ПЛР</i>	<i>ПЦР</i>
<i>consulate</i>	<i>консульство</i>	<i>консульство</i>
<i>Poland</i>	<i>Польща</i>	<i>Польша</i>
<i>Germany</i>	<i>Німеччина</i>	<i>Германия</i>
<i>Romania</i>	<i>Румунія</i>	<i>Румыния</i>
<i>Hungary</i>	<i>Угорщина</i>	<i>Венгрия</i>
<i>Moldova</i>	<i>Молдова</i>	<i>Молдова</i>
<i>Slovakia</i>	<i>Словаччина</i>	<i>Словакия</i>

It is to note that we observed that Ukrainians use the Russian language more often to search for terms than Ukrainian. This phenomenon is numerous: During the 19th century, the Russian government and in the 20th century, the Soviet Union promoted the spread of the Russian language among the native Ukrainian population by suppressing the Ukrainian (Magoscy 1996). In independent Ukraine, although Russian is not an official language, it is widely spoken, particularly in regions of Ukraine where Soviet Russification policies were the strongest, notably most of the urban areas of the east and south (Donetsk, Luhansk and Crimea) (Magoscy 1996).

It is also necessary to consider the overall social context and that even before the war, many Ukrainians emigrated to Poland. 1.4 million Ukrainians lived in

Poland before the war broke out (most fled after the Crimea crisis) (The Guardian 2019).

Limitations of the Methodological Concept

Like all data sources, Big Data also come with limitations. According to UNHCR, issues with Big Data mostly come from three sources: bias, inaccuracy, and low scalability (UNHCR Blog 2022). GT has some specific limitations when applied in the context of forced displacement. Namely, many displaced people do not have the time to conduct detailed searches on the Internet for a potential host country (UNHCR Global Data Service 2021, p. 19). But the biggest weakness of GT data, which is biased due to limited access to the Internet, is in the case of refugees from Ukraine not represented. As already mentioned, it is proved proven that refugees are ready to allocate the most resources for smartphones, i.e., connection with relatives and friends, but also the Internet.

Although penetration rates of the Internet are increasing globally, they still do not reach 100% of the Ukrainian population. Rural communities, women, children, and the elderly are underrepresented population segments and risk becoming invisible in Big Data sources (UNHCR Global Data Service 2021, p. 7). On the other hand, there are numerous indications that potential migrants and refugees largely gather or refine information through Google searches before emigrating (Wladyka 2017). According to EUROPOL (2018), there has been an exponential growth in the use of the Internet and social media by refugees and migrants arriving in Europe in the last few years. Even the UNHCR has concluded that smartphones, the Internet and social media are now a key tool for migrants, who spend up to a third of their total budget on staying connected (UNHCR 2017).

Although the data obtained with GT are robust data with large samples, which provide information qualitatively different from what can be obtained from the official databases, they are not representative of the observed population (Jurić 2022a). A significant restriction is that GT does not provide data on which population was sampled or how it was structured (Wladyka 2017). A problem also exists in the researchers' education who must be skilled in computational methods, be transparent about their methods to ensure repeatability, and be accustomed to the interdisciplinary environment (Jurić 2022b).

We are also aware that refugees may use other methods to gather information on living and working conditions in Poland and Germany, such as the accounts of friends or family members who have already immigrated to these countries. A further aspect is the availability of humanitarian help in a potential host country, a network within an existing diaspora in the likely host country, and available information on flight routes and corridors from earlier migrants (UNHCR Global Data Service 2021, p. 11).

This study has a descriptive character and is limited by the lack of data due to the actuality of the phenomenon being observed. The exponential growth in the first phase of the war needs to be modelled via the second phase to detrend the time series. The whole process should be checked with significance tests to make conclusive statements about the observed quantities.

Twitter data is only effective if the user uses the geotagging option, typically using only 3% of users. Furthermore, the basic idea of monitoring the moves of refugees from country to country through geotagged messages usually requires a more extended period, which in our study, due to actuality, could not be done.

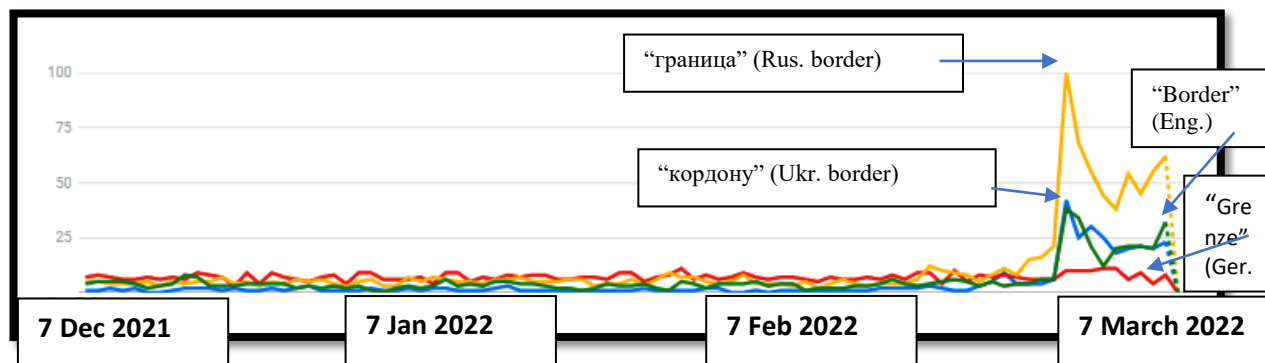
Regarding ethical issues, since the vast majority of Big Data is generated automatically by mobile and Internet users without informed consent and knowledge for what purpose those data are collected, its use may violate privacy and data protection (Ibáñez Sales 2021, p. 8). This situation also raises concerns about public-private cooperation because governments may become dependent on Big Data collection (Bircan and Korkmaz 2021).

Results

Use of the Google Trends Analytical Tool to Forecast Refugee Flows from Ukraine

Searching for queries in Ukrainian, Russian and English from Ukraine “border crossing”, “кордону” (Ukr. border), “граница” (Rus. border) from 7 December 2021 to 3 March 2022 (Figure 2) shows an upward trend since the outbreak of war. This is a strong indication that more and more Ukrainian citizens will emigrate, i.e., flee from Ukraine.

Figure 2. Search Queries in Ukrainian, Russian, German and English from Ukraine “Border Crossing”, “кордону” (Ukr. border), “граница” (Rus. Border) (7 December 2021 – 3 March 2022)



Source: Jurić (2022d).

A further indication is searching for terms related to registration in Poland, Hungary and Germany.

Figure 3. Fastest Growing Google Search Terms in Ukraine: Western Union, Asylum, refugee, Schengen (7 December 2021 to 7 March 2022)

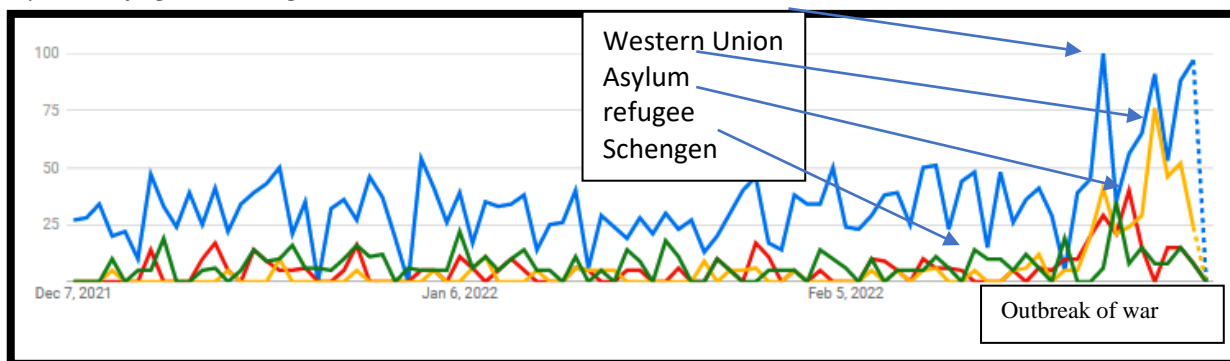
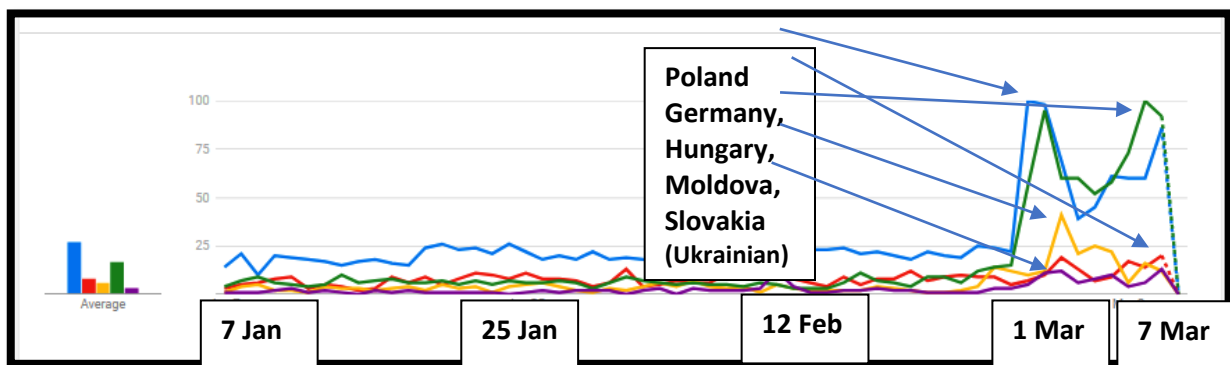


Figure 3 shows that the fastest-growing Google search terms in Ukraine (7 December 2021 to 7 March 2022), except mentioned term “border”, are Western Union, asylum, refugee and Schengen.

When looking at the showed interest in emigration countries, we notice that the internet traces correspond to the official data of the UNHCR. Namely, the GT, just like the UNHCR, shows that the interest is focused primarily on Poland, Hungary, Slovakia and Germany (UNHCR still does not confirm information about Germany, 11 March 2022).

Figure 4. Search queries in Ukrainian from Ukraine “Польща”, “Німеччина” “Угорщина”, “Молдова” “Словаччина” (Poland, Germany, Hungary, Moldova, Slovakia) (7 January 2021 – 3 March 2022)



Source: Jurić, T (2022d) Big (Crisis) Data in Refugee and Migration Studies – Case Study of Ukrainian Refugees, Comparative Southeast European Studies (forthcoming)

Figure 4 shows that the most searched countries from Ukraine in Ukrainian are Poland and Hungary. As already mentioned, the citizens of Ukraine use Russian during Internet searches more than Ukrainian. This could be explained, as mentioned, that the citizens of Ukraine probably expect more information in Russian, which is known as a world language, but also that the vast majority of Ukrainians learn and use Russian (cf. Jurić 2022d).

Figure 5. Search Queries in Russian from Ukraine “Польша”, “Германия”, “Румыния”, “Венгрия”, “Молдова” (Poland, Germany, Hungary, Moldova, Slovakia) (7 January 2021 – 3 March 2022)

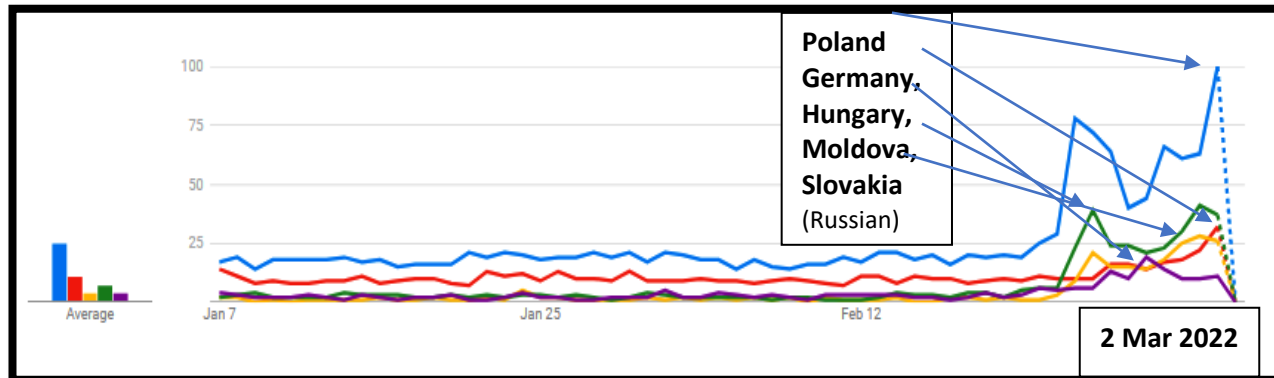
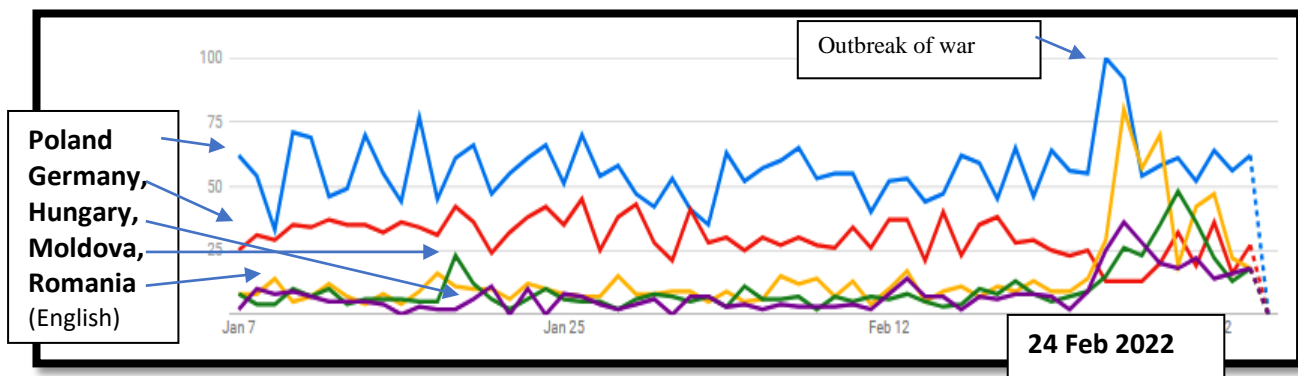


Figure 5 shows that Poland and Moldova are the most searched countries in Ukraine in Russian. The following Figure 6 shows that *Poland* and *Romania* are the most searched countries in English.

Figure 6. Search Queries in English Poland, Germany, Hungary, Moldova, Slovakia from Ukraine (7 January 2021 – 3 March 2022)

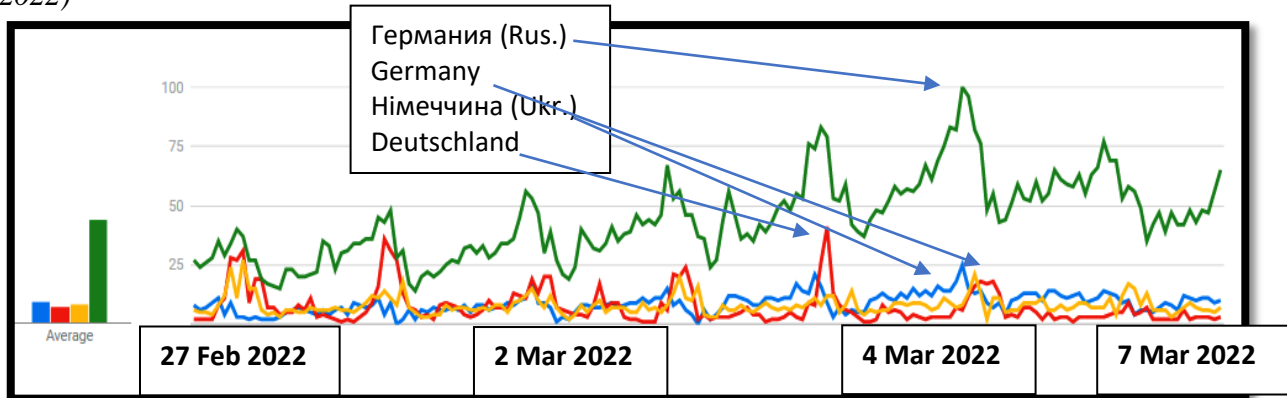


Figures 4, 5 and 6 show the increase in Internet searches for all neighbouring countries in Russian, Ukrainian and English; all three languages show that the search interest for Poland is the highest. Other languages, including German, are not significantly represented.

Our basic hypothesis is that one-third of all refugees from Ukraine will emigrate further to Germany after arriving in the original EU countries. As we stated in the introduction, this assumption is based on the movement of refugees during the Homeland War in Croatia, when it turned out that most Croatian citizens, after fleeing to nearby countries, Hungary and Slovenia, continued to Germany. Undoubtedly, the economic power of the receiving country and the assessment of when the war could stop also play an important role in this prediction. The data we received with the GT application undoubtedly show a high increase in interest in Germany (Figure 7), but also a set of inquiries related to the so-called “German way of life”, jobs opportunities, children’s enrolment in school

and other indicators that undoubtedly indicate the intention of refugees to stay longer in Germany.

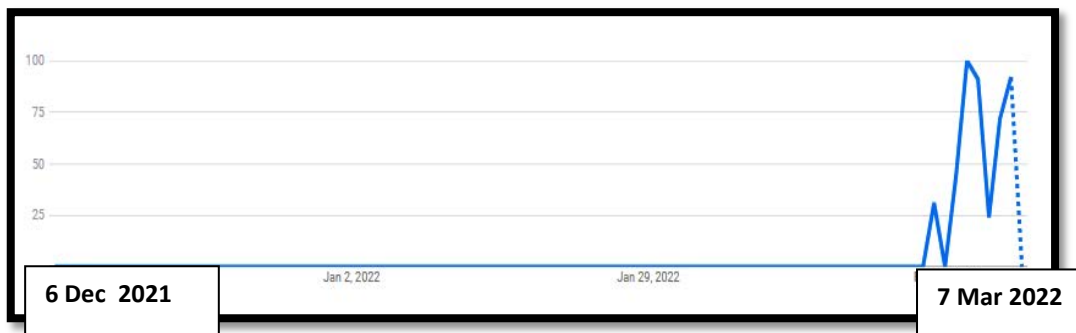
Figure 7. Interest in the Search Term “Germany” in Ukraine in English, German, Russian and Ukrainian since the Outbreak of War (28 February to 06 March 2022)



Source: Jurić, T (2022d)

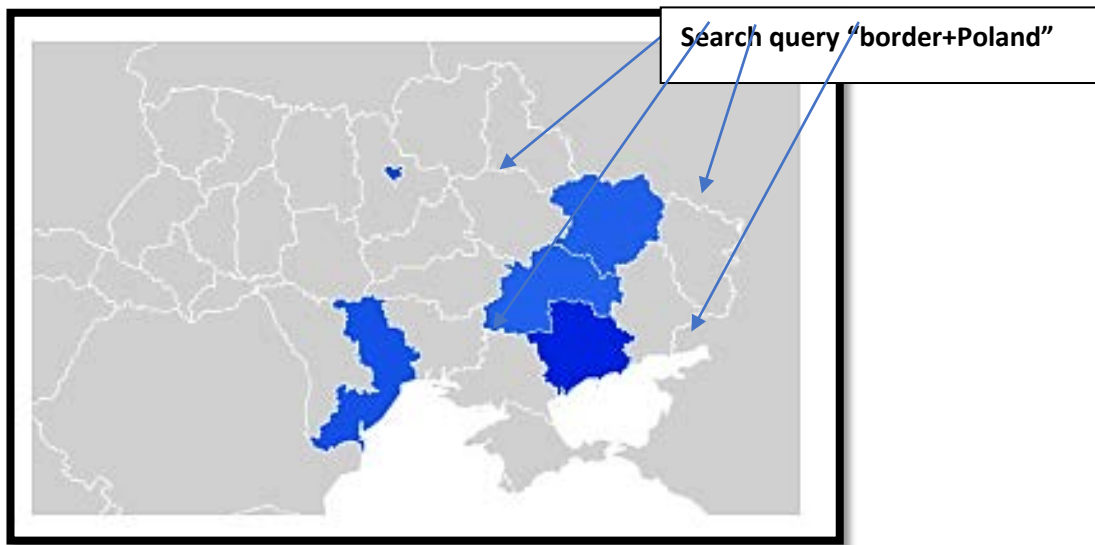
Further confirmation of this hypothesis is the increase in specific searches that undoubtedly reveal the intention to move or flee to Germany (Figure 8). One of the most common searches is just focused on the question “Германия принимает беженцев из Украины” (Does Germany accepts refugees from Ukraine?).

Figure 8. Increase in Search Query in Ukraine “Германия принимает беженцев из Украины” (Does Germany Accepts Refugees from Ukraine) (6 December 2021 - 08 March 2022)



This specific search also correlates with the regions most affected by the war, such as Donetsk, Luhansk, Kharkiv, Odesa and Kyiv (Figure 9).

Figure 9. Correlation of Google Search “border+Poland” with the Regions in Ukraine Most Affected by the War, such as Donetsk, Lugansk, Harkiv, Odesa and Kyiv (6 March 2022)



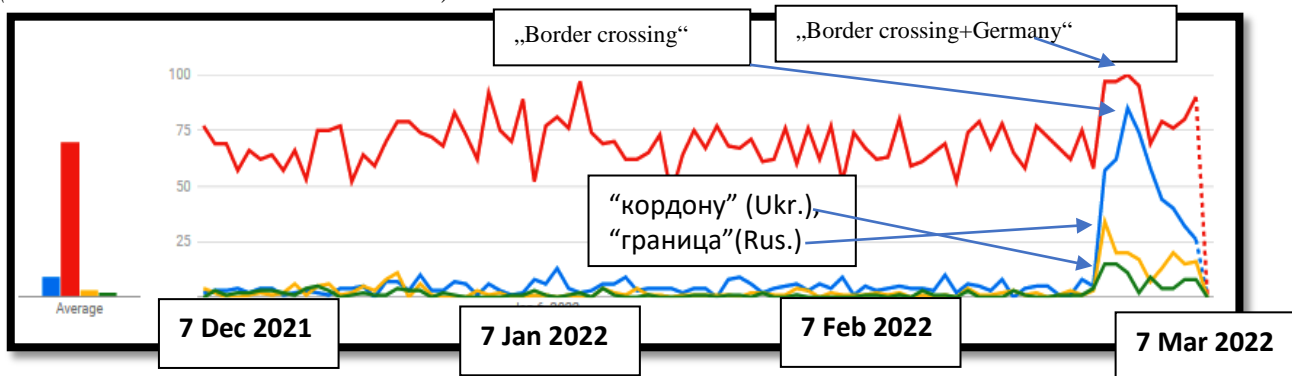
As time goes on, refugees usually realise that the war and its consequences will last longer than they expected. When they notice this, they continue their emigration to a country that offers the highest economic and financial security. In this case, as in the case of the Homeland War in Croatia, this is Germany. This assumption confirms the high growth of interest in Poland for the search query “Germany” (Figure 10).

Figure 10. The Search Term “Німеччина, Германия” (Germany) in Poland (from 11 July 2021 to 08 March 2022)



The most frequent search in Poland since the outbreak of the Russian-Ukrainian war is “Border crossing+Germany” (Figure 11) (cf. Jurić 2022d).

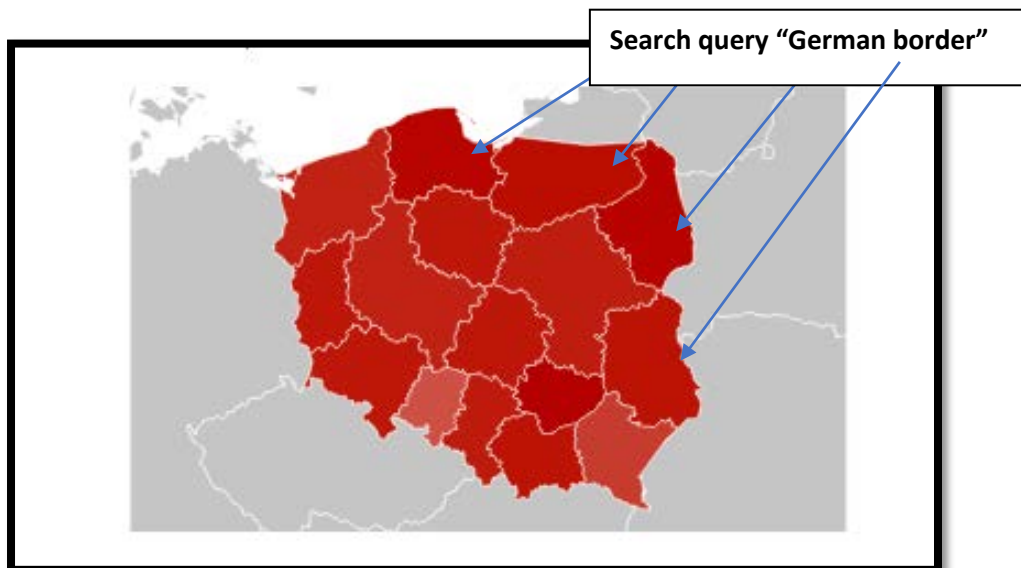
Figure 11. A Most Common Internet Search in Poland: “Border crossing+Germany”, “border crossing”, “кордону” (Ukr. Border), “граница” (Russ. Border) (7 December 2021 to 7 March 2022)



Source: Jurić, T. (2022d)

According to the German Federal Ministry of the Interior, 37,786 war refugees from Ukraine were registered in Germany by midday on 6 March 2022 (BMI 2022). The search for terms related to crossing the border into Germany in Poland is growing precisely in the regions located near Ukraine (Figure 12).

Figure 12. Regions in Poland Where Interest for “the German border” is Growing (6 March 2022)



Note: Dark red colour indicates an increase in the frequency of search keywords.

According to Düvell (2022), there are currently up to 24 million people affected by the war. Half of the population was displaced in the Crimea-Donbass conflict. According to the UNHCR, 2.3 million citizens were displaced in Ukraine during the Crimean crisis. Of that, 1.5 million escaped from Crimea (800,000 were of Russian ethnicity) (UNCHR nd). If half of the affected population flees like the

Crimean occupation in 2014, there will be up to 12 million Ukrainian refugees (Düvell 2022). That this hypothesis could be correct, we see when we compare the search query “Германия принимает беженцев из Украины” (Does Germany accepts refugees from Ukraine) during the Annexation of Crimea in 2014 with the current crisis.

Figure 13. Comparison of the Search Queries “Германия принимает беженцев из Украины” (Germany Accepts Refugees from Ukraine) During the Annexation of Crimea in 2014 with the Current Crisis in 2022

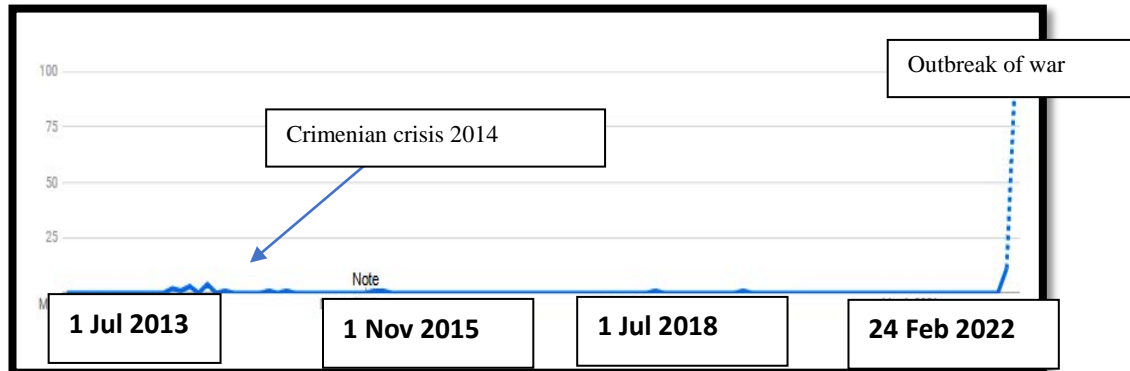
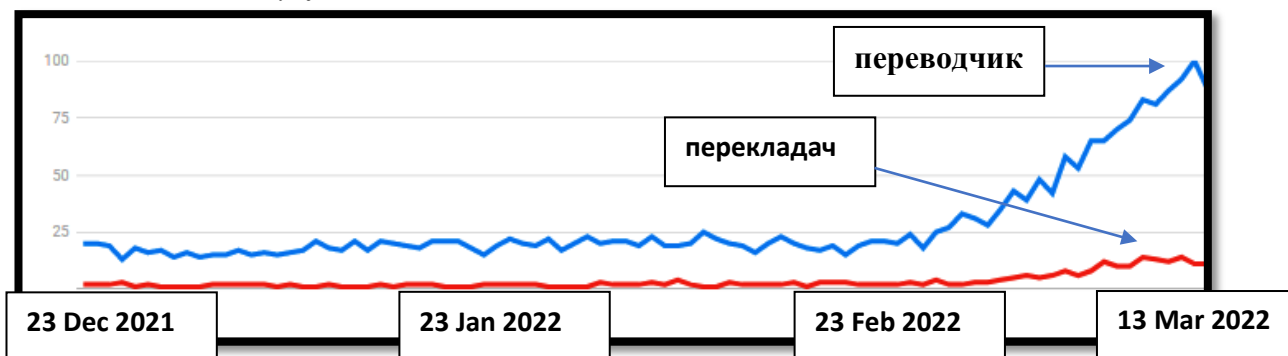


Figure 13 shows that during the peak of the Crimean crisis in 2014, the search index for query “Германия принимает беженцев из Украины” (Germany accepts refugees from Ukraine) was 4, and in 2022 it is 12, that is, the interest in fleeing the country is now three times higher.

Further keywords we tested in Germany were *Ukrainian-German translator* and *English-Ukrainian translator* (Figure 14). At 13 March 2022 all queries shows rapid increasing “точный переводчик з української на німецьку” (+250%), “перекладач з англійської на українську” (+190%), “українсько-німецький перекладач” (+150%) compared to a month ago (Jurić 2022d).

Figure 14. Querie “перекладач” (Ukr. translator) and “переводчик” (Rus. translator) in Germany (from 23 December 2021, to 13 March 2022)



Source: Jurić (2022d).

In further proceedings to standardise the data, we requested the data from 01 February to 11 March 2022, and divided the keyword frequency for the most

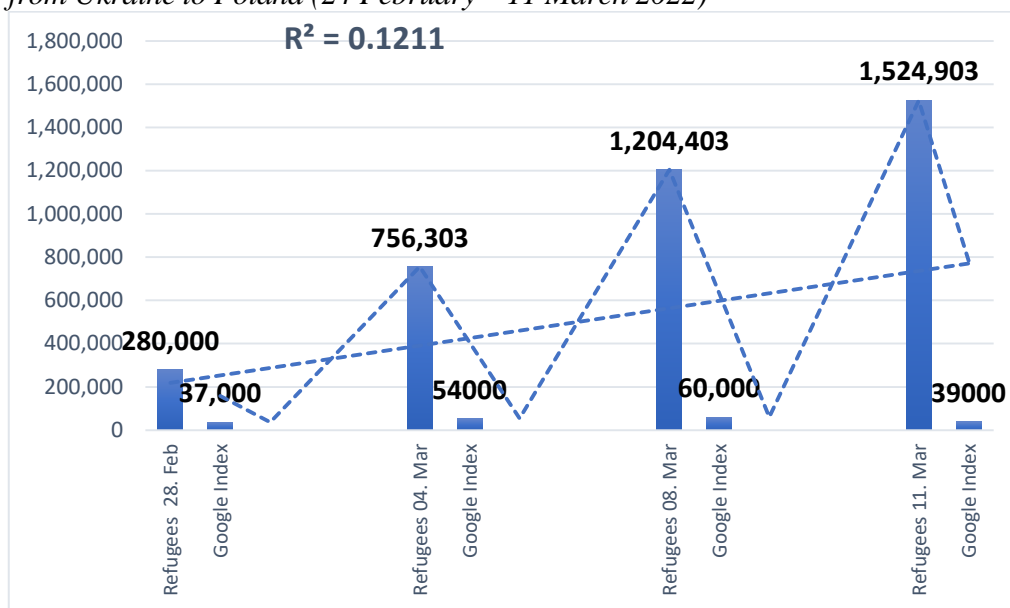
searched terms “граница” and “кордону” (border) and compared this search index with later available official statistics from UNHCR to prove the signification of results (Table 3) (cf. Jurić 2022d).

Table 3. *Refugees from Ukraine in Poland, Hungary and West European Countries*

Datum	Poland	Hungary	Other European countries (West)
11. Mar 2022	1,524,903	225,046	282,497
08. Mar 2022	1,204,403	191,348	210,239
04. Mar 2022	756,303	157,004	133,876
28. Feb 2022	280.000	85.000	34,600
Total: 2,504,893 (Data as 11 March 2022)			

Source: UNHCR, <https://data2.unhcr.org/en/situations/ukraine> and UNHCR, Ukraine Situation Regional Refugee Response Plan, <https://data2.unhcr.org/en/situations/ukraine>, authors creation.

Figure 15. *Correlation Between Google Search Index for Query “граница” (border) in Russian and the UNHCR Statistics for Externally Displaced persons from Ukraine to Poland (24 February – 11 March 2022)*



Source: Jurić (2022d).

Figure 15 shows that the increase in Google search for the query “граница” (border) in Russian correlates with the increase of externally displaced persons from Ukraine to Poland (Feb 24 – Mar 11, 2022). R2 is 0.1211 and shows a positive correlation. A p-value is statistically significant (Jurić 2022d).

Figure 16. Correlation Between Google Search Index for Query “кордону” (Border) in Ukrainian and the UNHCR Statistics for Externally Displaced Persons from Ukraine to Poland (24 February – 11 March 2022)

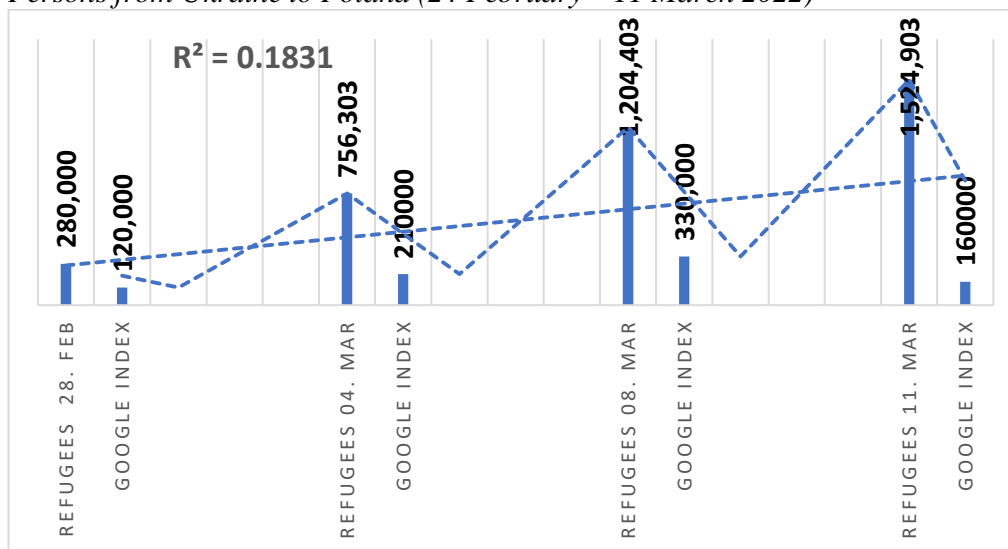


Figure 16 shows that there is a positive correlation as in the case of Google searches in Russian. The increase in Google search for the query “кордону” (border) in Ukrainian correlates with the rise of externally displaced persons from Ukraine to Poland (24 February – 11 March 2022). R^2 is 0.1831 and shows a positive correlation. A p-value is not statistically significant (Table 4).

Table 4. Correlation Analysis ANOVA

Regression Statistics	
Multiple R	0.417921
R Square	0.174658
Adjusted R Square	-0.23801
Standard Error	603084.3
Observations	4

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	1.54E+11	1.54E+11	0.423237	0.582079
Residual	2	7.27E+11	3.64E+11		
Total	3	8.81E+11			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	431691.1	839512.2	0.514217	0.658282	-3180438	4043820	3180438	4043820
X Variable 1	2.486396	3.821891	0.650567	0.582079	-13.9579	18.93067	13.9579	18.93067

Our previous studies (Jurić 2021a, Jurić 2022c) have shown that, in contrast, when the Google search decreases, there will be a decrease in externally displaced

persons from Ukraine. There is a positive linear association between the Google index and data from official statistics, UNHCR, enabling estimates and forecasting for the future.

Discussion

All tested migration-related search queries (20), which indicate emigration planning, show a positive linear association between the Google index and data from official statistics, UNHCR; $R^2 = 0.1211$ for queries in Russian and $R^2 = 0.1831$ for queries in Ukrainian (cf. Jurić 2022d).

The increase in Google search is correlated with the rise in the number of refugees in the EU. In contrast, according to our previous studies, the decrease in Google search will show a decrease in externally displaced persons.

Our testing shows an increase in Internet searches for all neighbouring countries in Russian, Ukrainian and English; all three languages show that the interest in *Poland* is the highest. It should be emphasised that when refugees arrive in nearby countries, the search for terms related to *Germany*, such as “crossing the border + Germany”, etc., is proliferating. This result confirms our hypothesis that up to one-third of all refugees will cross into Germany.

Regarding the estimated number of expected refugees, we compared the Crimean crisis in 2014 and the current crisis. During the peak of the Crimean crisis in 2014, the search index for the query “Германия принимает беженцев из Украины” (Germany accepts refugees from Ukraine) was 4, and in 2022 it is 12, that is, the interest in fleeing the country is now three times higher. Therefore, according to Big Data insights, it is to expect 5.9 Million refugees until July 2022.

The presented method contributes in a way that proves the feasibility of predicting further refugees waves from Ukraine, which allows reliable forecasts for the future and helps prepare humanitarian aid and civil infrastructure. This procedure also presents a new methodological approach to how data obtained through GT can be standardised for comparison with official UNHCR databases. The insights are particularly relevant for EU policymakers and can help governments to design appropriate strategies and prepare and better respond to this crisis in the future. Namely, even if the war ends quickly, Croatia’s previous experience with the war (Radelić et al. 2006) shows that the crisis in Ukraine will continue for many years to come, and it is to be expected that the abolition of military service will lead to the emigration of large numbers of men from the country.

This study has important limitations. It has a descriptive character and is limited by the lack of data due to the actuality of the observed phenomenon. Therefore, many more other studies are needed to perfect this method. The exponential growth in the first phase of the war needs to be modelled via the second phase to detrend the time series. The whole process should be checked with significance tests to make conclusive statements about the observed quantities.

Conclusion

This study was created due to the need to predict the migration flows of refugees from Ukraine to the EU and help crisis managers to prepare the humanitarian infrastructure and conditions for eventual integration. The usefulness and the main advantage of this approach is the timely identification of external migrations from Ukraine, which can be used to model projections and predict future trends.

Our approach to Big Data and analysis of digital traces showed that Ukrainian refugees most often searched for the term “border”. The data generated by our approach regarding the direction of refugee flows from Ukraine correlates with official UNHCR data, but there were available one to two months later after Big Data insights. Furthermore, our insights showed timely that 27% more refugees are expected than the UNHCR predicted in early stage of crisis. Here presented method showed that mid 2022 it is to expect 5.9 Million refugees and mid 2023 Germany can expect 1.5 Million Ukrainian refugees.

According to geolocations, the tested crosschecks of migration-related searches correspond to the subsequently published UNHCR data. Big (Crisis) Data show that the main regions of emigration from Ukraine are the provinces directly affected by the war (as expected). However, Ukrainian refugees will not necessarily stay in the countries of first immigration, and almost one fourth will continue their journey to Germany.

This study showed that Big (Crisis) Data has many advantages, but there are still significant open methodological issues. At the same time, more and more studies argue that samples obtained from Big Data do not significantly differ from samples obtained from more traditional recruitment and sampling techniques (Kalimeri et al. 2020). Therefore, Big Data holds enormous potential for humanitarian organisations, governments and crisis managers to improve their early warning systems, especially now in the Ukrainian refugee crisis.

Compliance with Ethical Standards

In this work, we use only anonymous, aggregate data. All data are collected following the applicable GDPR and ethical principles of personal data handling. This project’s database retains no information about the identity, IP address, or specific physical location of any user.

References

- Abcnews (2022) *Europe welcomes Ukrainian refugees — Others, less so*. Retrieved from: <https://abcnews.go.com/International/wireStory/europe-welcomes-ukrainian-refugees-83153021> [Accessed 30 March 2022]
- Agrawal A, Sahdev R, Davoudi H, Khonsari F, McGrath S (2016) *Detecting the magnitude of events from news articles*. Retrieved from: <https://ieeexplore.ieee.org/document/7817051>. [Accessed 24 February 2022]

- Almustafa M (2021) *Reframing refugee crisis: a “European crisis of migration” or a “crisis of protection”?* Available at: <https://doi.org/10.1177/2399654421989705>.
- Arbeitsagentur (2020) *Gemeldete Arbeitsstellen nach Berufen (Engpassanalyse)* (Registered jobs by occupation). Retrieved from: <https://statistik.arbeitsagentur.de/Statistikdaten/>. [Accessed 24 February 2022]
- Bergtora SK, Garnier A (4 Mar 2022) *Forced displacement from Ukraine: notes on humanitarian protection and durable solutions*. Retrieved from: <https://reliefweb.int/report/ukraine/forced-displacement-ukraine-notes-humanitarian-protection-and-durable-solutions>. [Accessed 7 March 2022]
- Berkay M (2020) *Sharing the burden: revisiting the EU-Turkey migration deal*. Retrieved from: <https://www.crisisgroup.org/europe-central-asia/western-europemediterranean/turkey/sharing-burden-revisiting-eu-turkey-migration-deal>. [Accessed 25 February 2022]
- Bertelsmann Stiftung (2015) *Zuwanderungsbedarf aus Drittstaaten in Deutschland bis 2050*. (Need for immigration from third countries in Germany by 2050). Retrieved from: <https://bit.ly/3cyw0u1>. [24 February 2022]
- Bircan T, Korkmaz EE (2021) Big data for whose sake? Governing migration through artificial intelligence. *Humanities Social Sciences Communications* 8(Oct): 241.
- BMI (2015) *A test for our country*. Retrieved from: <https://www.bmi.bund.de/SharedDocs/kurzmeldungen/EN/2015/09/measures-on-asylum-and-refugee-policy.html>. [Accessed February 2022]
- BMI (2022) *Befragung ukrainischer Kriegsflüchtlinge*. (Interrogation of Ukrainian war refugees). Retrieved from: <https://bit.ly/3Q8krZw>. [Accessed 6 March 2022]
- Böhme MH, Gröger A, Stöhr T (2020) Searching for a better life: Predicting international migration with online search keywords. *Journal of Development Economics* 142(Jan): 102347.
- Cesare N, Lee H, McCormick T, Spiro E, Zagheni E (2018) Promises and pitfalls of using digital traces for demographic research. *Demography* 55(5): 1979–1999.
- Charron A (2020) Somehow, we cannot accept it: drivers of internal displacement from Crimea and the forced/voluntary migration binary. *Europe-Asia Studies* 72(3): 432–454.
- Choi H, Varian H (2012) Predicting the present with google trends. *Econ Record* 88(s1): 2–9
- Connor P (2017) *The digital footprint of Europe’s refugees*. Methodology. Pew Research Center. Retrieved from: <https://www.pewglobal.org/2017/06/08/online-searches-eu-refugees-methodology/>. [Accessed 28 February 2022]
- Curry T, Croitoru A, Crooks A, Stefanidis A (2018) Exodus 2.0: crowdsourcing geographical and social trails of mass migration. *Journal of Geographical Systems* 21(Sep): 161–187.
- Deutsche Bahn (2022) *“Helpukraine” ticket allows refugees to travel to any station in Germany free of charge*. Available at: https://www.deutschebahn.com/en/presse/press_releases/-helpukraine-ticket-allows-refugees-to-travel-to-any-station-in-Germany-free-of-charge-7315140.
- Deutsche Welle (2022) *Ukrajinske izbjeglice kao konkurencija useljenicima sa zapadnog Balkana?* (Ukrainian refugees as competition to immigrants from the Western Balkans). Retrieved from: <https://www.dw.com/bs/ukrajinske-izbjeglice-kao-konkurencija-useljenicima-sa-zapadnog-balkana/a-61045150>. [Accessed 1 March 2022]
- Dubois A, Zagheni E, Garimella K, Weber I (2018) Studying migrant assimilation through Facebook interests. *Lecture Notes in Computer Science* 11186(Sep): 51–60.

- Düvell F (2022) *Krieg in der Ukraine: Wie ist die Flüchtlings-Situation?* (War in Ukraine: what is the refugee situation like? Universität Osnabrück, Mediendienst Integration. Retrieved from: <https://mediendienst-integration.de/artikel/krieg-in-der-ukraine-wie-ist-die-fluechtlings-situation.html>. [Accessed 7 March 2022]
- EUROPOL (2018) *Two years of EMSC: activity report January 2017 – January 2018*. European Migration Smuggling Centre, EUROPOL. Retrieved from: https://www.europol.europa.eu/cms/sites/default/files/documents/two_years_of_emsc_report.pdf. [Accessed 10 March 2022]
- Gabrilovich E (2020) *Using symptoms search trends to inform COVID-19 research*. Google Health. Retrieved from: <https://blog.google/technology/health/using-symptoms-search-trends-inform>. [Accessed 28 February 2022]
- Hawelka B, Sitko I, Beinat E, Sobolevsky S, Ratti PKC (2014) Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science* 41(3): 260–267.
- Ibáñez Sales M (2021) *Big data at the crossroads: seizing the potential of Big data to guide the future of EU migration policy*. Euromesco Policy brief, n. 116.
- Internet World Stats (2022) *Internet users distributions in the world - 2021*. Retrieved from: <https://www.internetworldstats.com/stats.htm>. [Accessed 2 March 2022]
- Jurić T (2021a) “*Gastarbeiter Millennials*”: exploring the past, present and future of migration from Southeast Europe to Germany and Austria with approaches to classical, historical and digital demography. Hamburg: Verlag Dr. Kovač.
- Jurić T (2021b) Google trends as a method to predict new COVID-19 cases and socio-psychological consequences of the pandemic. *Athens Journal of Mediterranean Studies* 8(1): 67–92.
- Jurić T (2021c) Medical brain drain from Southeastern Europe: using digital demography to forecast health worker emigration. *JMIRx Med* 2(4): e30831.
- Jurić T (2022a) Forecasting migration and integration trends using digital demography – A case study of emigration flows from Croatia to Austria and Germany. *Comparative Southeast European Studies* 70(1): 1–28.
- Jurić T (2022b) Facebook i Google kao empirijska osnova za razvoj metode digitalnog praćenja vanjskih migracija hrvatskih građana (Facebook and Google as an empirical basis for the development of a method of digital monitoring of external migration of Croatian citizens). *Ekonomski pregled* 73(2).
- Jurić T (2022c) Google search analysis in interdisciplinary research (case studies: COVID-19, birth rate and tourism demand), 4. In *Međunarodna naučna konferencija o digitalnoj ekonomiji DIEC 2021 - Zbornik radova*, 5–25. Tuzla, 2022.
- Jurić T (2022d) Big (crisis) data in refugee and migration studies – Case study of Ukrainian refugees. *Comparative Southeast European Studies* (forthcoming).
- Kalimeri KA, Beiro MG, Bonanomi A, Rosina A, Cattuto C (2020) Traditional versus Facebook-based surveys: evaluation of biases in self-reported demographic and psychometric information. *Demographic Research* 42(5): 133–148.
- Kuznetsova I, Mikheieva O (2020) Forced displacement from Ukraine’s war-torn territories: intersectionality and power geometry. *Nationalities Papers* 48(4): 690–706.
- Magoscy R (1996) *A history of Ukraine*. Toronto: University of Toronto Press.
- Önder I (2017) Forecasting tourism demand with Google trends: accuracy comparison of countries versus cities. *International Journal of Tourism Research* 19(6): 648–660.
- Radelić Z, Marijan DN, Barić A, Bing, Živić D (2006) *Stvaranje hrvatske države i Domovinski rat*. (The creation of the Croatian state and the Homeland War). Školska knjiga Zagreb 2006.

- Scrinic A (2014) Humanitarian aid and political aims in Eastern Ukraine: Russian involvement and European response. *Eastern Journal of European Studies* 5(2): 77–88.
- Sereda V (2020) Social distancing’ and hierarchies of belonging: the case of displaced population from Donbas and Crimea. *Europe-Asia Studies* 72(3): 404–431.
- Singh L, Wahedi L, Wang Y, Wei Y, Kirov C, Martin S, et al. (2019) Blending noisy social media signals with traditional movement variables to predict forced migration. In *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1975–1983.
- Spyratos S, Vespe M, Natale F, Weber I, Zagheni E, Rango M (2018) *Migration data using social media: a European perspective*. JRC Technical Reports European Commission. Luxembourg: Publications Office of the European Union.
- State B, Rodriguez M, Helbing D, Zagheni E (2014) Migration of professionals to the U.S. evidence from LinkedIn data. *Lecture Notes in Computer Science* 8851: 531–43.
- The Guardian (2019) “A whole generation has gone”: Ukrainians seek a better life in Poland. Retrieved from: <https://www.theguardian.com/world/2019/apr/18/whole-generation-has-gone-ukrainian-seek-better-life-poland-elect-president>. [Accessed on 14.04.2022]
- The Guardian (2022, March 5) *Ukraine has fastest-growing refugee crisis since second world war, says UN*. Retrieved from: <https://www.theguardian.com/world/2022/mar/06/ukraine-fastest-growing-refugee-crisis-since-second-world-war>. [Accessed 10 April 2022]
- UNHCR (2016) *Global trends. Forced displacements in 2015*. <https://www.unhcr.org/576408cd7.pdf> (accessed 1 April 2022). [Accessed 8 March 2022]
- UNHCR (2017) *From a refugee perspective: discourse of Arabic speaking and Afghan refugees and migrants on social media from March to December 2016*. Retrieved from: <https://www.unhcr.org/publications/brochures/5909af4d4/from-a-refugee-perspective.html>; <https://reliefweb.int/sites/reliefweb.int/files/resources/58018.pdf>. [Accessed 10 March 2022]
- UNHCR (2022a) *Ukraine RRR*. Retrieved from: <https://www.unhcr.org/ukraine-emergency.html> [Accessed 4 March 2022].
- UNHCR (2022b) *Internally displaced persons (IDP)*. Retrieved from: <https://www.unhcr.org/ua/en/internally-displaced-persons>. [Accessed 29 February 2022]
- UNHCR (2022c). *Ukraine situation regional refugee response plan*. Retrieved from: <https://bit.ly/3bBmIxp>. [Accessed 8 March 2022]
- UNHCR Blog, Pellandra A, Henningsen G (2022) *Predicting refugee flows with big data: a new opportunity or a pipe dream?* Retrieved from: <https://www.unhcr.org/blogs/predicting-refugee-flows-with-big-data-a-new-opportunity-or-a-pipe-dream/>.
- UNHCR Global Data Service (2021) *Big (crisis) data for predictive models: a literature review*. UNHCR Global Data Service.
- United Nations (2014) *The data revolution for human development*. Retrieved from: <http://hdr.undp.org/en/content/data-revolution-human-development>. [Accessed 28 February 2022]
- United Nations (2019) *Report of the global working group on big data for official statistics*. The Economic Social Council of the United Nations, UN Global Working Group on Big Data.
- Visitukraine (2022, March 5) *Refugees from Ukraine received the right to live in the EU for 3 years*. Retrieved from: <https://visitukraine.today/blog/154/refugees-from-ukraine-received-the-right-to-live-in-the-eu-for-3-years>. [Accessed 11 March 2022]

- Wang R, Wang W, daSilva A, Huckins JF, Kelley WM, Heatherton TF, et al. (2018) Tracking depression dynamics in college students using mobile phone and wearable sensing. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2(1): 43.
- Wanner P (2020) How well can we estimate immigration trends using Google data? *Quality and Quantity* 55(Oct): 1181–202.
- Wilde J, Chen W, Lohmann S (2020) *COVID-19 and the future of US fertility: what can we learn from Google?* IZA Discussion Papers 13776.
- Wladyka D (2017) Queries to Google Search as predictors of migration flows from Latin America to Spain. *Journal of Population and Social Studies* 25(4): 312–327.
- Zagheni E, Garimella VRK, Weber I, State B (2014) Inferring international and internal migration patterns from Twitter data. In *WWW '14 Companion: Proceedings of the 23rd International Conference on World Wide Web*, 439–444.
- Zagheni E, Weber I (2015) Demographic research with non-representative Internet data. *International Journal of Manpower* 43(9).
- Zagheni E, Weber I, Gummadi K (2017) Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review* 43(4): 721–734.

