*Joop Adema* and *Maitreyee Guha*

# Following the Online Trail of Ukrainian Refugees through Google Trends

## ABSTRACT

Online search behavior offers unique insights about human behavior, not only giving insights about people's state of mind or needs, but also giving insights about their location and mobility intentions. In times of a refugee crisis, such information is important for human well-being in the short run but can also guide integration policies in the long run. Google Trends enables tracking the location, needs, and plans of Ukrainian refugees that left their home country following the invasion of the Russian army. As official statistics and surveys lack granularity and are published with delay, online search data is a valuable addition in humanitarian crises. In this article, we document several use cases of online search data to the current refugee crisis.

Russia invaded Ukraine on February 24, 2022. In almost four months, more than 5.2 million individuals registered for Temporary Protection Schemes in the EU (UNHCR, updated June 21, 2022).[1] This constitutes the largest wave of refugees in Europe in decades. In comparison, during the Syrian refugee crisis, around 1.1 million refugees reached the EU. Although the refugees have predominantly migrated to developed countries, there are nevertheless policy-relevant knowledge gaps related to humanitarian needs, housing, work, and return in-tentions. Several surveys have aimed to fill these gaps.[2] These surveys found that typically over 80 per-

cent of adult refugees are women, the average age of refugees is below 40, and a considerable share of refugees are housed by family members or friends. An online survey administered in April in Poland and Germany found that many refugees residing in Poland did not officially register, suggesting that official statistics may underestimate the number of refugees present in Poland (GESIS 2022). A survey conducted in Moldova in March and April showed that most refugees intend to move to third countries (IOM 2022). In contrast, a survey in Germany showed that only a small portion of migrants want to go to another country and that one-third of respondents want to return to Ukraine soon (BMI 2022). However, these surveys rely on small sample sizes, require trained interviewers or recruitment through social media, and do not enable real-time monitoring. As more than 70 percent of Ukrainians use the internet, their digital footprint can be used as an additional source of information.[3]

## MIGRATION AND ONLINE SEARCH

Google Trends enables tracking the search intensity of a search term, relative to the total search volume in a geographic area. This Google Trends Index (GTI) can be disaggregated at the subnational level, which allows comparisons between regions and over time for common search terms.[4] In Ukraine, Google has a market share of 93 percent on the market for search engines.[5] Therefore, the GTI is representative of Ukrainians' online search interests. As users from different countries use different languages and have different interests, search behavior can be used to track migrants. In this article, we use the GTI of specific search terms in Ukrainian and Russian (the two most commonly spoken languages in Ukraine). In the following, we refer to GTI as the raw index obtained from Google Trends.[6] As we are often

**Joop Adema**

is Junior Economist and Doctoral Student at the ifo Center for International Institutional Comparisons and Migration Research.

**Maitreyee Guha**

is a master's student at the Paris School of Economics and currently an intern at the ifo Center for International Institutional Comparisons and Migration Research.

[1]   https://data.unhcr.org/en/situations/ukraine (updated June 21, 2022).
[2]   Example of surveys among Ukrainian refugees: an online survey among Ukrainian refugees in Germany (https://www.bmi.bund.de/ SharedDocs/kurzmeldungen/EN/2022/04/survey-ukraine.html), a survey in Moldova (https://eca.unwomen.org/en/digital-library/publications/2022/04/updated-displacement-survey-ukrainian-refugees-and-third-country-nationals-0), and an online survey in Poland and Germany (https://blogs.lse.ac.uk/europpblog/2022/05/12/preliminary-findings-from-an-online-survey-of-ukrainian-refugees-in-germany-and-poland/).
[3] https://datareportal.com/reports/digital-2022-ukraine.
[4] The GTI of a search term includes all queries containing the search term.
[5] https://gs.statcounter.com/search-engine-market-share/all/ukraine.
[6]   https://trends.google.com/trends/?geo=DE.

interested in the relative intensity of search terms among a subpopulation, we refer to this as the *relative search intensity*.

Well-known use cases of online search data are surveillance of flu outbreaks (Carneiro and Mylonakis 2009) and prediction of private consumption patterns (Vosen and Schmidt 2011). Related to international migration, Böhme et al. (2020) showed that migration intentions can be predicted using migration-related search terms and that these are also predictive of subsequent migration flows. Furthermore, Wanner (2021) showed that online search behavior is predictive of migration from Spain and Italy to Switzerland. Santamaria (2021) used online search behavior to infer the location of Venezuelan refugees in Colombia to study the labor market effects of immigration. The only application of online search behavior to the current refugee crisis shows that search for migration-related queries rose after the Russian invasion and correlates to subsequent migration (Jurić 2022). In this article, we complement these studies by showing the usefulness of online search data in the current crisis for a variety of cases.
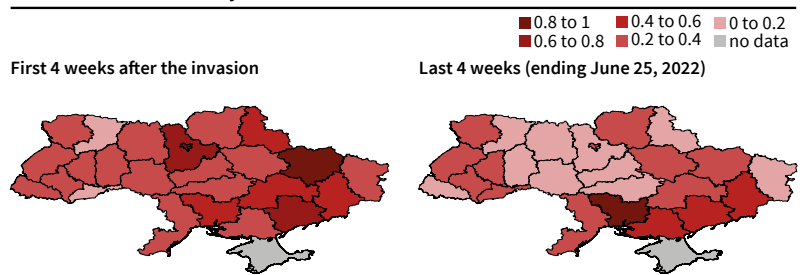
## DISPLACEMENT IN UKRAINE

As a first case, we show how urgent plans to move away can be mapped with online search behavior. Figure 1 shows the relative search intensity of the word "evacuation"[7] across space, for the first four weeks of the war (left panel), and the past four weeks (right panel). Relative search intensity is obtained by scaling both panels and is set to 1 for the region with the highest intensity. Relative search intensity in the whole of Ukraine in the first four weeks of the war was nine times higher than in the past four weeks, suggesting that most individuals looking for evacuation managed to do so. Initially, the interest in evacuation was concentrated in the east and the north (particularly Kharkov, Kyiv, and the region surrounding Kyiv). More recently, the interest in evacuation shifted, as the conflict concentrated in the eastern and southern parts of the country. Especially in the Mykolaiv region, the interest in evacuation is particularly high.[8] This shows that online search behavior has the potential to identify populations in distress, planning to move away. As such information is available at high frequency, it could guide humanitarian assistance.

## TRACKING REFUGEES

In this section, we show how refugees can be tracked across countries. The leading country-level (and for some countries subnational-level) data collection ef-

---

[7] We present the sum of the Google Trend Indices for the Ukrainian and Russian words in order to be able to compare areas with different proportion of Ukrainian and Russian speakers.
[8] In the week of June 26, 2022, relative search interest for evacuation was highest in the partially occupied regions of Kherson and Donetsk, as well as Kharkiv, after renewed shelling.

Figure 1

**Relative Search Intensity for "Evacuation" across Ukraine**



■ 0.8 to 1 ■ 0.4 to 0.6 ■ 0 to 0.2
■ 0.6 to 0.8 ■ 0.2 to 0.4 ■ no data

**First 4 weeks after the invasion**    **Last 4 weeks (ending June 25, 2022)**

Source: Google Trends (2022).                                    © ifo Institute

fort on Ukrainian refugee stocks is the Operational Data Portal of the UNHCR.[9] To detect Ukrainian refugees with online search data, we focus on the relative prevalence of the search term for "weather" ("погода" in both Ukrainian and Russian) in various receiving countries. As online search behavior for the weather is seasonally dependent, one has to carefully account for seasonal trends. Therefore, we compare online search behavior in 2022 in the same period as in 2019. Russians and Ukrainians residing in an area before the Russian invasion allow us to benchmark how much relative search queries for "weather" are generated per individual. We predict the number of Ukrainian and Russian speakers in May 2022 in the following way. To obtain the ratio of Russian and Ukrainian speakers between May 2022 and May 2019, we divide the average Google trend index in May 2022 by the average Google trend index in May 2019. We multiply this ratio by the number of Ukrainians and Russians holding residence permits in the respective country in 2019, giving the predicted number of Ukrainian and Russian speakers present in 2022. To obtain the predicted number of Ukrainian refugees, we subtract the last available number of Ukrainian and Russian residents.
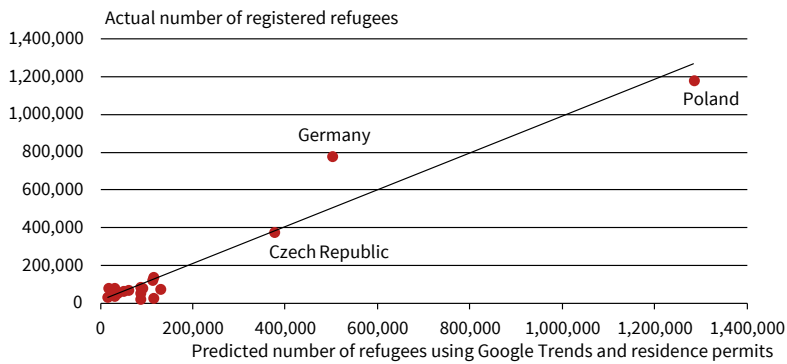
The two main caveats of this method are the inability to precisely pin down the pre-existing number of Ukrainian and Russian speakers in a region and the potentially different propensity to search for weather conditions between refugees and non-refugees. Nevertheless, we obtain a good prediction of actual numbers of registered refugees. Figure 2 shows the correlation between the predicted number of refugees and the actual numbers of registered refugees for the 18 EU countries with the largest number of registered Ukrainian refugees, as well as a diagonal line. We find a strong correspondence (slope = 0.95, R2 = 0.93) between the predicted and the actual number of refugees. However, for several smaller countries, we strongly overestimate the number of refugees. In Slovakia and Hungary, we overestimate the number of refugees by about a factor 2. This can be explained by the fact that many more refugees enter these countries from Ukraine to transit rather

---

[9] https://data.unhcr.org/en/situations/ukraine.

Figure 2

**Actual vs. Predicted Number of Refugees Based on Google Trends within Poland**



Actual number of registered refugees

Source: Google Trends (2022).

© ifo Institute

than register as refugees. The largest outlier is Finland, where searches for the weather increased by 260 percent despite officially only housing 65 percent more Russian and Ukrainian speakers (residents and registered refugees). This may be suggestive of a large inflow of Russian speakers fleeing Russia after the invasion, rather than Ukrainian refugees.[10]

Although the number of refugees registering for temporary protection schemes is well known in most countries receiving Ukrainian refugees, not all countries have information about their distribution within the country. Using the GTI for "погода" across regions and the population of Poland's 16 regions (with a population between 900,000 and 5 million), we can predict the share of total refugees per region. Figure 3 shows a scatterplot between the predicted and actual share of all refugees in each of the regions. We find a strong correspondence between predicted and actual shares (slope = 0.91, R2 = 0.92).[11] This suggests that countries – unlike Poland – that do not collect statistics about the subnational distribution of refugees can estimate it using online search behavior.

[10] https://www.bbc.com/news/world-60624500.
[11] For the data on the regional distribution of refugees in Poland, see https://dane.gov.pl/en/dataset/2715,zarejestrowane-wni-oski-o-nadanie-statusu-ukr/resource/39437/table?page=1&per_page=20&q=&sort=. For the data on the number of inhabitants per voivodship in 2020, see https://stat.gov.pl/en/topics/other-studies/cities-voivodship/.

Figure 3

**Actual vs. Predicted Number of Refugees Based on Google Trends within Poland**



Actual share of refugees as a % of all refugees in Poland
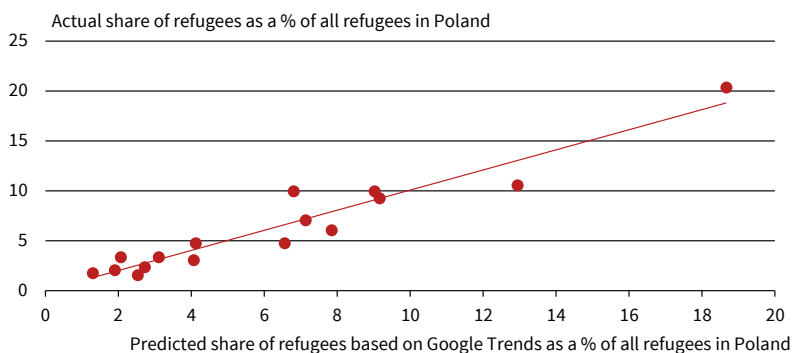
Source: Google Trends (2022).

© ifo Institute

There are drawbacks to Google Trends that limit its usability, especially for less frequently searched terms. Because of Google's policy not to provide many details on their sampling procedure and how small counts are left-censored, Google Trends provides information about relative search intensities rather than the absolute number of queries or unique individuals. To argue that despite the sampling and left censoring we can still detect relatively small refugee flows when there is well-calibrated prior data about the number of people present with similar search behavior, we turn to Croatia.

According to the UNHCR, close to 15,000 Ukrainians have requested shelter in Croatia. From official statistics, we obtain precise information on the number of nights tourists from Russia and Ukraine stayed in Croatia.[12] We infer that a single GT point in the weekly GTI series for "погода" between 2019 and 2022 corresponds to the presence of 350 people. Assuming the complete absence of Ukrainian and Russian tourists in 2022, the sharp increase at the end of March 2022 of around 35 GT points suggests the presence of about 13,000 additional Russian and Ukrainian speakers, corresponding well with the official statistics. In winter months, the GT index varies by less than 5 percentage points from week to week, suggesting a resolution of less than 2000 people, 0.05 percent of the Croatian population.

These validation exercises suggest that aggregate online search data can be powerful in cases when official statistics are unavailable. Provided that accurately calibrated data is available and that search behavior of migrants and natives is distinct, this implies that the number of migrants in a subnational region can be approximated on a weekly basis with Google Trends.

**MAPPING REFUGEES' NEEDS**

In this section, we show how refugees' online search behavior can be informative about their imminent needs. This is important from a humanitarian point of view and can guide authorities in aiding refugees with unmet needs.[13] In the following, we focus on search behavior in Ukrainian and Russian for the following terms: "work," "school," "housing," "medical doctor," "language course" (in the local language), and "refugee allowance." Comparing the prevalence of these keywords vis-à-vis may be deceiving, as there are multiple search queries using different terms to answer the same question. However, it enables comparing their relative prevalence over time and across places. To obtain a measure of the relative prevalence of these search terms among refugees, we compute the relative search intensity with respect to the search term for "weather" in Ukrainian/Russian. We use the

[12] https://www.htz.hr/sites/default/files/2020-07/HTZ percent20TUB percent20ENG_2019.pdf.
[13] UNHCR aims to monitor the needs of refugees and writes assessment reports, interviewing various stakeholders. For such a report, see https://data.unhcr.org/en/documents/details/91748.

relative search intensity of the same terms in the same time period in 2021 to correct for the contribution of Ukrainians and Russians residing in the destination already prior to the war and for term-specific seasonal patterns in relative search intensity. Furthermore, we correct for weekly weather shocks by using natives' relative search interest for weather in the local language.

Figure 4 shows the corrected weekly relative search intensity for aforementioned terms in Poland, the country with the largest inflow of Ukrainian refugees. In the first weeks after the start of the war search intensities for work and school were relatively prevalent and declined thereafter. This implies that efforts to search for employment and schooling decline after refugees begin to settle down in their host countries. In contrast, the relative search intensity for doctors started increasing several weeks after the beginning of the Russian invasion.

Figure 5 shows the relative prevalence of the items shown in Figure 4 between May 30 and June 25, 2022 for the three countries hosting most of the Ukrainian refugees: Poland, Germany, and Czech Republic. As most refugees were living in these host countries for several months by June 2022, search terms between countries are comparable. Therefore, these likely do not reflect search behavior at different stages of migration, but rather currently unmet needs and interests. Relative to the Czech Republic and Poland, Ukrainians in Germany search much more for allowances for refugees. This suggests that information provision about such allowances in Germany could be improved. Furthermore, Ukrainian refugees in Germany are most likely to search for language courses in the local language, suggesting a larger interest than the Czech Republic and Poland to learn the local language and integrate in the local society. As refugees in Germany are willing to stay for longer (BMI 2022), the relative search intensity for housing unsurprisingly exceeds search intensity for housing in the other two countries.

## RETURN AND TRANSIT INTENTIONS

Ukrainian refugees obtained the right to stay in the EU for at least one year.[14] As long as the war lasts, it is unlikely these rights will be revoked. Therefore, Ukrainians' intentions to stay are crucial for understanding whether their presence in receiving countries will be long-lasting or not. To probe the short-run intentions to return of those residing in Poland who did not return yet, we rely on search terms related to travel to Ukraine, as well as to Germany. Figure 6 shows the relative search intensity (corrected similarly to Figure 4 and 5 using auxiliary GTIs) for two different search terms in Poland: "Berlin" and "Ukrainian

[14] https://ec.europa.eu/info/strategy/priorities-2019-2024/stronger-europe-world/eu-solidarity-ukraine/eu-assistance-ukraine/information-people-fleeing-war-ukraine_en#your-rights-in-the-eu.

Figure 4

**Relative Intensities of Search Terms in Poland in 2022**
Corrected for relative intensities in 2021 and native "weather" searches
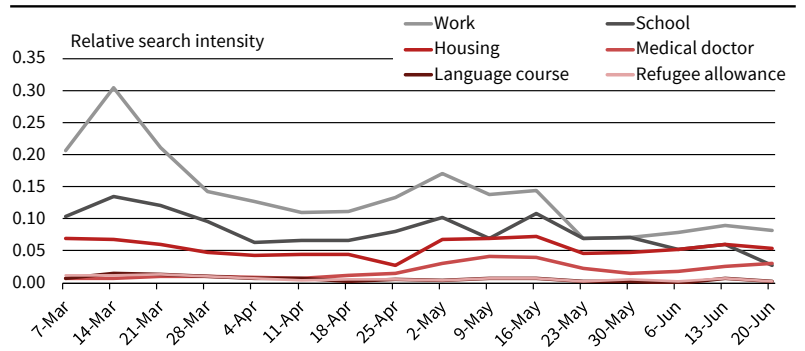


Source: Google Trends (2022). © ifo Institute

Railways." Whereas early in the refugee crisis many refugees transited to Germany, a substantial portion of refugees returned to Ukraine.[15] We observe that in May 2022, searches for "Ukrainian Railways" increased, following the retreat of Russian troops from the north of Ukraine. Search intensity for "Berlin" decreased considerably over the run of the crisis. This suggests that either refugees changed their mind and stayed in Poland, left for Berlin early on, or returned to Ukraine. Nevertheless, there is still some residual interest in Berlin, suggesting Germany could receive a continuing stream of refugees.
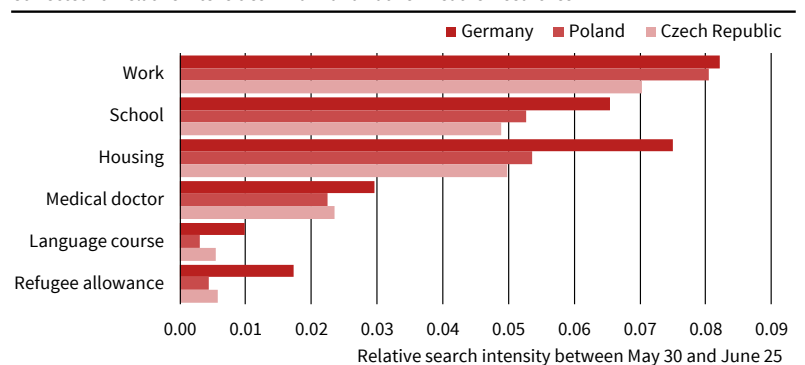
## CONCLUSION

Publicly available aggregate online search behavior can be helpful in mapping multiple aspects of a refugee crisis. Not only can online search behavior be used to identify populations under distress searching to flee – it can also be used to reasonably predict the stock of a group of people using group-specific search terms at any moment in time. Additionally, it enables monitoring of unmet needs that are particu-

[15] UNHCR collects information about how many people cross the borders to Ukraine. Although it is unknown how many individuals this concerns, more than 2.8 million border crossings into Ukraine were reported between February 24 and June 21.

Figure 5

**Relative of Intensities of Search Terms by Country in June 2022**
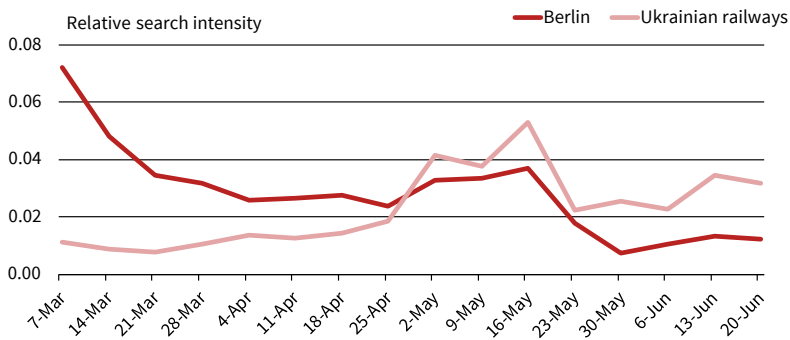Corrected for relative intensities in 2021 and native "weather" searches



Source: Google Trends (2022). © ifo Institute

Figure 6

**Relative Intensity of Transit and Return Searches in Poland in 2022**
Corrected for relative intensities in 2021 and native "weather" searches



Source: Google Trends (2022).                                                                    © ifo Institute

and can be obtained at higher frequency than administrative data and surveys. Therefore, methods based on aggregate online search data could be readily implemented by organizations such as the UNHCR or national authorities responsible for refugee crises to get a more complete picture of the situation of Ukrainian refugees.

**REFERENCES**

Böhme, M. H., A. Gröger and T. Stöhr (2020), "Searching for a Better Life: Predicting International Migration with Online Search Keywords", *Journal of Development Economics* 142, 102347.

BMI (2022), "Survey of Ukrainian War Refugees", https://www.bmi.bund. de/SharedDocs/kurzmeldungen/EN/2022/04/survey-ukraine.html.

Carneiro, H. A. and E Mylonakis (2009), "Google Trends: A Web-based Tool for Real-time Surveillance of Disease Outbreaks", *Clinical Infectious Diseases* 49, 1557–1564.

Dergiades, T., E. Mavragani and B. Pan (2018), "Google Trends and Tourists' Arrivals: Emerging Biases and Proposed Corrections", *Tourism Management* 66, 108–120.

GESIS (2022), "Geflüchtete aus der Ukraine – Erste Deskriptive Ergebnisse einer Onlinebefragung in Deutschland und Polen", https://blog. gesis.org/gefluchtete-aus-der-ukraine-erste-deskriptive-ergebnisse-einer-onlinebefragung-in-deutschland-und-polen/.

IOM (2022), "Updated Displacement Survey: Ukrainian Refugees and Third-Country Nationals", https://eca.unwomen.org/en/digital-library/publications/2022/04/updated-displacement-survey-ukrainian-refugees-and-third-country-nationals-0.

Jurić, T. (2022), "Predicting Refugee Flows from Ukraine with an Approach to Big (Crisis) Data: A New Opportunity for Refugee and Humanitarian Studies", *medRxiv preprint* no. 2022.03.15.22272428.

Santamaria, J. (2021), "When a Stranger Shall Sojourn with Thee': The Impact of the Venezuelan Exodus on Colombian Labor Markets", University of Minnesota, *Working Paper* 51422.

Vosen, S. and T. Schmidt (2011), "Forecasting Private Consumption: Survey-based Indicators vs. Google Trends", *Journal of Forecasting* 30, 565-578.

Wanner, P. (2021), "How Well Can We Estimate Immigration Trends Using Google Data?", *Quality & Quantity* 55, 1181–1202.

larly important to target humanitarian assistance and facilitate integration. Ultimately, it allows for probing the intentions for future transit and return migration of refugees.

However, the use cases presented in this article are relatively simple and leave room for improvement. Relative search intensities are susceptible to several biases, which can be mitigated using additional knowledge of the studied population (Dergidiades et al. 2018). Moreover, one can consider several additional use cases of Google Trends in the current Ukrainian refugee crisis. As Google Trends can be obtained with a one-minute resolution, it can be used for real-time tracking of larger displaced groups. Finally, relative search intensity among refugees for potential destination countries could shed light on refugees' preferred destinations.

Although the measures calculated in this article are not perfect proxies, these are less costly, faster,