

## IMMIGRATION TO CROATIA: BIG (MIGRATION) DATA APPROACH

**Summary:** This research paper deals with the phenomenon of increased immigration of foreign citizens to Croatia, raises the question of the validity of official data on the number of foreigners in Croatia, and offers an empirical basis for the development of a Big Data model for predicting the future trend of immigration to Croatia. This paper hypothesizes that Big Data sources could be useful for determining immigration flows to Croatia, as well as for estimating and forecasting. This approach is enabling first insights into the identification of trends of migrants in the intentions to immigrate to Croatia and refugee flows. The methodological concept of here presented approach is to monitor the so-called “digital trace” of immigrants and refugees left on Google search and social networks Facebook (FB), Instagram and YouTube, and their geo-locations. The approach focuses on users that use social network platforms in foreign languages in Croatia as well as Google search with specific migration-related queries worldwide. **Results:** The analysis shows disproportion and therefore unreliability of immigration data coming from the Croatian Ministry of Interior (MUP), Croatian Bureau of Statistics (DZS), and Croatian Employment Service. The Big (Migration) Data approach shows a much higher number of immigrants in Croatia than official data. This approach shows that the number of Facebook and Instagram users in specific Asian and African languages is growing rapidly in Croatia, as well as in Ukrainian (and Russian) after the war outbreak in Ukraine. The tested FB and Instagram index is correlated with stepped-up immigration. The analysis of the FB groups of Asian and African immigrants in Croatia shows that those groups can be a valuable source for studying migration. One of the contributions of the second used method, Google search, and YouTube insights, is that it shows that by searching for specific migration-related queries and video material on the YouTube platform, the intention of users to migrate to the EU i.e., Croatia can be estimated. The insights obtained through this method align with official data (regarding trends), which, in turn, comes with a one-year delay. This makes this method particularly useful for timely assessments of the number of immigrants and modelling trends in immigration to Croatia.

*Keywords:* immigration; EU; Croatia; Big Data; foreign workers; refugees.

## I. Introduction

The observed disproportion and unreliability of immigration data coming from Croatian official immigration data has emerged as a problem in Croatian demographic research.<sup>1</sup> This difficulty is also observed regarding external migration flows from Croatia.<sup>2</sup> In addition to Ukrainian refugees, there is an increasing number of foreign workers, illegal migrants, and asylum seekers, who temporarily and permanently settle in Croatia. Since 2023 Croatia represents the external Schengen border of the EU, this phenomenon should be of increased interest to the European public - especially since once foreign workers from Asia and Africa enter Croatia, they can freely move anywhere within the Schengen Area.

The development of statistical tools that combine traditional and new sources of information such as Big Data is likely to become an accepted approach to monitoring demographic trends of all kinds.<sup>3</sup> Web-assisted research has been consistently validated in several other scientific fields, such as public health, economics, etc.<sup>4</sup> In migration studies researchers have begun to use alternative data (mobile phone records, social media use, satellite maps, and internet-based platforms: so-called big data), particularly to understand migration and mobility in light of new methodological approaches.<sup>5</sup> With the insights that can be obtained in this way, it is possible to discover certain patterns and regularities that users express in the need for information when they intend to emigrate, as well as the interests that reflect the intention to migrate. Such a set of data allows researchers to observe certain regularities according to which it is possible to construct models for predicting migration intentions, as well as their timely observation. The main advantage of this approach is enabling the timely observation and determination of the extent of external migration, which is an extremely important analytical factor for various policies, from the labour market to the issue of undertaking demographic revitalization measures.<sup>6</sup>

---

<sup>1</sup> The paper refers to data until the end of 2022, and was submitted on January 10, 2023.

<sup>2</sup> Jurić, Tado, 2018. *Iseljavanje Hrvata u Njemačku. Gubimo li Hrvatsku?* (Emigration of Croats to Germany. Are we losing Croatia?). Zagreb: Školska knjiga.

<sup>3</sup> Spyrtatos, Spyridon, Vespe, M., Natale, F., Ingmar, W., Zaghene, E. and Rango, M. 2018. Migration Data using Social Media: A European Perspective. *EUR 29273 EN, Publications Office of the European Union*, Luxembourg, 2018, ISBN 978-92-79-87989-0, doi:10.2760/964282, JRC112310.

<sup>4</sup> Jurić, Tado. 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.

<sup>5</sup> Wanner, Philippe. 2020. How Well Can We Estimate Immigration Trends Using Google Data? *Quality and Quantity* 55, 1181-202.

<sup>6</sup> Jurić, Tado. 2022b. Facebook i Google kao empirijska osnova za razvoj metode digitalnog praćenja vanjskih migracija hrvatskih građana, *Ekonomski pregled*, Vol. 73 No. 2, 2022., <https://doi.org/10.32910/ep.73.2.2>.

Previous research has established that digital data can be employed to study migration, but there are still significant methodological issues and scepticism regarding the feasibility of using alternative data sources. Because more research is necessary to assess the value of using big data for such research, this paper represents an attempt to measure the extent to which Big Data gained from Google search activities and social media can predict immigration intentions. It focuses on immigration from non-EU countries to Croatia as a case study, but this approach can also be applied to all EU countries.

The paper is organized as follows: it begins with a brief analysis of the immigration situation in Croatia and presents the findings of pertinent studies on Big Data within migration studies. Following this, we elaborate on the methodology employed and delineate the limitations inherent in this conceptual framework. Subsequently, there is a discussion on immigration from Asia and Africa to the EU, with a specific focus on Croatia, unveiling the outcomes derived from this approach. In the results section, we highlight the correlation between the social media index and official Croatian statistics, delving into the exploration of how social media can be utilized to detect and forecast immigration trends to the EU, i.e. Croatia.

## II. Immigration to Croatia

According to the results of the 2022 population census, there are 28,784 foreigners living in Croatia (0.74%).<sup>7</sup> According to the data of the Croatian Employment Service (HZZ), by November 2022, 101,1047 residence and work permits were approved and issued to foreign workers. The same was confirmed by the Ministry of Interior for this research: "As of October 4, 2022, a total of 10,316 citizens of third countries in Croatia have approved permanent residence or long-term residence by the Aliens Act, and 96,526 have valid temporary residence - mostly for the purpose of work (85,342)."<sup>8</sup>

Before the 2021 census in Croatia<sup>9</sup>, there were also more than 80 thousand foreign workers. The annual quota for foreign workers for 2020 was filled with 39,385<sup>10</sup>, and for 2019 with 49,317.<sup>11</sup> The new Law cancelled the annual quotas of

---

<sup>7</sup> DZS, 2022. Konačni rezultati popisa, <https://dzs.gov.hr/vijesti/objavljeni-konacni-rezultati-popisa-2021/1270>

<sup>8</sup> MUP (Ministry of the Interior), Minister's office (20.12.2022). Department for Public Relations.

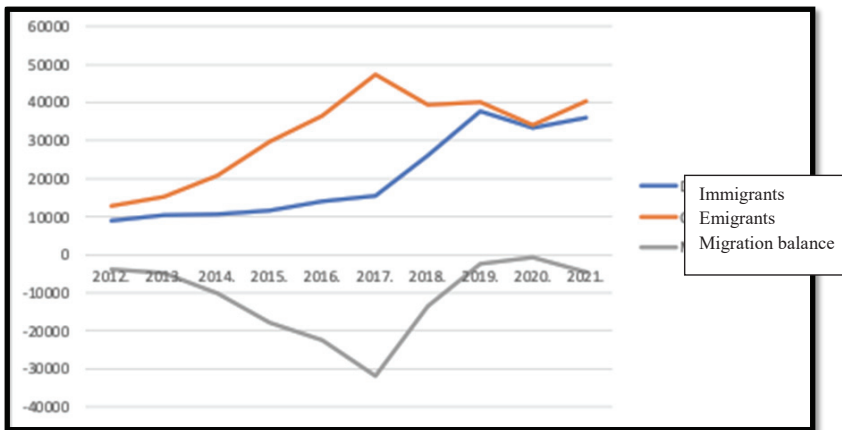
<sup>9</sup> DZS, 2022. Konačni rezultati popisa.

<sup>10</sup> MUP, 2021. Iskorištenost godišnje kvote na dan 31.12.2020.godine (Annual quotas for the employment of foreigners), <https://mup.gov.hr/gradjani-281562/moji-dokumenti-281563/stranci-333/statistika-169019/169019>

<sup>11</sup> MUP, 2020. Iskorištenost godišnje kvote na dan 27.12.2019. (Annual quotas for the employment of foreigners at 27.12.2019), <https://mup.gov.hr/UserDocsImages/statistika/2019/Iskori%C5%A1tenost%20kvota/ISKORISTENOST%20GODISNJE%20KVOTE%20ZA%202019:%20NA%20DAN%2031.12.2019..pdf>

permits for the employment of foreigners since 2021.<sup>12</sup> Thus, the rate of emigration from Croatia equalled the rate of immigration to Croatia in 2020.

Since 2022, immigration has been more intensive than emigration, and a positive migration balance of 11,685 has been registered<sup>13</sup> (it is important to note that a significant proportion of immigrants are Ukrainian refugees, 19,3%).<sup>14</sup> Croatia is therefore unquestionable turning from an emigrant into an immigrant country. The approach will show that most foreign workers immigrate and will continue to immigrate from Nepal, Philippines, and India, and the African countries: Ghana and Kenya. From the neighbouring countries, the highest number of new immigrants is expected from North Macedonia (in a relative ratio) and from Serbia and Bosnia and Herzegovina (in absolute ratio).



**Figure 1:** Emigration and immigration in the past decade (Migration balance Croatia 2012 – 2022)

Source: DZS, 2022. Migracija stanovništva Republike Hrvatske 2021., <https://podaci.dzs.hr/2022/hr/29030>

<sup>12</sup> Until January 1, 2021, residence and work permits for citizens of third countries were issued based on the government's decision on the annual quota. Source: HZZ, 2022, <https://www.hzz.hr/statistika/statistika-usluga-test-trzista-rada-i-radne-dozvole/> (22.12.2022).

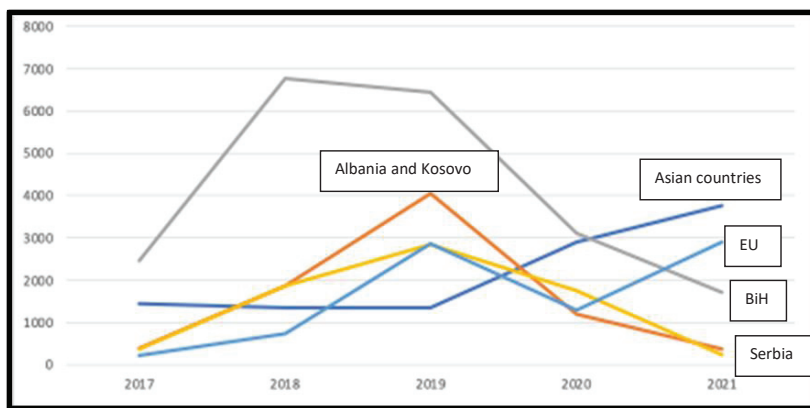
Instead of the quota system, a labour market test has been introduced. An employer who wishes to hire a foreign worker must, before applying for a residence and work permit, check if there is the possibility to employ a worker from the labour market of the Republic of Croatia. Source: HZZ, 2022. <https://www.hzz.hr/usluge/radne-dozvole-za-zaposljavanje-stranaca-i-test-trzista-rada/>.

<sup>13</sup> DZS, 2023. <https://podaci.dzs.hr/2023/hr/58061>.

<sup>14</sup> These data were available after the work was submitted for editorial review.

That there are problems with records and statistics is also confirmed by the Croatian Employment Service (HZZ), which states in the Employer Survey (2020): “The total estimate of the number of missing domestic workers in the population in 2020 is about 137,000, with about 25% of the shortage of domestic workers being filled by hiring foreigners.<sup>15</sup> By comparison with the number of used residence and work permits for foreigners from third countries at the end of 2020 (about 39,000), we can estimate that about 5,000 foreigners in 2020 (about 13% of newly employed foreigners) were employed without employers first looking for domestic workers labour market”.<sup>16</sup>

The phenomenon of immigration to Croatia from Asia is recent; the rise of this immigration occurs after 2017 and especially during 2020 and 2021. During 2022, there was a significant influx of Ukrainian refugees, and the absolute number of immigrants from Asia was almost equal to the number of immigrants from Ukraine in 2022 (11,874 compared to 11,121).<sup>17</sup> It is also noted that workers from Asia have begun to replace previous workers from the countries of Southeast Europe (BiH, Serbia, Kosovo, and Albania) (Figure 2).



**Figure 2:** Immigration of foreign citizens to the Republic of Croatia (2017-2021)

Source: DZS, 2022. Migracija stanovništva Republike Hrvatske 2021.,  
<https://podaci.dzs.hr/2022/hr/29030>

<sup>15</sup> HZZ, 2021. Anketa poslodavaca, <https://www.hzz.hr/app/uploads/2022/09/hzz-anketa-poslodavaca-2020-020321-1.pdf>

<sup>16</sup> Ibid.

<sup>17</sup> DZS, 2023, <https://podaci.dzs.hr/2023/hr/58061>.

Therefore, the question remains, how many foreign citizens live in Croatia regardless of confusing reports from different official databases?

### III. Big Data in Migration Studies - A Review

Traditional statistics often fail to measure the complexity of immigration flows to the EU.<sup>18</sup> This limitation is the reason for the development of new methods based on alternative sources, so-called Big Data.<sup>19</sup> A study by Singh et al. on internally displaced persons' movements between provinces in Iraq showed that a mix of social media data and traditional register data improves the predictive quality compared to predictions based on register data alone.<sup>20</sup> Several studies have used Big Data sources to analyse migration, i.e. refugee flows.<sup>21</sup> It is proven that migrants and refugees are more interested in information from the Internet than the average and that they generally use smartphones during migration.<sup>22</sup> Therefore, Big Data has the potential for governments to generate better statistics and improve their early warning systems.<sup>23</sup> Previous research has proven that Big Data can be used to study migration, but there are still significant methodological issues and scepticism regarding the feasibility of using alternative data sources.<sup>24</sup> Zagheni<sup>25</sup> was the first to show that Facebook (FB) can be an invaluable source of data unused for demographic research and that it offers a set of data that is a kind of "constantly updated census" but many authors also questioned this approach.<sup>26</sup>

<sup>18</sup> Wanner, Philippe. 2020.

<sup>19</sup> c.f. Jurić, Tado. 2022a.

<sup>20</sup> Singh, Lisa 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*, p. 1975–1983, <https://doi.org/10.1145/3292500.3330774>

<sup>21</sup> Connor, Phillip. 2017. The digital footprint of Europe's refugees. *Pew Research Center*, [https://www.pewresearch.org/global/wp-content/uploads/sites/2/2017/06/Pew-Research-Center\\_Digital-Footprint-of-Europes-Refugees\\_Full-Report\\_06.08.2017.pdf](https://www.pewresearch.org/global/wp-content/uploads/sites/2/2017/06/Pew-Research-Center_Digital-Footprint-of-Europes-Refugees_Full-Report_06.08.2017.pdf); Wanner Philippe. 2020.

<sup>22</sup> Ibáñez Sales, Matias. 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*, p. 2.

<sup>23</sup> UNHCR Global Data Service, 2021. Big (Crisis) Data for Predictive Models. A Literature Review, *UNHCR*, p 11., <https://www.unhcr.org/media/big-crisis-data-predictive-models-literature-review>.

<sup>24</sup> Jurić, Tado. 2022 a,b.

<sup>25</sup> Zagheni, E., M. Polimis, M. Alexander, I. Weber, and F. C. Billari. 2020. Combining Social Media Data and Traditional Surveys to Nowcast Migration Stocks in the United States. *Population Research and Policy Review*, <https://doi.org/10.1007/s11113-020-09599-3>.

<sup>26</sup> Cesare, N., H. Lee, T. McCormick, E. Spiro, and E. Zagheni. 2018. Promises and Pitfalls of Using Digital Traces for Demographic Research, *Demography* 55: 1979–99.; Schneider, D., and K. Harknett. 2019. What's to Like? Facebook as a Tool for Survey Data Collection, *Sociological Methods & Research* 20 (10): 1–33.

FB and Instagram (now “Meta”<sup>27</sup>) have devoted significant resources to assessing their users’ demographic characteristics, behaviours, and interests to improve targeted advertising, which is their primary source of revenue. For this purpose, FB has specifically developed a platform called *Adverts Manager* (FB ADS) and integrated it in 2022 with Instagram to *Meta Business Suite*<sup>28</sup>, allowing advertisers to select detailed user characteristics to which their ads should be displayed.<sup>29</sup> Socio-demographic data includes data such as the geo-location, age or gender of users, networks of friends, and related websites, as well as expressed interests. For example, FB and Instagram support the display of advertisements exclusively to Ukrainian emigrants/refugees who are adults and are now in Zagreb. Due to such a possibility, Meta tools offer beneficial insights for demographic research.<sup>30</sup>

Jurić has shown that Croatian would-be migrants often use online searching to get answers about the country they plan to emigrate to<sup>31</sup>, and Wanner that Google is the first source of information for most users planning to emigrate.<sup>32</sup> Jurić (2022 c,d)<sup>33</sup> tested the predictive capacity of Google Analytics regarding refugee flow from Ukraine and proved that these insights give reliable data and can help model future trends.

The first successful analysis of Google Trends migration data examined the searches of certain Arabic terms in Turkey and Germany via words such as “Greece” or “Germany” during the Migration Crisis of 2015.<sup>34</sup> This study showed that digital traces of internet searches can illuminate the movement of migrants.

Although Facebook (FB), Instagram, and YouTube are the most used social platforms<sup>35</sup>, very few studies have been written about their potential for migration studies (in South-eastern Europe). We will show that there are several

---

<sup>27</sup> MetaBusinessSuite, [https://business.facebook.com/latest/insights/people?asset\\_id=449256925513066&nav\\_ref=audience\\_insights](https://business.facebook.com/latest/insights/people?asset_id=449256925513066&nav_ref=audience_insights)

<sup>28</sup> Meta Business Suite, *ibid.*

<sup>29</sup> Jurić, Tado. 2022 b.

<sup>30</sup> Jurić, Tado. 2022 c. Big (Crisis) Data in Refugee and Migration Studies – Case Study of Ukrainian Refugees, *Comp. Southeast Europ. Stud.* 2022; 70(3): 540–553, <https://doi.org/10.1515/soeu-2022-0048>

<sup>31</sup> Jurić, T. 2022 a.

<sup>32</sup> Wanner Phillippe. 2020.

<sup>33</sup> Jurić, Tado. 2022 c.; Jurić, Tado. 2022 d. Predicting refugee flows from Ukraine with an approach to Big (Crisis) Data: a new opportunity for refugee and humanitarian studies, *Athens Journal of Technology & Engineering*, doi: 10.30958/ajte.9-3-1.

<sup>34</sup> Connor, Phillip. 2017.

<sup>35</sup> Statista.com, 2023. Most popular social networks worldwide as of October 2023, ranked by number of monthly active users, <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.

approaches that insights obtained from digital traces left on social networks can be used to identify and model migratory movements of immigrants to Croatia. Because a part of immigrants to Croatia are asylum seekers and refugees and therefore are a vulnerable group, social research with a typical approach of interviews and surveys at this time is not feasible and, in some cases, inappropriate due to the traumatisation of respondents – this is another advantage of this method.<sup>36</sup>

The question of the rate of internet penetration in observed societies is certainly important in this method. By mid-2022, 69% of the world's population was estimated to be internet users: Asia 67,4 %, Africa 40%, Europe 92%.<sup>37</sup> The Google search engine is by far the most popular in the world, preferred by 83,84% of users. The percentage of citizens with no experience using the internet increases with advancing age. Most people not using the internet are over 65.<sup>38</sup> Since members of this age group rarely emigrate, they are irrelevant to our study. In 2022 the most popular social networks worldwide are FB, YouTube, WhatsApp, and Instagram.<sup>39</sup>

#### IV. Methodology

The primary method we used for analysing digital traces on Google and YouTube is based on search language and geo-location, and for FB and Instagram on geo-location and the primary language of users. Both tools allow analysis according to the country where the specific group uses these applications.

Google Trends analytics is a search tool that highlights the popularity of particular search terms on Google, enabling researchers to discern emerging or waning trends. It's important to emphasize that Google Trends doesn't directly disclose the total search volume. Rather, it standardizes search volume on a scale of 0 to 100 for the specified period, with higher values signifying peak search volume. This standardization facilitates the provision of verifiable metrics. Normalizing the data is necessary because the number of people searching on Google constantly changes over time.<sup>40</sup>

---

<sup>36</sup> For a detailed review of the literature and the current discussion on the application of Big Data in migration and refuge studies, see my previous studies Jurić, 2022 b,c.

<sup>37</sup> Statista.com, 2022. Global internet penetration rate as of July 2022, by region, <https://www.statista.com/statistics/269329/penetration-rate-of-the-internet-by-region/>

<sup>38</sup> Jurić, Tado. 2021a.

<sup>39</sup> Statista.com, 2023.

<sup>40</sup> See: Google Trends, <https://trends.google.com>.



Using Google Trends and YouTube Insights, one should especially take care to choose specific keywords, i.e. migration-related queries. As mentioned, according to Wanner, Wilde et al.,<sup>41</sup> and Böhme et al.<sup>42</sup>, one can expect to find a relationship between the intention of the users and particular behaviours. Therefore, we have focused on monitoring the terms that indicate the intention of migrants to move to Croatia. For example, if someone in Nepal googles “Croatia” this does not necessarily imply an intention to move there; it might simply indicate a desire for information about the world championship in football. But if the search query is entered “Croatia + job” this is indices about emigration planning.

**Table 1.** List of Keywords

<b>English (note: with translations into Arabic, Russian, Ukrainian, French, Macedonian, Albanian, Africans-Swahili)</b>	
<i>asylum</i>	- <i>Learning Croatian</i>
<i>border</i>	- <i>Employment agency + Croatia</i>
<i>border control</i>	- <i>Life in Croatia</i>
<i>migrant</i>	- <i>Registration of residence in Croatia</i>
<i>refugee</i>	- <i>Job in Croatia</i>
<i>interpreter</i>	- <i>salary + Croatia</i>
<i>passport</i>	- <i>Wage + Croatia</i>
<i>Schengen</i>	- <i>Earning + Croatia</i>
<i>visa</i>	- <i>immigration policy in the EU</i>
<i>job</i>	- <i>Emigration to Croatia</i>
<i>PCR</i>	- <i>Employment agency + Croatia</i>
<i>consulate</i>	- <i>weather forecast + Croatia</i>
<i>citizenship</i>	- <i>immigration policy + Croatia</i>
<i>mosque</i>	- <i>corona + Croatia</i>
<i>Zagreb</i>	- <i>Visa + Croatia</i>
	- <i>Residence permit +Croatia</i>
	- <i>Language + Croatia</i>

<sup>41</sup> Wilde, J., W. Chen, and S. Lohmann. 2020. COVID-19 and the Future of US Fertility: What Can We Learn from Google?, *IZA Discussion Papers* 13776. [www.iza.org/publications/dp/13776/ covid-19-and-the-future-of-us-fertility-what-can-we-learn-from-google](http://www.iza.org/publications/dp/13776/covid-19-and-the-future-of-us-fertility-what-can-we-learn-from-google).

<sup>42</sup> Bohme, M. B., A. Groger, and T. Stohr. 2020. Searching for a Better Life: Predicting International Migration with Online Search Keywords. *Journal of Development Economics* 142, [https:// doi.org/10.1016/j.jdeveco.2019.04.002](https://doi.org/10.1016/j.jdeveco.2019.04.002).

With this approach, it is necessary to avoid periods characterised by changes in immigration policies, and major sports events and to include overall context such as the Covid-19 pandemic. The control mechanism for determining data deviations was performed by comparing official databases, including the Croatian Ministry of Interior (MUP), the Croatian Bureau of Statistics (DZS) and the Croatian Employment Service (HZZ). The main advantage of this approach lies in the timely manner in which the data can be obtained.

Meta Business Suite can provide insights into geo-location and particular interests of the observed population based on many signals such as likes, pages visited, and specific cultural interests. In the following, we will show that these insights can be used to model immigration flows to Croatia. The same approach can be used to analyse digital tracks on Instagram, while insights from YouTube, the second most visited social network, can be obtained by analysing keyword searches as in the case of Google search.

Although previous research in this area has established the feasibility of using digital data to study migration<sup>43</sup> methodological issues remain.<sup>44</sup>

### Limitations of the Concept

The primary limitations regarding FB and Instagram data in this study are as follows: Those data are not representative, they depend on the Internet penetration rate in a specific group and do not include all age groups equally (for example, due to GDPR provisions, children are not included at all). The problem also lies in the possibility that users have multiple unlinked FB and Instagram accounts, leading to data distortion. However, after 2018, such practices have been significantly reduced after the scandal with Cambridge Analytica because FB then shut down millions of fake profiles, and in suspicious cases, it requires verification of personal information.

By using this method, it is most important to pay attention to the limitations that can lead to misinterpretations of the results. *Meta* states that estimates are not intended to be consistent with population, census, or other sources. These estimates depend on how many accounts an individual has on the FB and Instagram Apps and Services; how many transient visitors are in a specific geographic location, and demographic information provided by FB users themselves.<sup>45</sup> A serious limitation of our entire approach is that many immigrants use

---

<sup>43</sup> Zagheni, E., I. Weber, and K. Gummadi. 2017.

<sup>44</sup> See: Cesare et al. 2018.

<sup>45</sup> Facebook, *Geschätzte Zielgruppengröße*, <https://www.facebook.com/business/help/1665333080167380?id=176276233019487> (09.04.2022)

these social platforms and Google search in other languages, and not primarily in the native language. For example, in the case of Ukrainians, many refugees use Russian.<sup>46</sup>

To track FB and Instagram users by native language and geo-location, we have collected data about users worldwide every week during 2022. This procedure was necessary because *Meta* offers data only for the present day with the ability to compare this day with the average of the past 12 months. Therefore, it is important to note that the researcher has to keep their own data archive – which is a significant limitation compared to data from the YouTube platform and Google Trends, which provides insights into every date from 2008 to date. The same issue also refers to the analysis of FB groups. Furthermore, a problem arises from Facebook's annual network estimates being available only for user age groups from 18 to 64 years of age.<sup>47</sup> There are also important limitations in using these tools (Facebook and Google Trends) due to the unequal representation of females and members of older age groups.

The YouTube and Google search Index insights do not show the exact number of searches in a specific country, so with this tool, the exact number of potential emigrants, cannot be calculated, but the increase in the trend can be noticed very precisely<sup>48</sup>. Furthermore, Google Trends offers no demographic data on population structure.

When interpreting the results of the YouTube and Google Trends analytics tools, a problem arises with the identification of the analysed group according to the search language used.<sup>49</sup> The issue is particularly pronounced in the case of the countries of the Western Balkans and affects the methodological procedure as it pertains to the Serbian, Albanian and Macedonian minorities in Croatia.

Using Google search i.e., Google Trends, one should pay attention to variations in the search index depending on the keywords of a given query. Certain delimiters such as “-” and “+” allow users to change the combinations of keywords searched. A search for a single keyword will yield the search frequency index counting all searches containing that keyword, including searches that contain other words.<sup>50</sup> It is equally important that one must account for the overall social

---

<sup>46</sup> 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. Source: Magoscy, R., 1996. *A History of Ukraine*, Toronto: University of Toronto Press.

<sup>47</sup> Spyrtatos et al. 2018.

<sup>48</sup> c.f. Jurić, Tado. 2022 c.

<sup>49</sup> c.f. Jurić, Tado. 2022 a,b.

<sup>50</sup> Wilde, J., W. Chen, and S. Lohmann. 2020.

context of both home and destination countries, as disruptive events like the COVID-19 pandemic and the varying national policies in response may influence such currents and trends.

Although these data are not representative of the observed population, studies have provided evidence that samples obtained from FB do not significantly differ from those obtained by more traditional recruitment and sampling techniques (Pötzschke and Braun 2017<sup>51</sup>; Zhang et al. 2018<sup>52</sup>). It is encouraging that more and more researchers are perfecting this method and that both Google (YouTube) and Meta (FB and Instagram) realise specific data collection problems and are improving their methods of data collection. Combining traditional and new data sources is certainly key to making progress in migration studies.

From an ethical point of view, it is questionable how ethical it is to use data that was not collected with user consent (even when no personal data is known, as in this study).<sup>53</sup>

## V. Results and Discussion

Before we present the data obtained through “Big Data” insight, at the beginning of this section, for the purposes of comparison, we list the official data of the DZS (CBS, Croatian Bureau of Statistics), HZZ (CES, Croatian Employment Service) and MUP (Croatian Ministry of Interior) and show the discrepancy in the official reports on the number of foreigners in Croatia.

---

<sup>51</sup> Pötzschke, S., and M. Braun. 2017. Migrant Sampling Using Facebook Advertisements: A Case Study of Polish Migrants in Four European Countries. *Social Science Computer Review* 35 (5), 633-53.

<sup>52</sup> Zhang, B., M. Mildenerger, P. D. Howe, J. Marlon, S. A. Rosenthal, and A. Leiserowitz. 2018. Quota Sampling using Facebook Advertisements, *Political Science Research and Methods* 8 (3), 558-64.

<sup>53</sup> For a complete list of restrictions, see Jurić, Tado. 2022 a,b.

**Table 2:** Immigration to Croatia by country of citizenship 2017 - 2022 according to DZS

Country of citizenship	2017	2018	2019	2020	2021	2022
	immigration					
<b>In total</b>	15 553	26 029	37 726	33 414	35 912	57 972
<b>Croatia</b>	7 911	8 619	9 882	8 460	10 622	10 340
<b>Europe</b>	6 891	16 109	24 194	22 220	20 223	34 367
<b>European Union</b>	2 087	2 108	2 314	3 814	4 753	5 167
<b>Other European countries: (BiH, Kosovo, Serbia, North Macedonia)</b>	4 804	14 001	21 880	18 406	15 470	29 200
<b>Asia</b>	401	880	3 084	2 159	4 358	11 874
<b>Africa</b>	70	83	139	135	177	675

Source: DZS, 2018 - 2023, edited by author.

When we compare the data of the DZS (CBS) from Table 2 with the data of the MUP (Table 3) about realised annual quota for foreign workers, it can be seen an opposite picture of immigration to Croatia, which for 2018 is (+16.17%), for 2019 (+43.54%), for 2020 (+ 36.64%) and for 2022 (+47,11 %). A particularly problematic issue in the official data recording is that Croatia does not record immigrants from Asia and Africa according to the country of immigration but only states the continent.

When comparing only the last two years, it can be observed that the number of immigrants from Asia has tripled, and the number of immigrants from Africa has increased sixfold. At the same time, there is a consistent rise in immigrants from the EU, as well as from neighbouring countries in Southeastern Europe.

**Table 3:** Difference in records of foreigners who immigrated to Croatia according to official databases 2018-2022

<b>CBS (DZS) data on immigrated foreigners</b>	<b>2018</b>	<b>2019</b>	<b>2020</b>	<b>2021</b>	<b>2022</b>
	17,410	27,844	24,954	28,784	47,652 <sup>54</sup>
<b>MUP data (realised annual quota) / CES</b>	<b>2018</b>	<b>2019</b>	<b>2020</b>	<b>2021</b>	<b>2022</b>
	20,769	49,317	39,385	* (quotas abolished) No data available	101,147 (Data from MUP <sup>55</sup> )
<b>Difference</b>	<b>3,359</b> (= 16,17 %)	<b>21,473</b> (= 43,54 %)	<b>14,431</b> (= 36,64 %)	---	<b>53,495</b> (=47,11 %)

Sources: Author according to data from: 1) HZZ, <https://www.hzz.hr/statistika/statistika-usluga-test-trzista-rada-i-radne-dozvole/>; 2) MUP, 2021-2023; 3) DZS, 2019-2023.

The presented analysis shows a difference in official databases of + 39,263 foreign workers who immigrated to Croatia in the period from 2018 to 2020. Ministry of Internal Affairs (MUP) data for 2021 are not available, while the difference in official databases for 2022 of + 53,495.

Since the quotas were abolished in 2021 and the import of foreign workers in 2022 increased by over 55% compared to 2019 and 2020, in recent times, we can start from the assumption that the error has also increased in the records of foreigners. To solve this problem, below we present a Big (Migration) Data concept.<sup>56</sup>

### Google Analytics as a source of immigration data

This section provides a descriptive analysis of Big Data sources to determine whether, based on past data, regularities of immigration trends to Croatia can be observed.

<sup>54</sup> DZS, 2023. <https://podaci.dzs.hr/2023/hr/58061>. Note: These data were published after the completion of the work.

<sup>55</sup> MUP, 2023.

<sup>56</sup> Official data subsequently published confirmed the validity of the estimates provided by the Big (Migration) Data method. See: DZS, 2023. <https://podaci.dzs.hr/2023/hr/58061>

As we mentioned, Croatian statistics do not make a distinction between the countries of Asia and Africa from which foreign workers come, so it is unknown from which country the largest number of foreign workers come, and from which is to expect the largest immigration. Our method fills in these gaps and shows that these are India and Nepal.<sup>57</sup> The argumentation based on the analysis shows that at the world level, the term “immigration to Croatia” is most searched in India and Nepal (in English) (Figure 3).

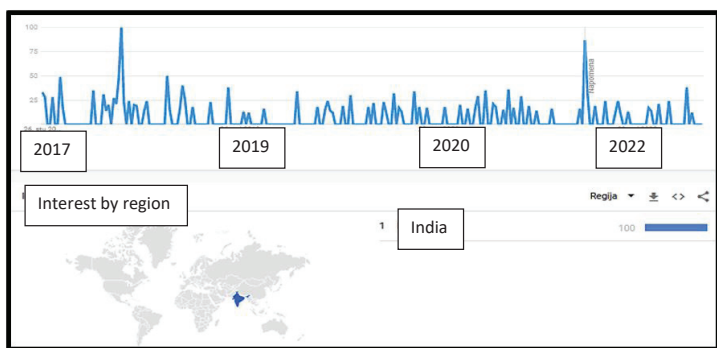


Figure 3: Search query „immigration to Croatia“ in English worldwide (2017 – 2022)

Another confirmation of the hypothesis that workers from India and Nepal are most interested in Croatia shows that at the world level in the Arabic language, the largest number of searches related to the term “Croatia + job” comes from these countries (Figure 4).

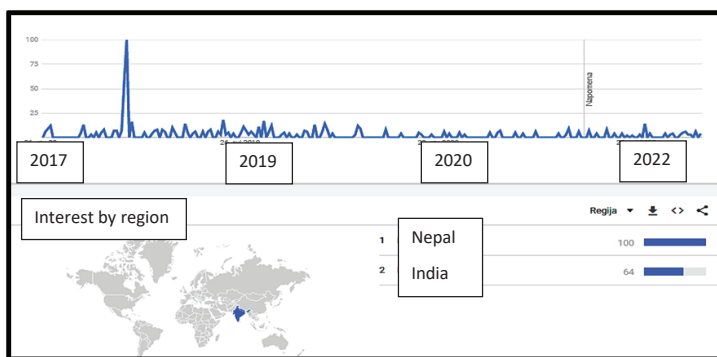
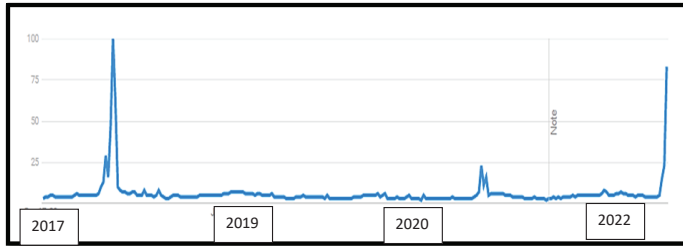


Figure 4: Search query „ايتاورك“ (Croatia) in Arabian worldwide (2017 – 2022)

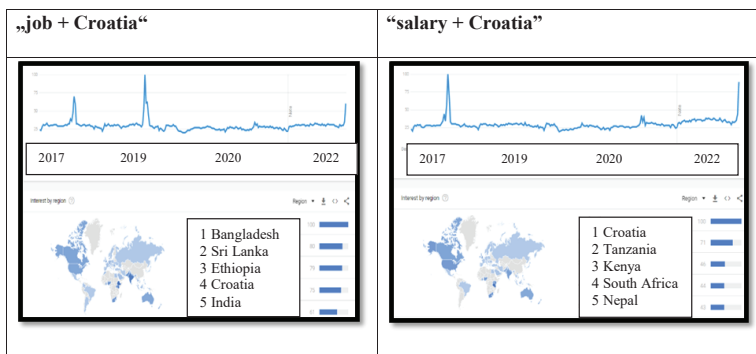
<sup>57</sup> Subsequently published official data have confirmed this.

That Croatia is becoming more and more interesting to migrants worldwide is also shown by the increase in interest in submitting asylum applications in Croatia (Figure 3). The analysis confirms again that Nepal and India are among the ten most represented non-European countries in terms of this expressed interest. Regarding interest in obtaining Croatian citizenship, Nepal stands out again, and there is an increase in interest at the level of numerous Asian and African countries.



**Figure 5:** Search query „asylum + Croatia“ in English worldwide (2017 – 2022)

The analysis of digital traces shows that potential immigrants to Croatia, when planning their migrations, most often search for the terms “job + Croatia”, and “salary + Croatia”. We noticed that those searches were particularly intensive after Croatia increased the annual quotas for the import of foreign workers after 2016.<sup>58</sup> We take this finding as initial evidence that the immigration flows to Croatia can be predicted. Search queries „job + Croatia“ and “salary + Croatia” at the world level (Figure 4) show that Croatia has been particularly focused on workers from Asia and Africa for the last five years.



**Figure 6:** Search queries a) „job + Croatia“ and b) “salary + Croatia” in English worldwide (2017 – 2022)

<sup>58</sup> MUP, Iskorištenost godišnjih kvota, <https://mup.gov.hr/gradjani-281562/moji-dokumenti-281563/stranci-333/statistika-169019/169019>



When we look in more detail at certain countries that are showing themselves as countries that send workers, we see a growing trend of interest in Croatia in Pakistan, India, Tunisia, and the Philippines.

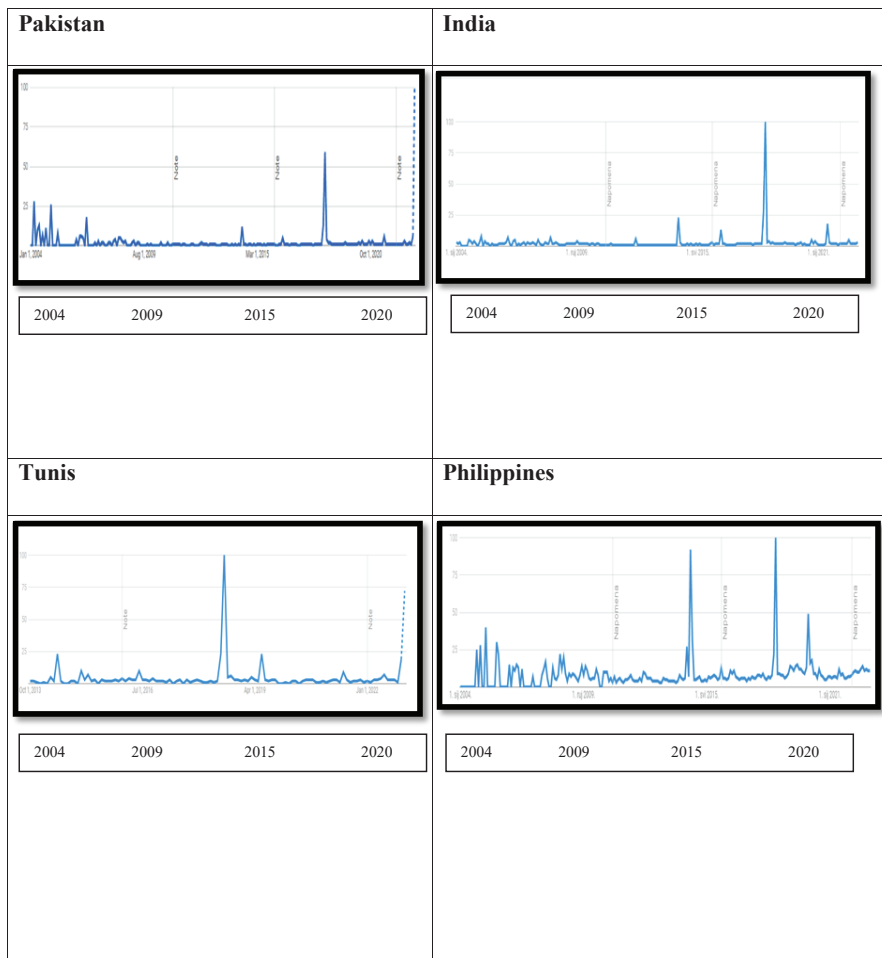
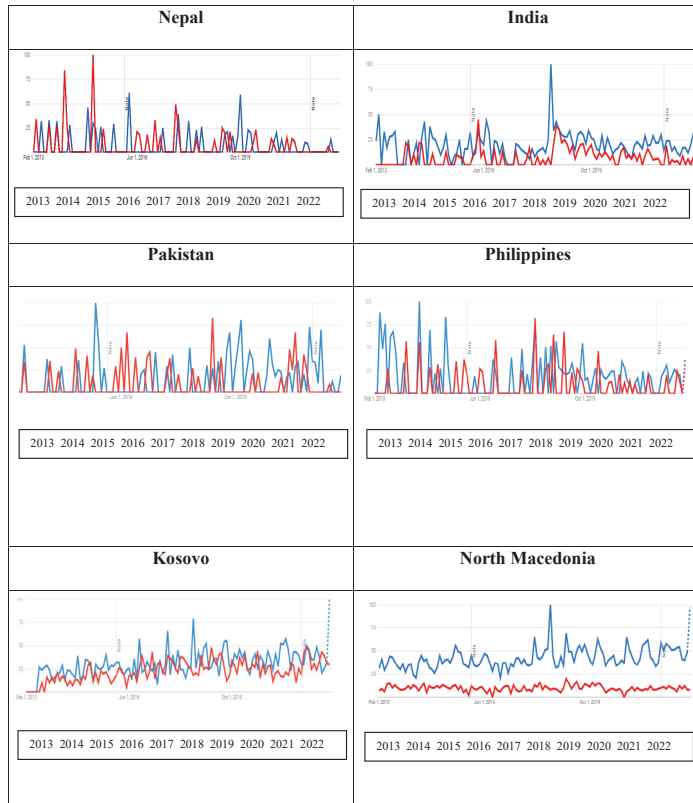


Figure 7: Search query „Croatia“ in English in Pakistan, India, Tunis, and Philippines 2004 - 2022

All analyzed examples (Figures 4 and 5) show that there is a correlation between increased migration-related search queries and increased immigration to Croatia after 2017. Increased immigration is confirmed by the number of issued quotas for foreign workers (2016 = 2,100; 2017 = 6,500; 2018 = 20,769; 2019 = 49,317, 2020 = 39,385).<sup>59</sup>

The indication of the increase in interest in Croatia is shown even more clearly by the following Figure 6, which shows a search for queries in the Croatian language - a language that is not, by any criteria, characteristic of the mentioned countries.



**Figure 8:** Interest in the terms "Hrvatska + posao" (Croatia + job) in Croatian in Nepal, India, Pakistan, Philippines, Kosovo, and North Macedonia from 2013 to 2022

Note: Red line „posao“ (job), blue line „Hrvatska“ (Croatia), Authors creation, Source: Google Trends

<sup>59</sup> MUP, 2021.

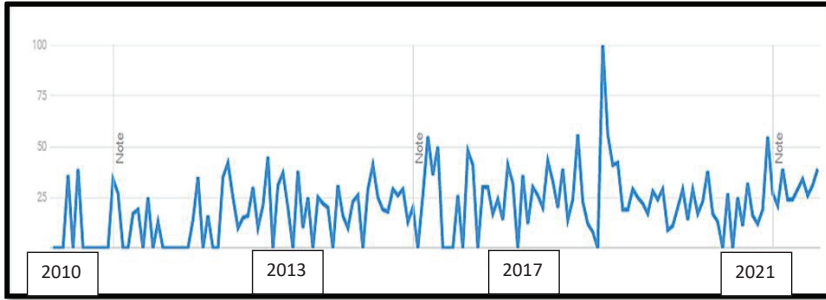
My analysis of the most relevant migration-related search queries (worldwide) that give clues about the planning of migration to Croatia shows that primarily African and Asian countries stand out (Table 4).

**Table 4:** Migration-related search queries „Employment + Croatia“, „Job + Croatia“, „Visa + Croatia“, „Residence permit + Croatia“, „Language + Croatia“ according to the geolocations worldwide (2017-2023)

Employment + Croatia		Job + Croatia	Visa + Croatia	Residence permit + Croatia	Language + Croatia
1	Nepal	Bangladesh	North Macedonia	North Macedonia	Nepal
2	India	Sri Lanka	Bangladesh	Bangladesh	Nigeria
3	Ghana	India	Nepal	Nepal	Ethiopia
4	Zambia	Ethiopia	Uganda	Uganda	Ghana
5	Kenya	Kenya	Kenya	Kenya	Burma
6	Uganda	Nepal	Zambia	Zambia	Kenia
7	Bangladesh	Pakistan	Nigeria	Nigeria	Bangladesh
8	North Macedonia	Uganda	Ghana	Ghana	Pakistan
9	Somalia	Nigeria	Albania	Albania	Tanzania
10	Nigeria	Ghana	India	Romania	Uganda

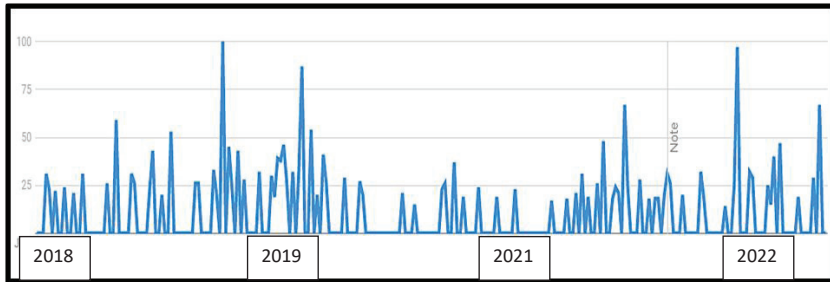
According to the analysis of Internet traces, the countries from which the largest immigration to Croatia can be expected are Asian countries Nepal and India, the African countries: Ghana and Kenya, while from the neighbouring countries, it is North Macedonia.

A further indication of the increase in immigration to Croatia using this method can be seen when analysing the search for the term “mosque” in Croatia from 2010 - 2022, which clearly shows an increase. Since many immigrants are of the Muslim faith, their search for a place of prayer in an unfamiliar environment is understandable.



**Figure 9:** Search query „mosque“ in Croatia from 2010 to 2022

This approach can also be used in security studies because it provides clues about the movement of illegal migrants. Namely, migrants on their way use internet data to find the fastest or unsecured route to enter the EU, i.e. Croatia. When such searches are followed, one can observe increased activity in a language uncommon for a particular area (for example, Arabic) as shown in Figure 9.



**Figure 10:** Search query „ايتاورك“ (Croatia) in Arabian in Bosnia and Herzegovina from 2015 to 2020

The increase in searches for the term “Croatia” in the Arabic language in BiH corresponds to increased illegal entries into Croatia during 2019, while with the onset of the pandemic in 2020, a decrease in activity can be clearly seen, but also a re-increase in early 2022.<sup>60</sup> These insights also provide the geolocation of potential migrants because they clearly locate them around the town of Bihać in BiH, which is located on the border with the Republic of Croatia, i.e. the Schengen border.

<sup>60</sup> Croatian Government, 2019. Božinović: Hrvatska neće dopustiti ilegalne migracije, <https://vlada.gov.hr/vijesti/bozinovic-hrvatska-neece-dopustiti-ilegalne-migracije/26678>

## Results: Social media as a method for determining and measuring immigration trends

In Croatia, numerous FB groups and portals have been created since 2019 where immigrants exchange information about life in Croatia and maintain contact with compatriots and satisfy the need for social contact, which is often lacking in the first stages of emigrant life. FB groups whose number of members has grown especially from 2020 to 2022 are Nepali people in Croatia, Indians in Croatia, and Filipinos in Croatia (Table 5).

**Table 5.** Facebook groups of immigrants in Croatia are a source of socio-demographic data

Facebook Group	Members	Link
<i>Nepali people in Croatia (Nepalci u Hrvatskoj)</i>	<b>3,005</b>	<a href="https://www.facebook.com/groups/188323796577370">https://www.facebook.com/groups/188323796577370</a>
<b>Indians in Croatia</b>	<b>3,303</b>	<a href="https://www.facebook.com/groups/2471039829882421">https://www.facebook.com/groups/2471039829882421</a>
<b>Filipinos in Croatia</b>	<b>23,161</b>	<a href="https://www.facebook.com/groups/Kabayanincroatia">https://www.facebook.com/groups/Kabayanincroatia</a>
<b>Makedonci u Hrvatskoj -- Македонци во Хрватска (информации и комуникација)</b>	<b>8,851</b>	<a href="https://www.facebook.com/groups/557350958419504">https://www.facebook.com/groups/557350958419504</a>
<b>SHQIPTARËT NË KROACI / ALBANCI U HRVATSKOJ</b>	<b>2,634</b>	<a href="https://www.facebook.com/groups/forumalbanacauhrvatskoj">https://www.facebook.com/groups/forumalbanacauhrvatskoj</a>
<b>Українці в Хорватії / Ukrajinці u Hrvatskoj</b> [Ukrainians in Croatia] Note: 6,700 at 01.04.2022	<b>8,400</b> (15.12.2022)	<a href="https://www.facebook.com/groups/756120944473505">https://www.facebook.com/groups/756120944473505</a>
<b>Deutsche in Kroatien</b>	<b>4,192</b>	<a href="https://www.facebook.com/groups/deutsche.in.kroatien">https://www.facebook.com/groups/deutsche.in.kroatien</a>
<b>Auswandern nach Kroatien</b>	<b>33,637</b>	<a href="https://www.facebook.com/groups/3479079895506344">https://www.facebook.com/groups/3479079895506344</a>
<b>Bangladeshi community In Croatia</b>	<b>1,473</b>	<a href="https://www.facebook.com/groups/771671490020362">https://www.facebook.com/groups/771671490020362</a>

Source: FB API, data collected and systematised by the author

The text analyses of those FB groups show that those groups serve as an essential source of information for those who immigrated to Croatia but also influence the decision to migrate to Croatia by many others. It is to be noted that

the number of members of a certain FB group does not necessarily mean only members who already live in Croatia.

Another approach to gain insights into the immigration of foreign citizens to Croatia give an analysis of users of FB and Instagram in selected foreign languages in Croatia (Table 6).

**Table 6.** Users of FB and Instagram in selected foreign languages in Croatia (2022)

Year	Macedonian	Albanian	Filipino	French	Ukrainian	Slovenian	German	Russian
2022								
Users	4,000-4,700	10,200-12,000	2,300-2,700	2,400-2,800	3,600-4,200	6,300-7,400	18,100-21,300	6.800-8.000
	Arabic	Portuguese	Turkish	Italian	Serbian	Bosnian	Spanish	
	1,500-1,800	1,300-1,500	1,900-2,300	17,300-20,400	9,400-11,100	5,500-6,400	5,000-5,900	

Data collected and systematised by the author.

With this approach, we encounter serious limitations of this method, which are reflected in the fact that 1) part of the users from the mentioned countries use FB and Instagram in English 2) that there is a national minority of Albanians, Italians, Slovenes, Serbs, and Bosniaks living in Croatia 3) that part of users who use FB and Instagram in Spanish or German certainly corresponds to the Croatian diaspora living in Europe and overseas countries.

Particularly useful insights are provided by testing this approach on the example of Ukraine. My approach to social media platforms shows that the number of FB and Instagram users in Ukrainian is growing rapidly in Croatia and other EU Member States after the war outbreak in Ukraine.<sup>61</sup>

<sup>61</sup> c.f. Jurić, Tado. 2022 c.

**Table 7.** The number of Facebook and Instagram users in Ukrainian (Ukrainian refugees) in Croatia, Poland, Slovakia, Hungary, Romania, and Germany (05. March – 20. Dec 2022)

Year 2022	Croatia	Slovakia	Hungary	Poland	Romania	Germany
05. March	1,000	no data available	no data available	282,000	no data available	51,400
05. April	1,400	17,800	16.400	396,500	2.600	72,700
20. Dec	4,200	32,700	22,200	651,100	7,100	234,600

Source: Meta Business Suite, data collected and systematised by the author.

Note: The maximum estimate value is displayed. Data for Slovakia, Hungary, and Romania were not reported because, at the time of the study, they did not exceed 1,000 users, which is the minimum for Meta to register.

From Table 7 we can see that Meta Analytics correctly noticed an increase in the trend of immigration of Ukrainians in selected countries. For example: Croatia: 1,000 (05. Mar) - 4,200 (20. Dec); Poland 282,000 (05. Mar) - 651,100 (20. Dec); Germany: 51,400 (05. Mar) - 234,600 (20. Dec). By comparing those data with official statistics from UNHCR<sup>62</sup> and IOM<sup>63</sup> about Ukrainian refugees, we can conclude that in the first phase, FB and Instagram, on average, register 17% - 20% of the immigrant population in these countries. Although they do not show exact figures, these insights are helpful because they showed three months before the first official reports, the correct trend in the movement of Ukrainian refugees. The testing of the method turned out to be correct also regarding the projection of the total number of Ukrainian refugees in Germany.<sup>64</sup>

Other useful insights of this approach are demographic insights. Table 5 shows the age and sex of FB and Instagram users in Ukrainian (Ukrainian refugees) which we have compared to official data to test the use value of this data.

<sup>62</sup> UNHCR, 2022, <https://data.unhcr.org/en/situations/Ukraine> (14.11.2022).

<sup>63</sup> IOM, 2022. Almost 6.5 Million People Internally Displaced in Ukraine: IOM, <https://www.iom.int/news/almost-65-million-people-internally-displaced-ukraine-iom> (12.11.2022)

<sup>64</sup> Jurić, Tado. 2022c.

**Table 8.** Age and sex of Facebook and Instagram users in Ukrainian (Ukrainian refugees) in Poland, Germany, and Croatia after the war outbreak (Mar - Dec 2022)

Poland			Germany			Croatia		
age	female	male	age	female	male	age	female	male
18-24	8,7 %	5,2 %	18-24	4,4 %	2%	18-24	5,2%	2,4 %
25-34	21,7 %	14,4 %	25-34	19,3 %	8,7 %	25-34	19 %	8 %
35-44	18,8 %	10,4 %	35-44	25,9 %	10,8 %	35-44	27,3 %	9 %
45-54	9 %	5,3 %	45-54	11,3 %	6%	45-54	12,2 %	4,4 %
55-64	3,3 %	1,5 %	55-64	5,6 %	2,3 %	55-64	6,7 %	2,3 %
65+	1,2 %	0,5 %	65+	2,5 %	1,2 %	65+	2,5 %	1,0 %
<b>Average:</b> <b>62,7 % female, 37,3 % male</b>			<b>Average:</b> <b>69 % female, 31 % male</b>			<b>Average:</b> <b>27,1 % female, 27,1 % male</b>		

Source: Meta Bussines Suite, data collected and systematised by the author

Table 5 shows that the users are most represented in the age group 25-34 and 35-44 and that there are almost triple as many women in this age group - which overlaps with official Poland and German data<sup>65</sup> and data from the Croatian Ministry of the Interior (16,96 % M, 49,81 % F, 33,23 Children).<sup>66</sup> According to Meta, the share of women in the refugee population from Ukraine is 62.7% in Poland and 69% in Germany. When we set the filter to limit the specific interests that only Ukrainians in Poland and Germany could have and keep the Ukrainian language option, we get a more accurate picture showing a demographic structure with 73.1% of women in Germany and 68,2 % in Poland. A limitation of this comparison is that Meta does not show data for children (see limitations). In the case of Poland, it should be remembered that many Ukrainian workers were employed in Poland many years before the start of the war.<sup>67</sup>

Furthermore, this approach is useful in determining the destination of immigrants and Ukrainian refugees, according to which most immigrants come to live (and work) in Zagreb and its surroundings (30%-50% of all) and in the

<sup>65</sup> Bundesministerium des Innern und für Heimat, 2022. Pressemitteilung, 04.04.2022, Befragung ukrainischer Kriegsflüchtlinge, <https://www.bmi.bund.de/SharedDocs/kurzmeldungen/DE/2022/04/umfrage-ukraine.html>.

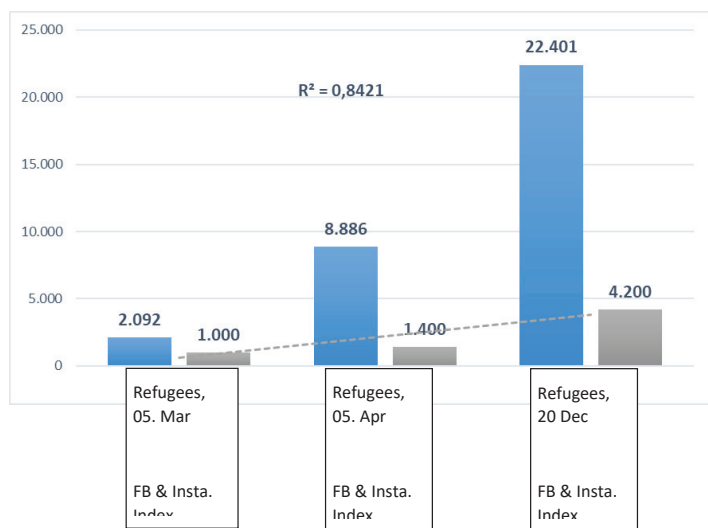
<sup>66</sup> see: <https://hrvatskazaukrajinu.gov.hr/> (20.12.2022)

<sup>67</sup> Euractive, 2022. Poland, Czech Republic struggle to replace the Ukrainian workers, <https://www.euractiv.com/section/economy-jobs/news/poland-czech-republic-struggle-to-replace-the-ukrainian-workers/> (20.12.2022)



other two largest Croatian cities, Split and Rijeka. These data (with a delay of half a year) are confirmed by the DZS (CBS).<sup>68</sup> All these findings show that FB and Instagram undoubtedly capture valuable socio-demographic insights on immigrants i.e. Ukrainian refugees and that this data source is of great use in a situation where there is no official data or delay in data.

In the following (Figure 11), we tested the correlation between the number of refugees from Ukraine in Croatia and Facebook and Instagram users in Ukrainian in Croatia during the year 2022.



**Figure 11.** Correlation between Refugees from Ukraine in Croatia and Facebook and Instagram users in Ukrainian in Croatia on 05. March, 05. April and 20 December 2022.

This tested correlation shows that the increase in the FB and Instagram index frequency is correlated with stepped-up emigration from Ukraine.  $R^2$  is 0.8421 and shows a positive correlation. Despite the mentioned shortcomings of the method, it is unquestionably that Big Data correctly observes recent trends.

### YouTube as a source of immigration data

Since FB and Instagram do not register the socio-demographic structure of people under 18, additional applications such as YouTube could be used to gain

<sup>68</sup> In the case of Germany, these are the cities: Berlin, Munich and Hamburg, and in the case of Poland, the cities Warsaw, Breslau and Krakow - which again corresponds to official indicators.

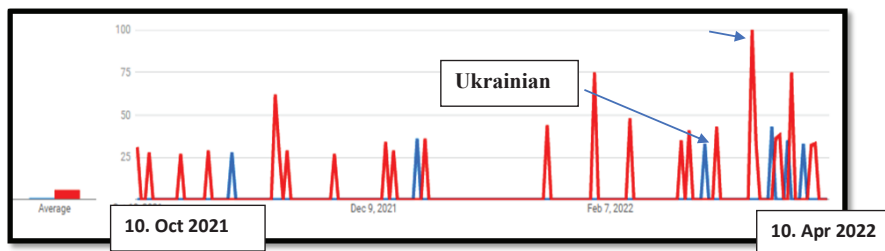
insights into this population. Using YouTube Analytics as a source of immigration data we start with an assumption that informing by watching videos on YouTube Platform worldwide about Croatia and its way of life gives indications of the intention to migrate to this country. For this purpose, we analysed the search for interests related to life in Croatia, videos about video lessons that offer the opportunity to learn Croatian, etc. In the case of wealthy countries, the presumption is that the interests of users are focused on spending vacations in Croatia, while such searches from users in developing countries indicate stronger indications of migration planning.

**Table 9:** Search queries „learning Croatian“, “life in Croatia”, and “travel to Croatia” on YouTube according to the geolocations worldwide (2017-2023)

Language + Croatia		Life in Croatia	Travel to Croatia
1	Nepal	Nepal	Ireland
2	Ethiopia	Nigeria	United Kingdom
3	Ghana	Bangladesh	United Arab Emirates
4	Nigeria	India	Australia
5	Kenya	Pakistan	South Africa
6	Somalia	-	Canada
7	Tanzania	-	-
8	Sri Lanka	-	-
9	Burma	-	-
10	Pakistan	-	-

As in the case of Google search analysis, YouTube Analytics also shows that the most common search interests come from Nepal and India, but also many African countries.

Figure 12 shows a high increase in interest in YouTube searches in Ukrainian and Russian after the war outbreak in Ukraine, which is again an indication that confirms the validity of such measurements.



**Figure 12.** YouTube searches in Ukrainian “Хорватія” and Russian “Хорватия” (Croatia) in Croatia from 10. October 2021 to 10. April 2022

Note: Ukrainian refugee uses both Ukrainian and Russian during Internet searches. We have explained this phenomenon in the section Limitations. It can be noticed that the activity is primarily in the coastal part of Croatia, mainly in Istria, Zadar, Šibenik Split, Zagreb and Karlovac and Varaždin County which overlaps with Croatian media reports on the distribution of refugees in the country.<sup>69</sup>



**Figure 13.** Overlap of YouTube data and official data on the prevalence of Ukrainian refugees in Croatia

Note: darker blue indicates a higher percentage of Ukrainian-language searches in Croatia.

<sup>69</sup> Slobodna Dalmacija, 2022. Ukrajinci smješteni u Splitu mahom su se počeli vraćati u domovinu, <https://slobodnadalmacija.hr/split/ukrajinci-smjesteni-u-splitu-mahom-su-se-poceli-vracati-u-domovinu-nisu-mogli-izdržati-bili-su-ovdje-fizicki-ali-u-mislama-tamo-no-neki-su-se-kod-nas-od-licno-snasli-1226311> (15.11.2022).

## VI. Conclusion

This research paper shows that here presented Big (Migration) Data approach is useful for determining immigration flows to Croatia, as well as for estimating and forecasting. This approach is enabling first insights into the identification of migrants' intentions to immigrate to Croatia and refugee flows - when there is no official data or conflicting data output. The analysis of different official Croatian statistic databases from 2018 to 2020 shows a difference of + 39,263 foreign workers: 2018 (+16.17%), 2019 (+43.54%), and 2020 (+36.64%). A particularly problematic issue in the official data recording is that Croatia does not record immigrants from Asia and Africa according to the country of immigration but only states the continent.

The methodological concept of here presented approach is to monitor the so-called "digital trace" of immigrants and refugees left on Google search and social networks Facebook (FB), Instagram, and YouTube and their geo-locations. We focus on users that use social network platforms in foreign languages in Croatia as well as Google search with specific migration-related queries worldwide.

There was a clear correlation between the FB and Instagram index and immigration flows to Croatia. The Big (Migration) Data approach shows a much higher number of immigrants in Croatia than official data. This approach shows that the number of FB and Instagram users in specific Asian and African languages has been growing rapidly in Croatia since 2017, as well as in Ukrainian (and Russian) after the war outbreak in Ukraine. According to the analysis of Internet traces, the countries from which the largest immigration to Croatia can be expected are Asian countries: Nepal, Philippines, and India, and the African countries: Ghana and Kenya, while from the neighbouring countries, it is North Macedonia.

The Big (Migration) Data correlates with official Croatian statistics regarding the distribution, age, and sex of foreign workers. The analysis of the FB groups of Asian and African immigrants in Croatia shows that those FB groups can be a valuable source for studying migration. One of the contributions of the second used method, Google search, and YouTube insights, is that it shows that by searching for specific migration-related queries and video material on the YouTube platform, the intention of users to migrate to the EU i.e., Croatia can be estimated. The analysis of digital traces shows that potential immigrants to Croatia, when planning their migrations, most often search for the terms "job + Croatia", and "salary + Croatia". Those searches were particularly intensive after Croatia increased the annual quotas for the import of foreign workers after 2017.

This approach can also be used in security studies because it provides clues about the movement of illegal migrants. The increase in searches for the term "Croatia" in the Arabic language in BiH corresponds to increased illegal entries into Croatia during 2019, while with the onset of the pandemic in 2020, a decrease in activity can be clearly seen, but also a re-increase in early 2022. These insights also provide the geolocation of potential migrants because they locate them around the town of Bihać in BiH, which is located on the border with the Republic of Croatia, i.e. the Schengen border.

The main advantage of this approach is the timely identification of immigration trends, which provides insights into migration patterns one year before official data becomes available. Therefore they can be used to model projections and predict further trends. Furthermore, they glean numerous insights that official statistics often fail to collect.

Since 2023 Croatia represents the external Schengen border of the EU, the increased immigration to Croatia is of interest to the European decision-makers - especially since once foreign workers from Asia and Africa enter Croatia, they can freely move anywhere within the Schengen Area. On the one side, this method allows governments to estimate how many migrants and refugees entered or intend to enter Croatia and Croatia's labour market and, on the other, to better respond to the recent Ukrainian humanitarian crisis.

## Sažetak

### Useljavanje u Hrvatsku: Pristup *Big (Migration) Data*

Rad obrađuje fenomen povećanog trenda useljavanja stranih državljana u Republiku Hrvatsku, postavlja pitanje valjanosti službenih podataka o broju stranih državljana u Hrvatskoj i pruža empirijsku osnovu za razvoj modela *Big (Migration) Data* za predviđanje budućih trendova imigracije u Hrvatsku. Postavlja se hipoteza da bi taj pristup mogao biti od koristi za određivanje migracijskih tokova prema Hrvatskoj te za procjenu i prognoziranje budućih trendova. Navedeni pristup omogućuje prve uvide u namjeru migranata da imigriraju u Hrvatsku te modeliranje trendova migracijskih i izbjegličkih tokova. Metodološki koncept ovdje razvijenog pristupa jest praćenje tzv. digitalnih tragova imigranata i izbjeglica ostavljenih na *Google* pretraživaču te društvenim mrežama *Facebook* (FB), *Instagram* i *YouTube* uključujući geolokacije. Fokusiramo se na korisnike koji se koriste društvenim mrežama na stranim jezicima u Hrvatskoj te na *Google* pretraživanje diljem svijeta. Rezultati: Analiza pokazuje nerazmjernost i nepouzdanost podataka o imigraciji koji dolaze iz Ministarstva unutarnjih poslova Republike Hrvatske, Državnog zavoda za statistiku i Hrvatskog zavoda za zapošljavanje. Pristup *Big (Migration) Data* pokazuje znatno veći broj imigranata u Hrvatskoj nego što sugeriraju službeni podaci. Taj pristup pokazuje da u Hrvatskoj rapidno raste broj korisnika *Facebooka* i *Instagrama* na određenim azijskim i afričkim jezicima te na ukrajinskom (i ruskom) nakon izbijanja rata u Ukrajini. Testirani indeks *Facebooka* i *Instagrama* korelira s pojačanom imigracijom. K tome analiza *Facebook* grupa azijskih i afričkih imigranata u Hrvatskoj pokazuje da te grupe mogu biti vrijedan izvor za proučavanje migracija. Jedan od doprinosa drugoga korištenog pristupa, *Google* pretraživanja i *YouTube* uvida, jest to da pokazuje da se pretragom određenih migracijskih upita i videomaterijala može procijeniti namjera korisnika da migriraju u Europsku uniju, odnosno Hrvatsku. Uvidi dobiveni tom metodom poklapaju se sa službenim podacima (vezano uz trendove), koji pak dolaze s godinom dana zakašnjenja. Time je navedena metoda od osobite koristi za pravovremene procjene broja imigranata i modeliranje trendova useljavanja u Hrvatsku.

*Ključne riječi:* imigracija; EU; Hrvatska; Big Data; strani radnici; izbjeglice.