ESTIMATING MIGRATION FLOWS USING ONLINE SEARCH DATA

PARTNER: UNITED NATIONS POPULATION FUND (UNFPA)
PROGRAMME AREA: ECONOMIC WELLBEING

SUMMARY

This study was conducted in partnership with the United Nations Population Fund (UNFPA) to explore how online search data could be analysed to understand migration flows. Using Australia as a case study, Google search query data from around the world was disaggregated by country and compared to historical official monthly migration statistics provided by UNFPA. Correlations were observed between relevant search queries (for example, searching for 'jobs in Melbourne') and official migration statistics (number of people who migrated to Melbourne). In particular, queries from specific locations in Australia related to local employment opportunities showed highest correlation. The research findings point toward new possibilities for further exploration into using online and other digital search data as proxy for migration statistics.

BACKGROUND

Information on the spatial distribution and population movement is an integral aspect of global development. Migrants contribute to local development by transferring both cultural and economic capital. Remittances are one of the most tangible and measurable outcomes of migration. By working abroad and sending money to their families back home, migrant workers play an important role in enhancing development and reducing poverty in their countries of origin, as well as contributing to the economic prosperity of their countries of destination.

Reliable and timely information on international migration flows is critical for developing policies that represent the needs of the population, from transportation to social services, programme planning and implementation, and monitoring and evaluation. However, there are many inconsistencies in international migration data. (UNSD, 2004) It is often out of date, consolidated on a yearly basis. Definitions of migrants are inconsistent between countries. Demographic information, especially on gender, is rarely tracked. In some cases, migration data is simply non-existent.

To address these challenges, a new field of migration research is emerging. Big data presents new opportunities for understanding migration through these 'digital traces.' It is hypothesised that migration flows can be inferred from volumes of online search queries implying the intent or interest in migration. Previous work has had proven success in using search volumes to predict flu and Dengue fever outbreaks, economic indicators and unemployment.

The United Nations Population Fund (UNFPA) is committed to working with governments, the UN System, non-governmental organizations and civil society to improve migration data by building capacity to plan and implement policies and programmes and strengthening partnerships to enhance understanding of the complexity of migration flows and their links to development. Global Pulse conducted this study in partnership with UNFPA and in alignment with these objectives.

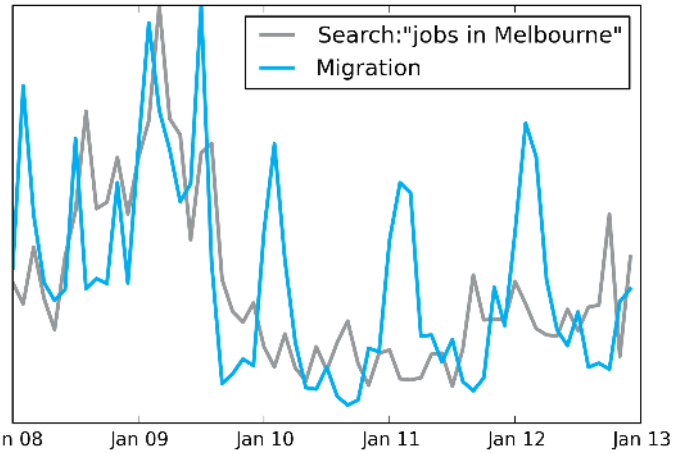
HOW TO CITE THIS DOCUMENT:
UN Global Pulse, 'Estimating Migration Flows Using Online Search Data ', Global Pulse Project Series no. 4, 2014.

ANALYZING ONLINE SEARCH QUERIES TO COMPLEMENT MIGRATION DATA

This study explores the feasibility of using online search data for insights on migration, and if search data can serve as proxies for migration indicators to augment existing information gathered by household survey or other official statistical analysis.

This research analysed search queries against official data on migrant arrivals into Australia from January 2008 to December 2013, disaggregated by country. The query data was extracted from the Google search engine. Monthly volumes of searches were collected for a set of queries relevant to employment (such as 'work visa' and 'jobs in Melbourne') in the same time period as the migration statistics.

Relevant search queries that were derived from Google Correlate, a tool used to find search terms that match a time series, and originated in Australia from 2008 to 2013 were analysed over time and correlated with monthly migration data from the Australian Bureau of Statistics.

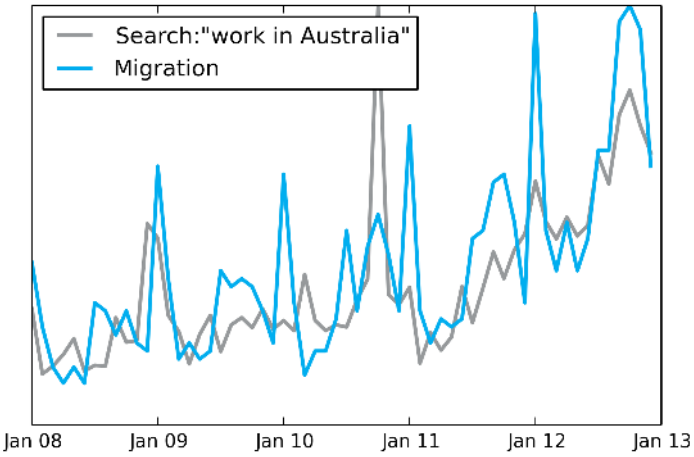


The graph above shows the migration trend from India to Australia from January 2008 to December 2013 (blue line) and Google search activity from India for the query 'jobs in Melbourne' (grey line).

INSIGHTS & OUTCOMES

The outcomes of this study highlight the potential relationships between online search queries and employment and job opportunities.

- **Online searches are a potential proxy indicator of employment opportunities:** In many cases, a clear correlation was found between search queries relating to specific places in Australia and actual job opportunities. This outcome could help augment official economic statistics about where employment opportunities are located.
- **Search queries can help measure the underemployed:** A strong correlation was found in how people in different countries search for job opportunities, particularly across regions. This could support efforts to estimate not only the number of people who are unemployed, but also people that are underemployed (official statistics often do not account for factors such as partial employment or job dissatisfaction).
- **Local searches correlated highest with employment statistics:** Queries from specific locations related to employment opportunities in those same locations showed the highest correlation to official migration statistics. This insight could help local governments assess unemployment more frequently, instead of waiting for annual survey data, which could drive more proactive and informed policy decisions.
- **Online search as proxy for migration statistics:** The results of this study demonstrate the potential for online search volumes to be used as proxy for migration statistics. This implies that people interested in migrating conduct online searches to explore employment just prior to migrating, and thus search data could be used as proxy for intent to migrate.



The graph above shows the trend in actual migration from Italy to Australia from January 2008 to December 2013 (blue line) and Google search activity from Italy for the query 'work in Australia' (grey line). The correlation value for migration from Italy to Australia with search query 'work in Australia' is r=0.74, p<0.001.

It is important to note that confounding factors complicate the analysis; an increase in relevant search queries does not always translate into reliable migration information. For example, citizens of other countries are not all equally able to migrate. Visa regulations differ, quotas may apply and individuals may be dissuaded from migrating for reasons such as distance or lack of an established expatriate community.

CONCLUSIONS

The results of this study demonstrated that online search data could complement official migration statistics, or even serve as a proxy indicator, helping reveal migration flows and insights such as underemployment.

While the outcomes of this research are promising, further research and more detailed projects are recommended to strengthen the case for using online search data in the international migration development and policy context. It is recommended to replicate this study with other destination countries that can provide comparable ground truth data. In addition, the ground truth data could be improved by disaggregating the purpose of the migration event, providing a breakdown of the number of entrants holding educational and work visas or seeking asylum. These measures would help test the robustness of the research results and refine the methodology.

By partnering with companies like Google, Baidu and Bing to leverage resources like online search data for public good, the possibilities of using big data for challenges such as measuring unemployment can be expanded. In working with UN Global Pulse for this study, UNFPA gained an increased understanding of the challenges, opportunities and implications associated with big data for development.

This study of migration flows to Australia showed both the potential utility and limits of using online search data to help improve the reliability, accessibility and frequency of migration data.

IMPLICATIONS & RECOMMENDATIONS

- This study evidenced the potential of online search data, specifically online search queries, to be used as proxy indicator for migration statistics, specifically in revealing job opportunities and underemployment.
- Further research is recommended to test this methodology in other country contexts, and disaggregate (? Not clear) the official migration data used to further refine the results

REFERENCES

United Nations Demographic Yearbook Review (2004) United Nations Department of Economic and Social Affairs Statistics Division, Demographic and Social Statistics Branch. http://unstats.un.org/unsd/demographic/products/dyb/techreport/migration.pdf

Using Web Search Query Data to monitor Dengue Epidemics: A New Model for Neglected Tropical Disease Surveillance (2011) E. Chan et al PLoS One

Predicting the Present with Google Flu Trends (2012) H. Choi and H. Varian Economic Record

Quantifying the Advantage of Looking Forward (2012) T. Preis et al Nature Scientific Reports

Google Econometrics and Unemployment Forecasting (2009) A. Nikolaos and Z. Klaus