AI for Social Good: Social Media Mining of Migration Discourse

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

DOKTOR DER NATURWISSENSCHAFTEN

Dr. rer. nat.

genehmigte Dissertation von

M. Sc., M. Tech. Aparup Khatua

2023

Referent: Prof. Dr. techn. Dipl.-Ing. Wolfgang Nejdl, Leibniz Universität Hannover

Korreferent: Prof. Dr. Prasenjit Mitra, Penn State University

Tag der Promotion: 11-Sep-2023

Acknowledgment

This Ph.D. was an exciting and rewarding journey. I was fortunate to come across wonderful mentors and colleagues over the last few years. I would like to thank my Ph.D. supervisor, Prof. Wolfgang Nejdl. I had the opportunity to interact with him on multiple occasions that started from my Ph.D. interview to weekly research seminars to co-authoring papers. He was extremely patient and supportive during these interactions. I am thankful to him for giving me the liberty to select my research trajectory. His constant encouragement and guidance immensely helped me in this journey – an academic version of *Les Rites de Passage*.

My research papers that are not part of this dissertation have also helped me to evolve as a researcher. Hence, I would also like to convey my heartfelt thanks to my past and present collaborators and mentors. Specifically, I would like to mention Prof. Erik Cambria, Nanyang Technological University (NTU) Singapore, Singapore and Prof. Ingmar Weber, Saarland University, Germany. I would also like to thank my colleagues at L3S Research Centre. My interactions with them made this journey joyful, specifically during the depressing Covid-19 crisis.

Finally, I would like to dedicate this thesis to my family, who encouraged me to pursue this arduous journey. Most importantly, I would like to thank my brother Apalak Khatua, who was a constant source of support.

Aparup Khatua

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AI4SG	– AI for Social Good
AUC	– Area Under the ROC Curve
BERT	– Bidirectional Encoder Representations from Transformers
Bi-LSTM	– Bidirectional LSTM
BS	– Batch Sizes
CNN	– Convolutional Neural Network
DL	– Deep Learning
DR	– Dropout Rate
F1	– F1 Score
GloVe	– Global vectors for word representation
IOM	– International Organization for Migration
LR	– Learning Rates
LSTM	– Long Short-Term Memory
ML	– Machine Learning
NGO	– Non-Government Organization
NLI	– Natural Language Inference
NLP	– Natural Language Processing
PR	– Precision
RC	– Recall
RNN	– Recurrent Neural Networks
RoBERTa	– Robustly Optimized BERT Pre-training Approach
SGD	– Sustainable Development Goals
SGD	– Stochastic Gradient Descent
UN	– United Nations
UNHCR	– United Nations High Commissioner for Refugees
ZSLM	– Zero-Shot Learning Models

Abstract

The number of international migrants has steadily increased over the years, and it has become one of the pressing issues in today's globalized world. Our bibliometric review of around 400 articles on Scopus platform indicates an increased interest in migration-related research in recent times but the extant research is scattered at best. AI-based opinion mining research has predominantly noted negative sentiments across various social media platforms. Additionally, we note that prior studies have mostly considered social media data in the context of a particular event or a specific context. These studies offered a nuanced view of the societal opinions regarding that specific event, but this approach might miss the forest for the trees. Hence, this dissertation makes an attempt to go beyond simplistic opinion mining to identify various latent themes of migrant-related social media discourse.

The first essay draws insights from the social psychology literature to investigate two facets of Twitter discourse, i.e., perceptions about migrants and behaviors toward migrants. We identified two prevailing perceptions (i.e., sympathy and antipathy) and two dominant behaviors (i.e., solidarity and animosity) of social media users toward migrants. Additionally, this essay has also fine-tuned the binary hate speech detection task, specifically in the context of migrants, by highlighting the granular differences between the perceptual and behavioral aspects of hate speech.

The second essay investigates the journey of migrants or refugees from their home to the host country. We draw insights from Gennep's seminal book, i.e., Les Rites de Passage, to identify four phases of their journey: Arrival of Refugees, Temporal stay at Asylums, Rehabilitation, and Integration of Refugees into the host nation. We consider multimodal tweets for this essay. We find that our proposed theoretical framework was relevant for the 2022 Ukrainian refugee crisis – as a use-case.

Our third essay points out that a limited sample of annotated data does not provide insights regarding the prevailing societal-level opinions. Hence, this essay employs unsupervised approaches on large-scale societal datasets to explore the prevailing societal-level sentiments on YouTube platform. Specifically, it probes whether negative comments about migrants get endorsed by other users. If yes, does it depend on who the migrants are – especially if they are cultural others? To address these questions, we consider two datasets: YouTube comments before the 2022 Ukrainian refugee crisis, and during the crisis. Second dataset confirms the Cultural Us hypothesis, and our findings are inconclusive for the first dataset.

Our final or fourth essay probes social integration of migrants. The first part of this essay probed the unheard and faint voices of migrants to understand their struggle to settle down in the host economy. The second part of this chapter explored the viability of social media platforms as a viable alternative to expensive commercial job portals for vulnerable migrants.

Finally, in our concluding chapter, we elucidated the potential of explainable AI, and briefly pointed out the inherent biases of transformer-based models in the context of migrant-related discourse. To sum up, the importance of migration was recognized as one of the essential topics in the United Nation's Sustainable Development Goals (SDGs). Thus, this dissertation has attempted to make an incremental contribution to the AI for Social Good discourse.

Keywords: Social Media, Migration, AI for Social Good

Zusammenfassung

Die Zahl der internationalen Migranten hat im Laufe der Jahre stetig zugenommen und ist zu einem der drängendsten Probleme in der heutigen globalisierten Welt geworden. Unsere bibliometrische Analyse von rund 400 Artikeln auf der Scopus-Plattform zeigt, dass das Interesse an migrationsbezogener Forschung in letzter Zeit zugenommen hat, aber die vorhandenen Forschungsergebnisse sind bestenfalls dünn gestreut. Bisher hat die KI-basierte Sentimentanalyse vor allem negative Stimmungen auf verschiedenen Social-Media-Plattformen aufgezeigt. Darüber hinaus stellen wir fest, dass frühere Studien Social-Media-Daten meist im Zusammenhang mit einem bestimmten Ereignis oder in einem spezifischen Kontext betrachtet haben. Diese Studien bieten einen nuancierten Blick auf die gesellschaftlichen Meinungen zu einem bestimmten Ereignis, aber dieser Ansatz könnte dazu führen, dass größere Zusammenhänge übersehen werden. In dieser Dissertation wird daher versucht, über eine einfache Sentimentanalyse hinauszugehen und verschiedene latente Themen des migrationsbezogenen Diskurses in den sozialen Medien zu identifizieren.

Der erste Aufsatz stützt sich auf Erkenntnisse aus der sozialpsychologischen Literatur um zwei Facetten des Twitter-Diskurses zu untersuchen, nämlich Einstellungen zu und Verhaltensweisen gegenüber Migranten. Wir haben zwei vorherrschende Einstellungen (d. h. Sympathie und Antipathie) und zwei vorherrschende Verhaltensweisen (d. h. Solidarität und Feindseligkeit) von Nutzern sozialer Medien gegenüber Migranten ermittelt. Darüber hinaus wird in diesem Aufsatz die binäre Aufgabe zur Erkennung von Hassrede speziell im Kontext von Migranten verfeinert, indem die feinen Unterschiede zwischen den Einstellungs- und Verhaltensaspekten von Hassreden herausgestellt wurden.

Der zweite Aufsatz untersucht die Reise von Migranten und Flüchtlingen von ihrem Heimatland in das Aufnahmeland. Wir ziehen Erkenntnisse aus Genneps einflussreichem Buch, Les Rites de Passage, heran, um vier Phasen ihrer Reise zu identifizieren: Ankunft der Flüchtlinge, vorübergehender Aufenthalt im Asyl, Rehabilitation und Integration der Flüchtlinge in das Gastland. Wir betrachten multimodale Tweets für diesen Aufsatz. Wir stellen fest, dass der von uns vorgeschlagene theoretische Rahmen für die ukrainische Flüchtlingskrise im Jahr 2022 als Anwendungsfall relevant ist.

Unser dritter Aufsatz zeigt auf, dass eine begrenzte Stichprobe von kommentierten Daten keinen Einblick in die vorherrschenden Meinungen auf gesellschaftlicher Ebene bietet. Daher werden in diesem Aufsatz unüberwachte Ansätze auf großen gesellschaftlichen Datensätzen angewandt, um die vorherrschenden gesellschaftlichen Stimmungen auf der Plattform YouTube zu untersuchen. Besonderes Augenmerk wird darauf gelegt ob negative Kommentare über Migranten von anderen Nutzern gebilligt werden. Wenn dies der Fall ist betrachten wir zusätzlich ob diese Billigung von der Identität der Migranten abhängt - vor allem, wenn es sich um Kulturschaffende handelt. Um diese Fragen zu beantworten, betrachten wir zwei Datensätze: YouTube-Kommentare vor der ukrainischen Flüchtlingskrise 2022 und während der Krise. Der zweite Datensatz bestätigt die Cultural-Us-Hypothese, während unsere Ergebnisse für den ersten Datensatz nicht schlüssig sind.

Unser letzter bzw. vierter Aufsatz befasst sich mit der sozialen Integration von Migranten. Der erste Teil dieses Aufsatzes untersuchte die ungehörten und leisen Stimmen der Migranten, um ihre Anstrengungen zu verstehen, in der Wirtschaft des Aufnahmelandes anzukommen. Im zweiten Teil dieses Aufsatzes wird untersucht, inwieweit Social-Media-Plattformen eine brauchbare Alternative zu teuren kommerziellen Jobportalen für gefährdete Migranten darstellen.

In unserem abschließenden Kapitel erläutern wir das Potenzial erklärbarer KI und weisen kurz auf die inhärenten Bias Transformer-basierter Modelle im Kontext des migrationsbezogenen Diskurses hin. Zusammenfassend lässt sich sagen, dass die Bedeutung der Migration als eines der wichtigsten Themen in den Zielen für nachhaltige Entwicklung der Vereinten Nationen (SDGs) anerkannt wurde. Daher versucht diese Dissertation, einen schrittweisen Beitrag zum Diskurs über KI für soziales Wohlbefinden zu leisten.

Schlagwoerter: Soziale Medien, Migration, KI für das Gemeinwohl

Chapter 1: Migration: An AI-based Approach

1.1 From AI to AI for Social Good (AI4SG)

The genesis of the Artificial Intelligence (AI) domain was triggered by Alan Turing's seminal article *Computing Machinery and Intelligence* in 1950. Interestingly, he published this article not in a traditional computer science journal but in *Mind – A Quarterly Review of Psychology and Philosophy*. The journal selection emphasizes the AI domain's interdisciplinary nature from its inception. In this seminal article, Turing raised the question - "Can machines think?" and introduced the concept of a *learning machine* (Turing, 2009). However, he was fully aware that "the idea of a learning machine may appear paradoxical to some readers," and thus, he pointed out that "an important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil's behaviour" (Turing, 2009, p. 458). A few years later, Marvin Minsky (from Massachusetts Institute of Technology) and John McCarthy (from Stanford University) officially coined the phrase *Artificial Intelligence (AI)*. Turing, Minsky, and McCarthy are considered the founders of the AI domain.

The initial successes, such as solving the *Tower of Hanoi* problem or developing ELIZA(an early natural language processing computer program created by the Massachusetts Institute of Technology AL Laboratory), have elucidated the immense potential of AI. For instance, ELIZA employed the *pattern matching* logic to communicate with a human being for the first time. This was followed by a series of pathbreaking theoretical works such as Rosenblatt's perceptron model, the first neural network model in 1958, or the application of backpropagation for recognizing handwritten Zip Code by LeCun in 1989. On the application front, some notable achievements were the industrial robot arm for the General Motors plant and the first mobile robot (from Stanford Research Institutes) to interpret instructions in the early 1960s (Kaul et al., 2020).

Around this time, the medical fraternity also started to explore various AI-based applications. Some of the initial AI-based applications were the development of MYCIN to "provide a list of potential bacterial pathogens and then recommend antibiotic treatment options" in 1972 or the consultation program for glaucoma using the CASNET model, i.e., causal–associational network, in 1976 (Kaul et al., 2020, p. 809). Probably, the domain of AI attracted the attention of popular media when IBM developed its Deep Blue Computer in 1997. Deep Blue's chess program "was able to beat the world champion Gary Kasparov ... Deep Blue was reportedly able to process 200 million possible moves per second and to determine the optimal next move looking 20 moves ahead through the use of a method called tree search" (Haenlein & Kaplan, 2019, p. 8). Since then, the AI domain has attracted the attention of academia, corporates, and regulators.

Over the years, researchers realized that all is not well with AI – it may have potentially adverse consequences if not used properly. For instance, Algorithmic Justice League[1], a USA-based non-profit organization, warns that AI-based applications can potentially "harm vulnerable and marginalized people, and threaten civil rights. Unchecked, unregulated and, at times, unwanted, AI systems can amplify racism, sexism, ableism, and other forms of discrimination" (¶ 1). For instance, the face recognition algorithm has many practical applications, from unlocking a smartphone to tagging families and friends on the Facebook platform. However, these face recognition algorithms are mostly biased even today. Klare et al. (2012) explored six different face recognition algorithms for eight cohorts based on gender (male and female), race/ethnicity (Black, White, and Hispanic), and age group (18 to 30, 30 to 50, and 50 to 70 years old). They observed that these algorithms have lower accuracies for cohorts like females, Blacks, and aged 18 to 30, and they primarily associated these lower accuracies with biased training data. The seminal study by Buolamwini & Gebru (2018, p. 1) also explored publicly available datasets and pointed out that these "datasets are overwhelmingly composed of lighter-skinned subjects." Interestingly, a follow-up audit of the previous research revealed that target companies released new API versions, and these companies "reduced accuracy disparities between males and females and darker and lighter-skinned subgroups" (Raji & Buolamwini, 2019, p. 429). Overall, algorithmic biases can have serious adverse consequences. For example, Google's photo app became controversial when it labeled a black couple as gorillas. Google immediately apologized for the racist blunder of its algorithm[2]. Similarly, a report by ProPublica[3] also noted racial biases in predicting future/potential criminals. This report found that accuracies are broadly the same for black and white defendants, but unfortunately,

- "the formula was particularly *likely to falsely flag black defendants as future criminals* (emphasis added), wrongly labeling them this way at almost twice the rate as white defendants" (¶ 15)

[1] Available at https://www.ajl.org/, Accessed on July 1, 2022
[2] Available at https://www.bbc.com/news/technology-33347866, Accessed on July 1, 2022
[3] Available at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, Accessed on July 1, 2022

– “*white defendants were mislabeled as low risk* (emphasis added) more often than black defendants” (¶ 15)

Not only face recognition algorithms but also large language models (LLMs) are biased. Applications of large LMs gained momentum in the last few years, and it seems “institutions (are) seemingly competing to produce ever larger LMs” (Bender et al., 2021, p. 610). For instance, the training dataset size for initial models like BERT, DistilBERT, ALBERT, and ERNIE-Gen was 16 GB, but the dataset size became more than 100 GB for models like XLNet, RoBERTa, MegatronLM, and T-NLG. Later models, such as GPT-3 (570 GB), T5-11B (745 GB), or Switch-C (745 GB), are even bigger (Bender et al., 2021). However, a recent paper by Abid et al. (2021) published in *Nature* observes that GPT-3 associates Muslims with violence. The authors gave a sentence completion prompt - *Two Muslims walked into a …,* and 66% of completions indicate Muslims are committing violence. However, this percentage drastically comes down for other religious groups like Christians, Sikhs, Jews, or Buddhists. Bender et al. (2021, p. 610) also pointed out that “large datasets based on texts from the Internet overrepresent hegemonic viewpoints and encode biases potentially damaging to marginalized populations.”

In the medical domain, AI-based drug discovery has attracted the attention of researchers and pharmaceutical companies for faster speed and lower cost. A recent paper by Jayatunga et al. (2022), published in *Nature*, studied 20 ‘AI-native’ drug discovery companies and reconstructed their product pipeline from 2010 to 2021. Interestingly, “the combined pipeline of these 20 AI companies contains ~160 disclosed discovery programmes and preclinical assets and about 15 assets in clinical development” (Jayatunga et al., 2022, p. 175). This article concludes that AI-based approaches can be “a game-changer for pharmaceutical R&D, especially for small-molecule drug discovery” (Jayatunga et al., 2022, p. 176). Contrarily, another article from another *Nature* journal by Urbina et al. (2022, p. 189) gave “a wake-up call” and demonstrated that AI-based approaches for drug discovery “could be misused for de novo design of biochemical weapons.” Generally, AI-based drug-discovery models *penalize* predicted toxicity and *reward* predicted target activity. This study inverted this conventional logic and *rewarded toxicity and bioactivity and* explored it in the domain of nerve agent VX – “one of the most toxic chemical warfare agents developed during the twentieth century — a few salt-sized grains of VX (6–10 mg)[5] is sufficient to kill a person” (Urbina et al., 2022, p. 189). Shockingly, this tweaked model generated 40,000 molecules in less than 6 hours – it generated VX and *other known chemical warfare agents* more toxic than VX. This unnerving anecdote elucidates the darker side of AI. Thus, Cowls et al. (2021) and Floridi et al. (2020) warned that human values shape AI applications like any other technology. Hence, if we are not careful about the end objectives, we may face an unwanted *good-AI-gone-bad* situation. Floridi et al. (2020, p. 1172) further argued that there can be an *accidental success,* but "lacking a clear understanding of AI4SG means that this success is accidental and cannot be repeated systematically." Hence, we need to use AI for Social Good (AI4SG) consciously.

Subsequently, the paper by Cowls et al. (2021, p. 112), published in *Nature Machine Intelligence*, defined AI4SG "as the design, development and deployment of AI systems in ways that help to (i) prevent, mitigate and/or resolve problems adversely affecting human life and/or the wellbeing of the natural world, and/or (ii) enable socially preferable or environmentally sustainable developments, while (iii) not introducing new forms of harm and/or amplifying existing disparities and inequities." The intriguing question is – where to apply AI4SG? In response, another recent article by Tomašev et al. (2020, p. 1) in *Nature Communications* suggests,

> *" … artificial intelligence (AI) present an opportunity to build better tools and solutions to help address some of the world's most pressing challenges, and deliver positive social impact in accordance with the priorities outlined in the United Nations' 17 Sustainable Development Goals (SDGs). The AI for Social Good (AI4SG) movement aims to establish interdisciplinary partnerships centred around AI applications towards SDGs."*

1.2 UN Sustainable Development Goals (SDGs) and Migration

Neither Rome nor SDGs were built in a day. Hence, we need to understand the precursors to SDGs. According to the United Nations[4], the starting point was the 1992 Earth Summit in Rio de Janeiro, Brazil, where "more than 178 countries adopted *Agenda 21*, a comprehensive plan of action to build a global partnership for sustainable development to improve human lives and protect the environment" (¶ 2). Nearly a decade later, i.e., in 2000, *Millennium Development Goals (MDGs)* were adopted to reduce extreme poverty by 2015 at the Millennium Summit in New York. Subsequently, these goals were reemphasized in the 2002 World Summit on Sustainable Development in South Africa.

In the 2012 United Nations Conference in Rio de Janeiro, Brazil, member states agreed to develop *a set of SDGs* that would expand the scope of previous MDGs, and member

[4] Available at https://sdgs.un.org/goals, Accessed on July 1, 2022

states officially adopted the outcome document - *The Future We Want*. Next year, i.e., in 2013, a 30-member Open Working Group was formed to develop a proposal on the SDGs. The landmark moment came two years later. Member states adopted the 2030 Agenda for Sustainable Development, with 17 SDGs and 169 targets, at the UN Sustainable Development Summit in 2015. The United Nations[5] says,

> *"The 2030 Agenda for Sustainable Development … provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries - developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests*" (¶ 1).

A report titled *Migration and the SDGs: Measuring Progress*[6] by IOM pointed out that the importance of migration-related issues was recognized as an essential topic for the first time in this 2030 development agenda. Goal 10, out of 17 SDGs, aims to *reduce inequality within and among countries*. Specifically, from the perspective of migration, Target 10.7 emphasizes the need to *facilitate orderly, safe, regular, and responsible migration and mobility of people, including through the implementation of planned and well-managed migration policies*. Some of the proposed indicators of Target 10.7 are as follows:

- *Indicator 10.7.1:* Recruitment cost borne by employee as a proportion of monthly income earned in country of destination
- *Indicator 10.7.2:* Number of countries with migration policies that facilitate orderly, safe, regular and responsible migration and mobility of people
- *Indicator 10.7.3:* Number of people who died or disappeared in the process of migration toward an international destination
- *Indicator 10.7.4:* Proportion of the population who are refugees, by country of origin

However, associating only Target 10.7 of SDGs with migration issues might be myopic. Migration issues can have an overwhelming influence on ensuring sustainable

[5] Available at https://sdgs.un.org/goals, Accessed on July 1, 2022
[6] Available at https://publications.iom.int/system/files/pdf/SDG-an-edited-volume.pdf, Accessed on July 1, 2022

development outcomes at the global level. Accordingly, the Migration Data Portal[7] also says,

> *"The 2030 Agenda for Sustainable Development recognizes for the first time the contribution of migration to sustainable development. Migration is a cross-cutting issue in the 2030 Agenda, relevant to all of the Sustainable Development Goals (SDGs). Further, the SDG's motto to 'leave no one behind' is a clear call for sustainable development to be inclusive, including for migrants. At least ten out of 17 goals contain targets and indicators that are directly relevant to migration or mobility"* (¶ 1).

Considering the SDG's overarching motto to *leave no one behind*, Table 1.1 provides a mapping between various SDG Goals and Migration issues. It is evident that migrants not only face inequalities (SDG 10), but also, they struggle with poverty (SDG 1) in the host country. Often, they get discriminated against in terms of employment opportunities (SDG 8). Consequently, many of them can't afford proper accommodation and primarily stay in crowded households (SDG 11). Adult migrants might not have access to health facilities (SDG 3). Similarly, young migrants don't get education opportunities (SDG 4) during their stay in temporal asylums and even after that. Literature suggests that negative perceptions are more associated with male migrants. Female migrants are more likely to get exploited at workplaces (SDG 5). Generally, migrants are vulnerable to sexual exploitation or face trafficking risks (SDG 16). Thus, policymakers need to work in coordination to address these inequalities (SDG 17). Tomašev et al. (2020, p. 2) also emphasized this interconnectedness of the SDGs and argued that "principle for fair and inclusive AI for social good: AI applications should aim to maximise a net positive effect on as many SDGs as possible, without causing avoidable harm to other SDGs." This justifies the selection of migration-related deliberations on social media platforms, as an application domain, for taking a baby step from AI to AI4SG.

Following prior research, such as Bondi et al. (2021), Cowls et al. (2021), Floridi et al. (2020), and Tomašev et al. (2020), this dissertation aims to develop and deploy AI-based approaches to explore issues adversely affecting human life, especially from the perspective of migrants or refugees – one of the world's most pressing issues. Following SDG 10, the overarching theme of this dissertation is to investigate the underlying reasons behind inequalities faced by migrants. This dissertation makes a humble attempt to

[7] Available at https://www.migrationdataportal.org/sdgs?node=0, Accessed on July 1, 2022

contribute to the vibrant AI4SG literature. Hopefully, the findings of this dissertation will have a positive social impact in accordance with the SDGs.

Table 1.1 Mapping SDGs and Migration Issues

SDGs	Targets by 2030	Migrant-specific Aspects
SDG 1: End Poverty in all its form everywhere	– Eradicate extreme poverty for all people everywhere, currently measured as people living on less than USD 1.25 a day – Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable	– 35% of migrants and 23% of non-migrants were on average in or at risk of poverty in 36 countries in 2015 – Literature and our pilot study pointed out the poverty and misery of migrants
SDG 3: Good Health and Well-being	– Reduce the global maternal mortality ratio to less than 70 per 100,000 live births – End preventable deaths of newborns and children under 5 years of age – End the epidemics of AIDS, tuberculosis, malaria, and neglected tropical diseases and combat hepatitis, water-borne diseases, and other communicable diseases – Reduce by one-third premature mortality from non- communicable diseases through prevention and treatment and promote mental health and well-being – Substantially increase health financing and the recruitment, development, training, and retention of the health workforce in developing countries, especially in LDCs and SIDS	– In 111 countries in 2018-2019, 90% of governments say they provide essential and emergency health care to all non-nationals, regardless of their migratory status. – However, migrant-related literature strongly indicates that migrants don't have access to proper healthcare facilities and the stress level is high for both young and adult migrants – Hence, it is crucial to ensure healthy lives and promote well-being for all
SDG 4: Quality Education	– Ensure that all girls and boys complete free, equitable, and quality primary and secondary education leading to relevant and effective learning outcomes – Ensure equal access for all women and men to affordable and quality technical, vocational and tertiary education, including university – Ensure that all youth and a substantial proportion of adults, both men, and women, achieve literacy and numeracy – Substantially expand globally the number of scholarships available to developing countries, in particular LDCs, SIDS, and African countries, for enrolment in higher education, including vocational training and ICT, technical, engineering, and scientific	– In 2017, 19% of foreign-born people aged 18 to 24 left school early compared to 10% of native-born people in the EU. – Prior studies, as well as Chapters 4 and 6 of this dissertation, have also pointed out the struggles faced by migrant kids or adult migrants to learn a new skill or new language in foreign countries

	programs, in developed countries and other developing countries	
SDG 5: Gender Equality	– Ensure women's full and effective participation and equal opportunities for leadership at all levels of decision-making in political, economic, and public life	– Only 69% of 111 countries have mechanisms to endure that migration policies are gender-responsive (2021) – Literature pointed out the need to empower migrant women and girls – Interestingly, Chapter 4 of this dissertation highlights that the perception of male migrants is often negative
SDG 8: Decent Work and Economic Growth	– Promote development-oriented policies that support productive activities, decent job creation, entrepreneurship, creativity, and innovation, and encourage the formalization and growth of micro-, small- and medium-sized enterprises, including through access to financial services – Achieve full and productive employment and decent work for all women and men, including for young people and persons with disabilities, and equal pay for work of equal value – Substantially reduce the proportion of youth not in employment, education, or training – Take immediate and effective measures to eradicate forced labor, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labor, including recruitment and use of child soldiers, and by 2025 end child labor in all its forms – Protect labor rights and promote safe and secure working environments for all workers, including migrant workers, in particular, women migrants, and those in precarious employment – Strengthen the capacity of domestic financial institutions to encourage and expand access to banking, insurance, and financial services for all	– Compared to 2 out of 100,000 native-born, 5 out of 100,000 foreign-born experienced a fatal injury at work in 2015 in selected European countries. – Chapter 6 of this dissertation also highlighted the job-related inequality/unemployment of foreign-born
SDG 10: Reduced Inequality	– Empower and promote the social, economic, and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status	– Probably, this is the overwhelming issue in the context of migrants – Extant literature, irrespective of their domains and disciplines,

	– Ensure equal opportunity and reduce inequalities of outcome, including by eliminating discriminatory laws, policies, and practices and promoting appropriate legislation, policies, and action in this regard – Facilitate orderly, safe, regular and responsible migration and mobility of people, including through the implementation of planned and well-managed migration policies – Reduce to less than 3 percent the transaction costs of migrant remittances and eliminate remittance corridors with costs higher than 5 percent	has unanimously pointed out this aspect – Findings from all the chapters of this dissertation have strongly indicated the inequality faced by the migrants across all spheres of their life
SDG 11: Sustainable Cities and Communities	– Ensure access for all to adequate, safe, and affordable housing and essential services and upgrade slums	– 30% of foreign citizens in European countries lived in overcrowded households in 2018, compared to 18% of citizens – In other words, it is a long journey from temporal asylums to a safe settlement
SDG 13: Climate Action	– Promote mechanisms for raising capacity for effective climate change-related planning and management in the least developed countries and small island developing States, including focusing on women, youth, and local and marginalized communities	– In 2021, 30.7 million total new displacements were triggered by climate-related disasters – So, we need to take urgent action to combat climate change and its impacts
SDG 16: Peace, Justice, and Strong Institutions	– Significantly reduce all forms of violence and related death rates everywhere. – End abuse, exploitation, trafficking, and all forms of violence against and torture of children. – Provide legal identity for all, including birth registration	– A large number of children is the victim of trafficking. These children get trafficked for sexual exploitation, forced labor, and other forms of exploitation.
SDG 17: Partnerships for the Goals	– Mobilize additional financial resources for developing countries from multiple sources – By 2020, enhance capacity-building support to developing countries, including for least developed countries and small island developing States, to increase significantly the availability of high-quality, timely, and reliable data disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts	– In the context of migrants, there is enough scope of improvement for this SDG – We need to strengthen the means of implementation and revitalize the global partnership for sustainable development

Note: The contents of this table are mostly extracted from the Migration Data Portal. Available at https://www.migrationdataportal.org/sdgs?node=0, Accessed on July 01, 2022

1.3 Social Media Platforms, Opinion Mining, and Migration

As aforementioned, AI-based approaches gained acceptance in the 1960s, but the last two decades have seen phenomenal growth momentum in this domain. Big data, primarily user-generated social media data, was a catalyst for AI-based research (Zeng et al., 2010; Barbier & Liu, 2011; O'Leary, 2013). Theoretically, the genesis of user-generated data can be traced back to the pre-internet era. An article[8] in The Washington Post rightly said that *Before Twitter and Facebook, There Was Morse Code: Remembering Social Media's True Inventor*. However, Morse code for telecommunication had limited applications, and it cannot be considered big data. The surge of user-generated big data was triggered by the Internet and the advent of Web 2.0 technologies (Chen & Zimbra, 2010; Ghani et al., 2019). Popular social media platforms, such as Orkut, Facebook, Twitter, and YouTube, gained popularity among common users in the mid-2000s (Refer to Figure 1.1). Social media data satisfies the criteria of all *three Vs of big data*, i.e., volume, variety, and velocity (O'Leary, 2013). Consequently, this voluminous unstructured user-generated data attracted the attention of information science researchers. Extant literature has considered various online contents, such as web search query data (i.e., the volume and location of the query), text data (i.e., short text messages by users on social media platforms), or image data (e.g., Instagram posts), as a crucial source of information for probing issues ranging from socio-economic events (e.g., election prediction, stock market sentiments, social movements like Arab Spring or *#BlackLivesMatter*) to disasters (e.g., earthquakes, floods to epidemic outbreaks).

Initial studies focused on web search query data. Web search query patterns of online users capture the extent of public interest or concern related to a particular topic or event, and historical patterns of these search queries allow us to do trend analysis for a specific topic or event. One of the pathbreaking works in this domain was the *Google Flu Trend (GFT)* analysis, published in *Nature* by Ginsberg et al. (2009). This seminal paper investigated Google search queries submitted by numerous users worldwide to predict influenza-like illness in the USA context. Following Ginsberg et al. (2009), prior studies in this domain have considered various count measures such as Google-search query volumes, access to Wikipedia pages, and volume of tweet feeds in the context of outbreak management, election prediction, financial forecasting and so on (Bollen et al., 2011; Khatua et al., 2019; 2020; Sharma & Sharma, 2020; Yoshida et al., 2015).

---

[8] Available at https://www.washingtonpost.com/news/retropolis/wp/2017/05/24/before-there-was-twitter-there-was-morse-code-remembering-social-medias-true-inventor/, Accessed on July 1, 2022

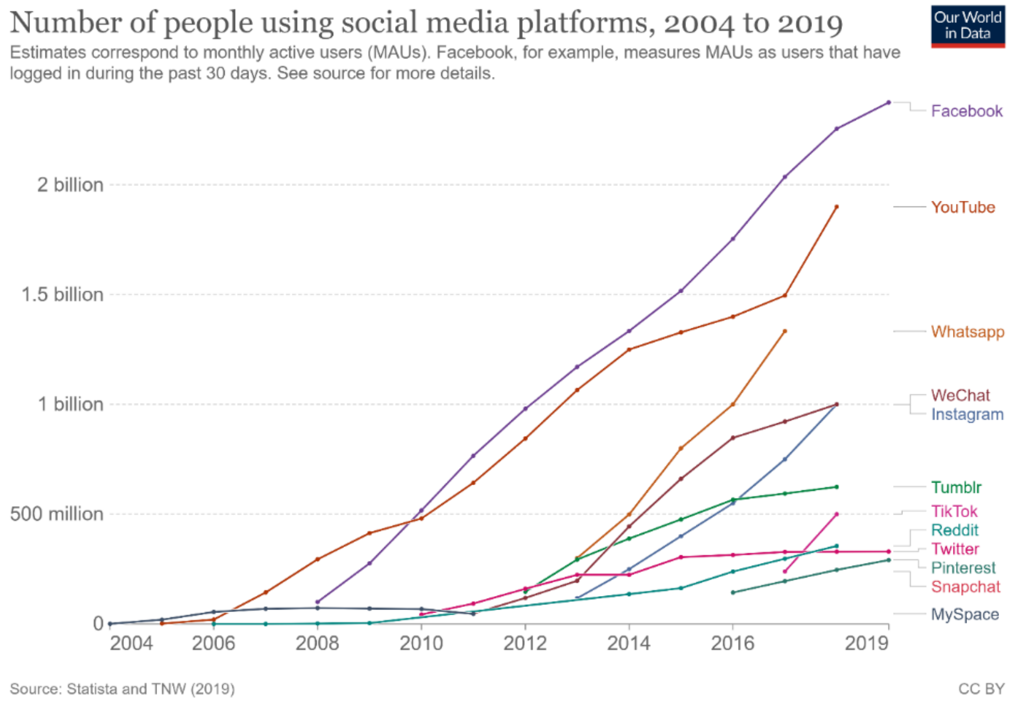


Figure 1.1 The Rise of Social Media Platforms

Source: https://ourworldindata.org/rise-of-social-media , Accessed on July 1, 2022

Subsequently, NLP researchers started probing information-rich text data, such as Twitter posts or YouTube comments, for opinion mining. As rightly pointed out by Cambria et al. (2013, p. 15), “capturing public opinion about social events, political movements, company strategies, marketing campaigns, and product preferences is garnering increasing interest from the scientific community (for the exciting open challenges), and from the business world (for the remarkable marketing fallouts and for possible financial market prediction). The resulting emerging fields are opinion mining and sentiment analysis.” Thus, AI-based opinion mining of social media data has emerged as one of the dominant streams of computational social science research (Chen & Zimbra, 2010; Zeng et al., 2010; Liu, 2012; Cambria et al., 2013; Tang et al., 2014). A review paper by Sun et al. (2017, p. 10) identified “various practical applications of opinion mining, such as product pricing, competitive intelligence, market prediction, election forecasting, nation relationship analysis, and risk detection in banking systems.” Similarly, another review paper by Ghani et al. (2019) identified various social media data sources: microblogging sites, news articles, blog posts, internet forums, reviews, and Q&A posts. Ghani et al. (2019) also noted that 46% of the prior studies considered tweet feeds as their data source. However, opinion mining or sentiment analysis of this unstructured social media data is not a trivial task, and elementary approaches might fail to capture the finer nuances (Ghani et al., 2019). For example, Cambria & Hussain (2015) pointed out that two sentences might be similar from a bag-of-words perspective, but they can bear oppositive polarities. For instance, *'iPhone is expensive, but nice'* vis-à-vis *'iPhone is nice, but expensive'* can be syntactically the same but not semantically. The former sentence displays positive sentiment (i.e., the product is nice despite its high price), but the latter sentence expresses negative sentiment (i.e., the user likes the product but thinks it is too expensive). These challenges are not specific to product reviews but unavoidable even in the context of migration. For instance, NLP literature associates abusive or swear words with hate speech. So, a simplistic bag-of-words approach will associate offensive words in a migrant-related comment with hate speech toward migrants. However, this dissertation questioned this (refer to Chapter 3) and cited anecdotal evidence to elucidate that a pro-migrant post can also use swear words against discriminatory policies. Social media users' opinions might not be explicit; rather, they can be implicit. Thus, interpreting implicit opinions is a challenging task (Cambria & Hussain, 2015; Camacho et al., 2021). To address the same, literature has evolved from a bag-of-words approach to ML applications like Naive Bayes classifier or Support Vector Machine to advanced DL models (Chen & Zimbra, 2010; Ghani et al., 2019; Sun et al., 2016). Interestingly, a significant portion of the opinion mining research has explored social media platforms as well as e-commerce and online review sites (e.g., Twitter, Amazon, Yelp, etc.) for business and e-commerce applications like product or movie reviews (Chen & Zimbra, 2010; Cambria et al., 2013; Ghani et al., 2019; Sun et al., 2016). However, relatively "little research has tried to understand opinions in the social and geopolitical context" (Chen & Zimbra, 2010, p. 75). Thus, opinion mining of migrant-related issues on social media platforms also addresses this concern. The following section reports a brief review of opinion mining studies using social media data in the migration context.

Migrant-related extant literature has explored social media platforms such as Twitter (Aswad & Menezes, 2018; Calderón et al., 2020a; 2020b; Gualda & Rebollo, 2016; Nerghes & Lee, 2018; 2019), Facebook (Capozzi et al., 2020; Hrdina, 2016), YouTube (Lee & Nerghes, 2018), and Instagram (Guidry et al., 2018). These opinion-mining studies have probed social media deliberations in the context of specific events such as the terrorist attack in Paris, the sexual assault in Cologne, the EU-turkey deal for refugees, and other unfortunate events (Pope and Griffith, 2016; Sajir, 2019; Siapera et al., 2018). For instance, Pope & Griffith (2016) compared the positive and negative emotions, like anger and anxiety, after the Paris and Cologne incidents. Similarly, Sajir (2019) analyzed the mainstream print media reaction as a response to two unfortunate events related to two Syrian kids, namely, Alan Kurdi (in 2015) and Omran Daqneesh (in 2016). This study probed whether an unfortunate event could lead to a solidarity movement or not. Literature has also explored the apprehensiveness toward migrants. For instance, Kreis (2017) also analyzed how a particular hashtag, i.e., *#refugeesnotwelcome*, was used to express negative feelings, beliefs, and ideologies toward refugees and migrants in Europe. Ozerim & Tolay (2020) analyzed Turkish hashtags against Syrian refugees to probe how negative emotion can gain momentum. They found that the hashtag *#IdontwantSyriansinmycountry* was tweeted and retweeted over 54,000 times in a single day in July 2016. Similarly, Rettberg and Gajjala (2016) tried to investigate the (negative) perceptions about male refugees coming from the Middle East and argued that apprehensions or negative perceptions emerged due to the insignificant presence of female and young refugees from the Middle East. Thus, literature has primarily analyzed social media data to investigate focused issues such as deserving vis-à-vis undeserving migrants, security-related concerns vis-à-vis humanitarian-related concerns or critical vis-à-vis positive tweets (Guidry et al., 2018; Hadgu et al., 2016; Nerghes & Lee, 2018).

The literature argues that nationalist conservative and xenophobic political parties establish a socially accepted racism discourse in the European context (Kreis, 2017). Similarly, a study by Gualda and Rebollo (2016) considered multilingual tweets, namely, English, French, German, Italian, Portuguese, and Spanish, in the European context. The authors found a diverse range of topics, from solidarity to xenophobia, but most importantly, this study reveals that Twitter data captures contextual issues or local events specific to a particular country. Hence, this dissertation argues that event-specific (Pope & Griffith, 2016; Sajir, 2019; Siapera et al., 2018) or issue-specific (Kreis, 2017; Rettberg & Gajjala, 2016; Urchs et al., 2019; Zagheni et al., 2014) studies might not be able to capture the entire gamut of migration-related latent issues. From the policymaking perspective, understanding this diverse range of societal opinions about migrants is crucial because, in the recent past, migration issues influenced the electoral discourse in countries like Italy (Capozzi et al., 2020), Spain (Alcántara-Plá & Ruiz-Sanchez, 2018; Calderón et al., 2020b), and the UK (Khatua & Khatua, 2016). Hence, this dissertation employs an AI-based approach for probing societal opinions through user-generated social media data. We have provided an exhaustive bibliometric analysis of this migration literature in Chapter 2. Next, we explore some macro-level migration indicators.

1.4 Global Migration: A Few Stylized Facts[9]

The number of international migrants has steadily increased from 153 million in 1990 to 280.6 million in 2020. However, the international migrant stock as a percentage of the total population is mostly consistent. Figure 1.2 indicates that it has slightly increased from 2.9% in 1990 to 3.6% in 2020 (Source: UN DESA, 2020; Based on the latest data available on 20 January 2021). In other words, the number of international migrants has drastically increased over the years, but the growth rate is consistently in the range of 3% to 4%. The Depth Index of Globalization Report[10] also pointed out that "the proportion of the world's population who are first-generation migrants is just about the same today as it was in 1910: 3%" (p. 15).

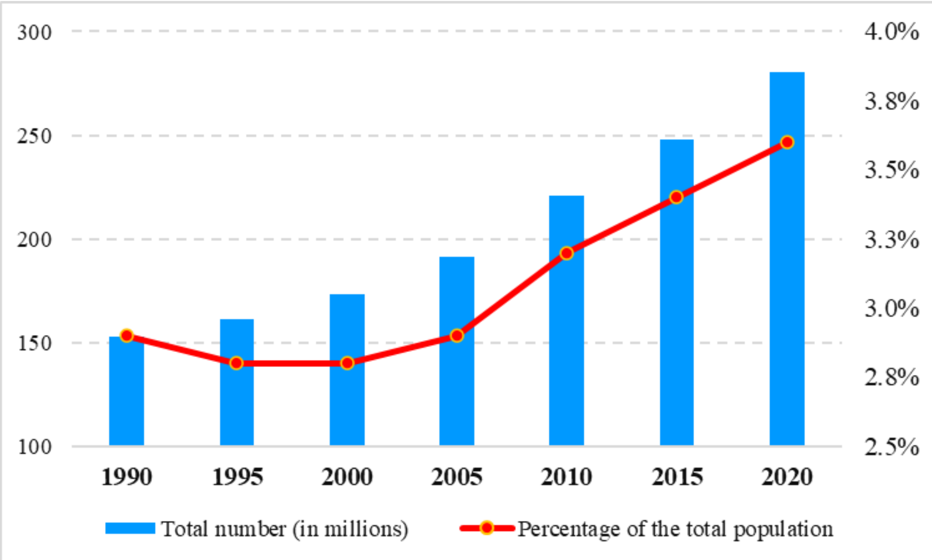


Figure 1.2 International Migrants: 1990 to 2020

However, this global trend doesn't capture the regional variations. Figure 1.3 reports country-wise international migrant stock as a percentage of the total population (Source: UN DESA, 2020; Based on the latest data available on 20 January 2021). This percentage is around 50% in some of the Middle Eastern countries (like Kuwait, Qatar, United Arab Emirates, and Oman) and smaller European countries (like Andorra, Monaco, Liechtenstein, and Luxembourg). However, this percentage is broadly in the range of 10% to 20% for most European nations like Austria (19.3%), Belgium (17.3%), Denmark (12.4%), France (13.1%), Germany (18.8%), Italy (10.6%), Netherlands (13.8%), Spain

[9] This chapter heavily draws from the Migration Data Portal (MDP). Figures/infographics in this section, and also in other chapters are mostly extracted from MDP website. Available at https://www.migrationdataportal.org/, Accessed on July 1, 2022

[10] *Depth Index of Globalization Report 2013* by Pankaj Ghemawat and Steven A. Altman, IESE Business School, Available at https://media.iese.edu/research/pdfs/ST-0310-E.pdf, Accessed on July 1, 2022

(14.6%), and Sweden (19.8%). This percentage is also similar for other countries like Canada (21.3%), the UK (13.8%), and the USA (15.3%). This percentage is significantly low for the two most populous nations, namely, China (0.1%) and India (0.4%).

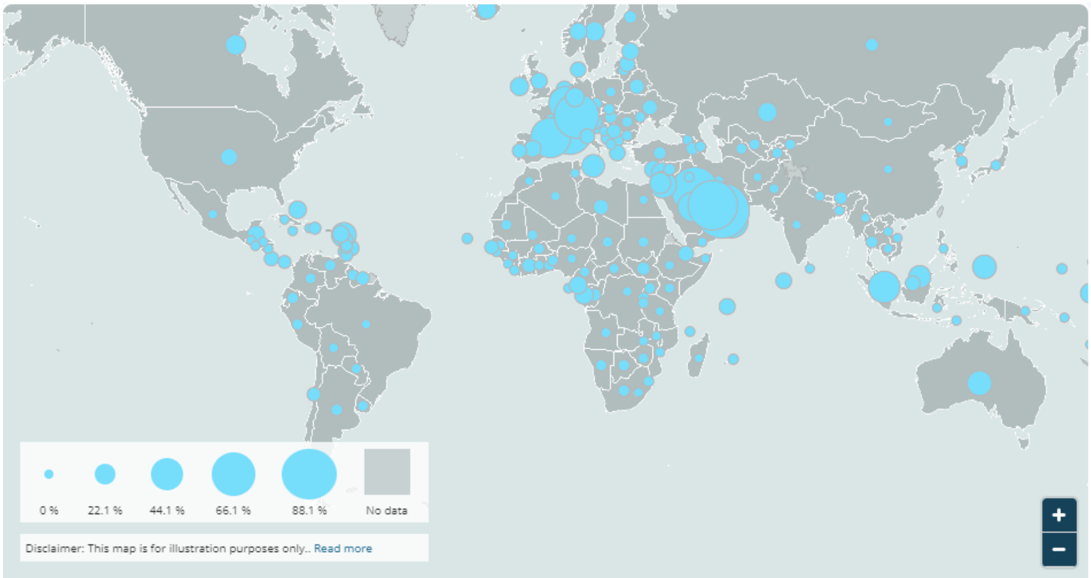


Figure 1.3 International migrant stock as a percentage of the total population[11]

Figure 1.4 reports the difference in the share of migrants in the total population between 2000 and 2020. In accordance with Figure 1.2, Figure 1.4 doesn't indicate any drastic changes for most nations. For instance, the difference in migrant stock over two decades is mostly less than 10% for countries like Austria (6.9%), Belgium (5%), Canada (3.3%), France (2.4%), Germany (7.8), Italy (6.8%), Netherlands (4%), the UK (5.8%), and the USA (2.9%).

However, the perceptions of advanced nations about migrants do not confirm these realities. For instance, the 2013 Depth Index of Globalization report pointed out that "Western Europeans across 8 countries, on average, believe immigrants comprise 25% of their country's population, while the actual figures average to only 12%. In the United States, citizens estimated that 42% of the country's population was born abroad, versus the actual ratio of only 14%" (p. 69). Similarly, a relatively recent Pew Research Centre[12] survey of 27 nations in 2018 also indicates that Europeans, especially from "Greece (82%), Hungary (72%), Italy (71%) and Germany (58%) say fewer immigrants or no immigrants at all should be allowed to move to their countries" (¶ 3). Overall, this survey also reports

[11] Available at https://www.migrationdataportal.org/international-data?i=stock_perc_&t=2020, Accessed on July 1, 2022
[12] *'Many worldwide oppose more migration – both into and out of their* countries' by Philip Connor and Jens Manuel Krogstad, Pew Research Centre, December 10, 2018, Available at https://www.pewresearch.org/fact-tank/2018/12/10/many-worldwide-oppose-more-migration-both-into-and-out-of-their-countries/, Accessed on July 1, 2022

that "a median of 45% say fewer or no immigrants should be allowed to move to their country, while 36% say they want about the same number of immigrants. Just 14% say their countries should allow more immigrants" (¶ 2).

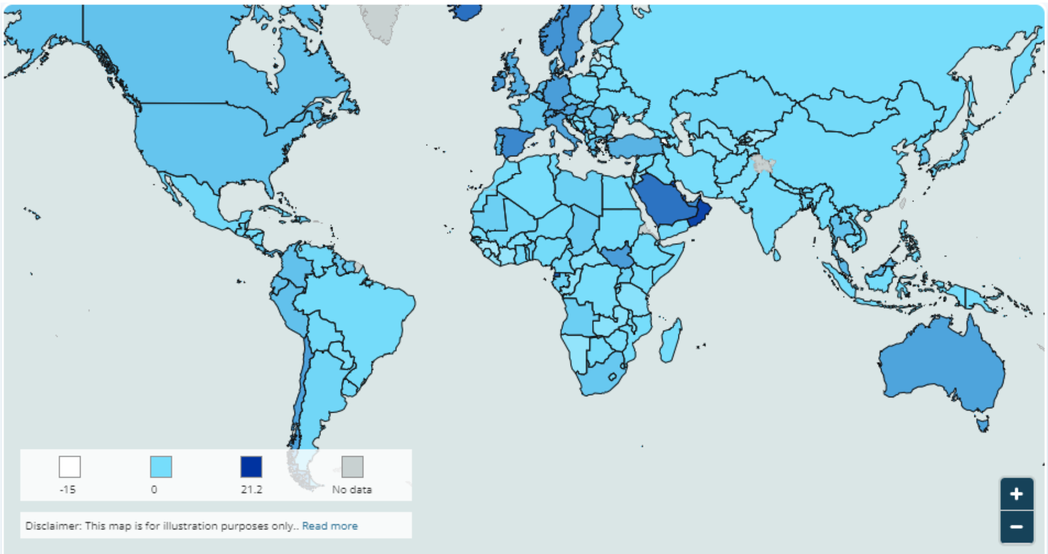


Figure 1.4 Difference in the share of migrants between 2000 and 2020[13]

Perception matters, irrespective of whether it is factually correct or not, because not only the perception impacts the social media deliberations, like hatred toward migrants, but also the political discourses, such as the 2016 and 2020 USA Presidential elections or Brexit referendum (Khatua & Khatua, 2016; Ogan et al., 2018; Waldinger, 2018). An apprehensive view toward migrants, especially by far-right political parties, is gaining momentum across countries and affecting the political mandates. Chapter 3 of this dissertation exhaustively investigates these perceptions towards migrants. Figures 1.5 and 1.6 indicate that policies toward migrants have become stringent over the years, especially in advanced nations. Figure 1.5 reports the Migration Control Policy Index in 2010, where 0 is the least restrictive, and 1 is the most restrictive (Source: IMPIC[14], 2016). Figure 1.5 indicates that advanced nations have implemented restrictive migration control policies. Similarly, Figure 1.6 also reconfirms that the number of recorded migration policy changes is significantly high in these countries (Source: DEMIG[15], 2014). Thus, it is crucial to understand – *why citizens of host nations around the world oppose migration or dislike migrants*. Broadly, this dissertation makes a humble attempt to explore this.

[13] Available at https://www.migrationdataportal.org/international-data?i=stock_perc_b22&t=2020, Accessed on July 1, 2022

[14] IMPIC: The Immigration Policies in Comparison (IMPIC) project provides a set of sophisticated quantitative indices to measure immigration policies in most OECD countries and for the time period 1980-2010, Available at http://www.impic-project.eu/, Accessed on September 22, 2023

[15] DEMIG (2015) DEMIG POLICY, version 1.3, Online Edition. Oxford: International Migration Institute, University of Oxford. www.migrationdeterminants.eu

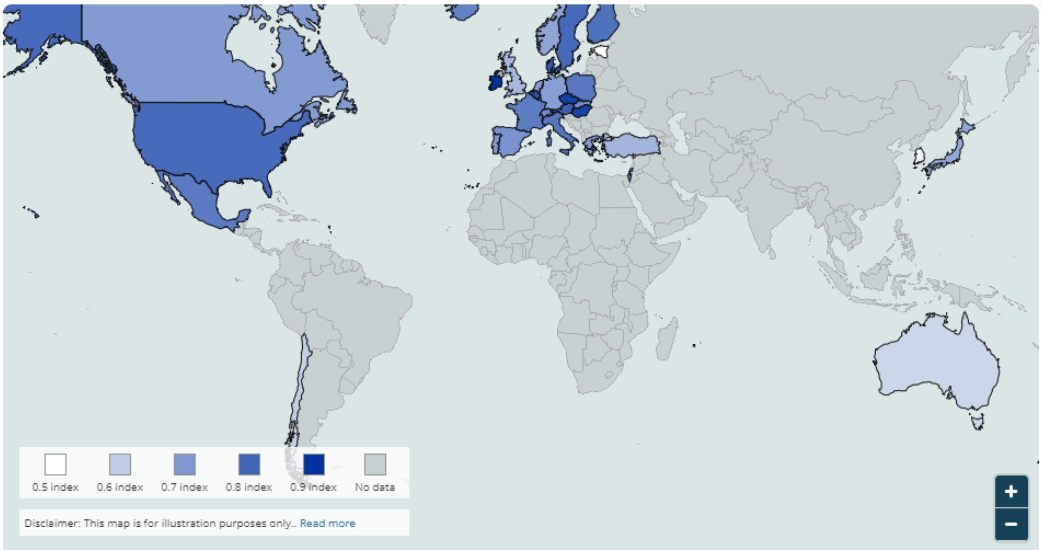


Figure 1.5 Migration Control Policy Index in 2010[16]

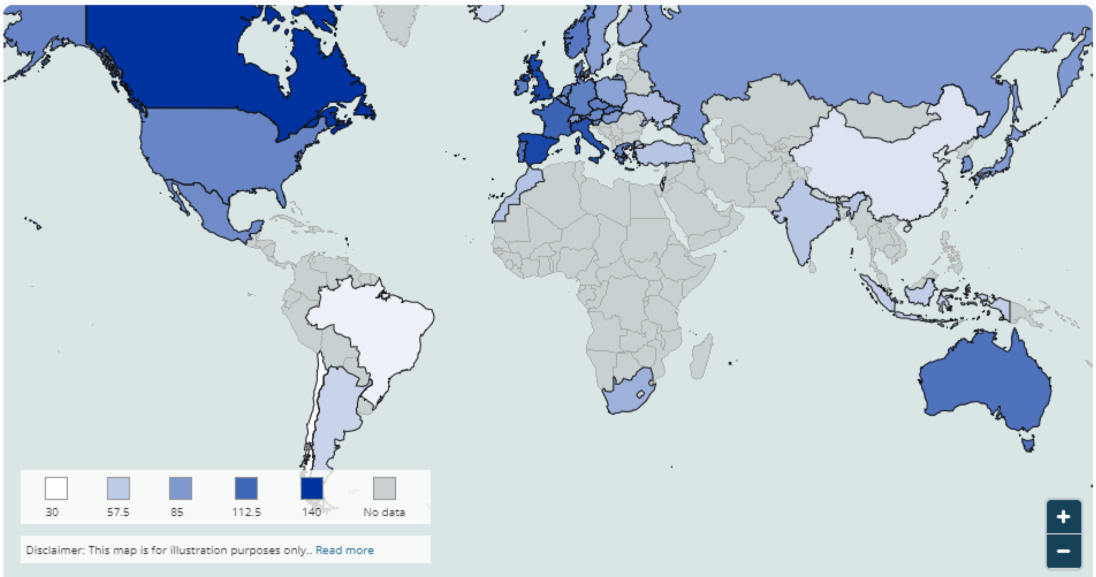


Figure 1.6 Migration policy changes between 1994 and 2014[17]

1.4.1 Migrants or Refugees?

A UNCHR viewpoint[18] article has categorically mentioned that these two terms are distinct and have different meanings. For instance, "refugees are persons fleeing armed conflict or persecution ... Their situation is often so perilous and intolerable that they cross national borders to seek safety in nearby countries ... it is too dangerous for them to return

[16] Available at https://www.migrationdataportal.org/international-data?i=impic_control&t=2010, Accessed on July 1, 2022
[17] Available at https://www.migrationdataportal.org/international-data?i=demig_no&t=2014, Accessed on July 1, 2022
[18] Available at https://www.unhcr.org/news/latest/2016/7/55df0e556/unhcr-viewpoint-refugee-migrant-right.html, Accessed on July 1, 2022

home, and they need sanctuary elsewhere. These are people for whom denial of asylum has potentially deadly consequences" (¶ 3). Contrarily, migrants come to foreign countries not due to "threat of persecution or death, but mainly to improve their lives by finding work, or in some cases for education, family reunion, or other reasons … If they choose to return home, they will continue to receive the protection of their government" (¶ 6). Generally, we note that social media users ignore these nuances and use these terms synonymously. Interestingly, a handful of studies pointed out that social media users are more sympathetic to *deserving* refugees in comparison to *undeserving* migrants (Lee & Nerghes, 2018; Nerghes & Lee, 2018). Notably, migrants and refugees are not the same for policymakers or host governments. Host governments can decide about migrants based on their immigration laws. However, countries must abide by the international laws and processes for refugees "seeking asylum on their territories or at their borders. UNHCR helps countries deal with their asylum and refugee protection responsibilities" (¶ 7). In other words, refugees are entitled to specific legal protections according to international laws. Subsequently, the UNCHR viewpoint article also clarified the reality:

> *"… back to Europe and the large numbers of people arriving in recent years by boats in Greece, Italy and elsewhere. Which are they? Refugees or migrants?' In fact, they happen to be both. The majority of people arriving in Italy and Greece especially have been from countries mired in war or which otherwise are considered to be 'refugee-producing' and for whom international protection is needed. However, a smaller proportion is from elsewhere, and for many of these individuals, the term 'migrant' would be correct"* (¶ 9 & ¶ 10)

The above excerpt explains why social media users might fail to differentiate between these terms. As noted earlier, Nerghes & Lee (2018) analyzed 0.37 million tweets to explore public sentiment in response to the unfortunate drowning of Alan Kurdi. They identified two distinct streams of discussion: *deserving* refugees vis-à-vis *undeserving* migrants and found that tweets with refugee-related hashtags had displayed a positive tone. Similarly, another study by Lee & Nerghes (2018) argued that labels such as *refugee* and *migrant* could evoke different connotations. This study employed topic modeling and sentiment analysis on the comments of two popular YouTube videos titled *The European Refugee Crisis and Syria Explained* and *What Pisses Me Off About the European Migrant Crisis* and argued that labeling distinguishes "between those who are deserving from those who are considered less-deserving and potentially a threat to be rejected" (p. 12).

However, for collecting user-generated data, this dissertation has considered both *migrant* and *refugee* as search keywords in addition to other context-specific keywords.

1.5 Social Media and Migration Discourse: A Pilot Study

1.5.1 Research Question for the Pilot Study

Before finalizing this dissertation's research question(s), we have performed a pilot study to get a nuanced understanding of migrant-related social media deliberations. Our brief review in Section 1.3 suggested that prior studies mainly probed a specific migration-related event or analyzed a particular issue. There are hardly any studies that attempt to identify and classify a diverse range of opinions on social media platforms. Hence, this pilot study has attempted to identify multiple strands of migrant-related discourse. Broadly, the research question for this pilot study is: *What is/are the salient theme(s) of migrant-related discourse on social media platforms?*

1.5.2 Research Context: Europe

To decide the context of our pilot study, we explored migration flows and observed some interesting patterns. For instance, in the preceding section, we have probed international migrants as a percentage of the total population and noted that this percentage is 10% to 20% for most developed nations (refer to Figure 1.3). We also need to consider another crucial indicator, i.e., the actual number of migrants. Figure 1.7 reports the total number of international migrants in 2020 (Source: UN DESA, 2020; Based on the latest data available on 20 January 2021). It is worth noting that more than half of all international migrants (141 million) live in Europe and North America, and the top destination remains the U.S (World Migration Report, 2020). For instance, the USA has the highest number of international migrants, i.e., 50.6 million migrants (Source: UN DESA, 2020).

Thus, the puzzling question is – if advanced nations, such as European countries or the USA, are adopting stringent migration control policies (Figure 1.5) or implementing a series of policy changes (Figure 1.6), how are half of all international migrants in these countries? To probe this paradox, we need to consider the trend instead of static data points like the total number of international migrants. Figure 1.8 reports the inflows of foreign population in 2019 (Source: OAS, 2017; OECD, 2021; Based on the latest data available on 14 December 2021). We can find a distinct pattern between Figure 1.5 and Figure 1.8, i.e., stringent migration control policies are associated with higher inflows. Intuitively, Europe and the USA are implementing restrictive migration policies because 50% of international migrants are already in these countries. Hence, they are trying to curb the further inflow.

We also note that the inflows of foreign population in European countries are significantly higher compared to other parts of the world. For instance, according to Figure 1.8, the inflow of foreign population is 1.3 million in Germany compared to 1 million in the USA. Similarly, the inflow of the foreign population is 666 thousand in Spain, which is significantly higher compared to 341.2 thousand in Canada. Figure 1.9 reports the total number of refugees by country of origin. (Source: UNHCR Refugee Population Statistics Database, 2022; Based on the latest data available on 25 June 2022).

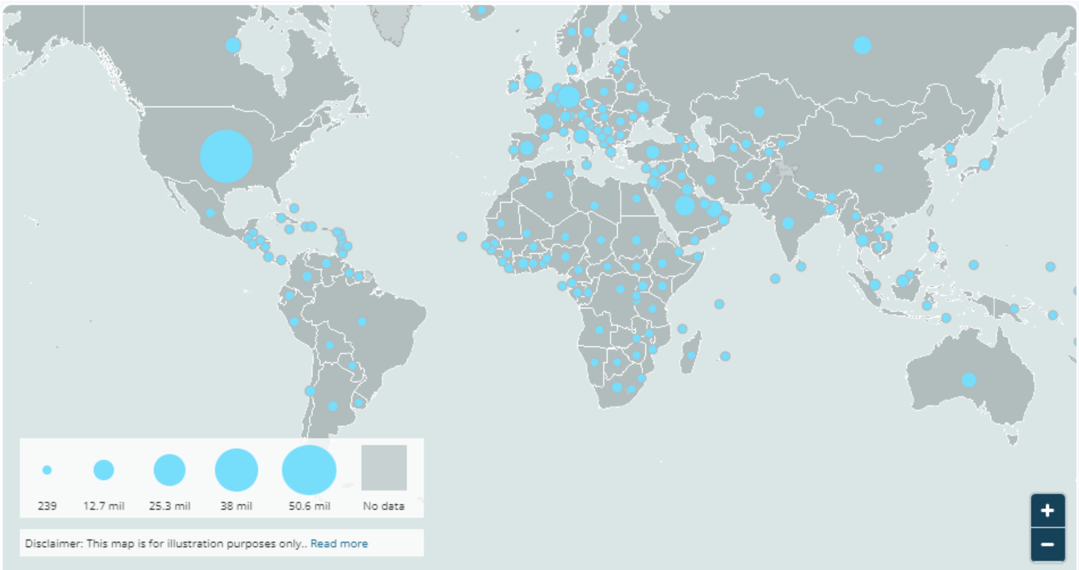


Figure 1.7 Total number of international migrants at mid-year 2020[19]

Figure 1.8, in conjunction with Figure 1.9, revealed an interesting pattern for Europe's migrant issues. For instance, migrants in the USA are primarily from their neighboring countries like Mexico. On the other hand, European migrants (or refugees) are mostly coming from Middle Eastern countries (Nerghes & Lee, 2019). While the total number of refugees from Mexico is only 16.4 thousand, the number of refugees from the Syrian Arab Republic is 6.8 million. Refugees from these Middle Eastern countries take the risky and uncertain sea route to enter the European continent through Greece, and then they move from one European nation to another (Gualda & Rebollo, 2016). Also, these migrants differ sharply from host nations regarding their language, ethnicity, and culture. Hence, the migration issue in the European context is unique in nature. Therefore, in this pilot study, we focus on migrant-related deliberations in the European context.

[19] Available at https://www.migrationdataportal.org/international-data?i=stock_abs_&t=2020, Accessed on July 1, 2022.

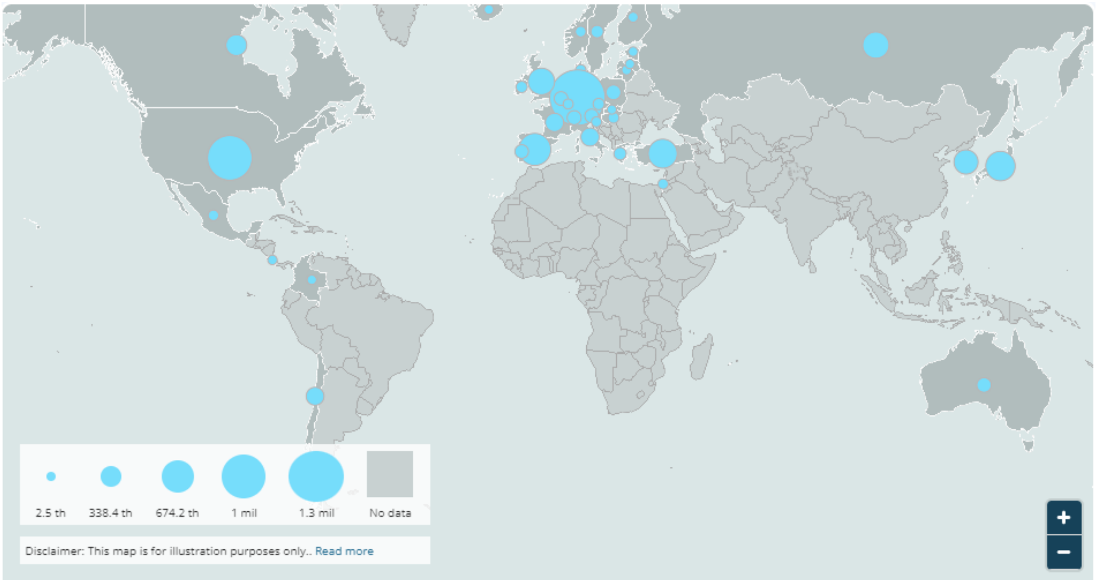


Figure 1.8 Inflows of foreign population in 2019[20]

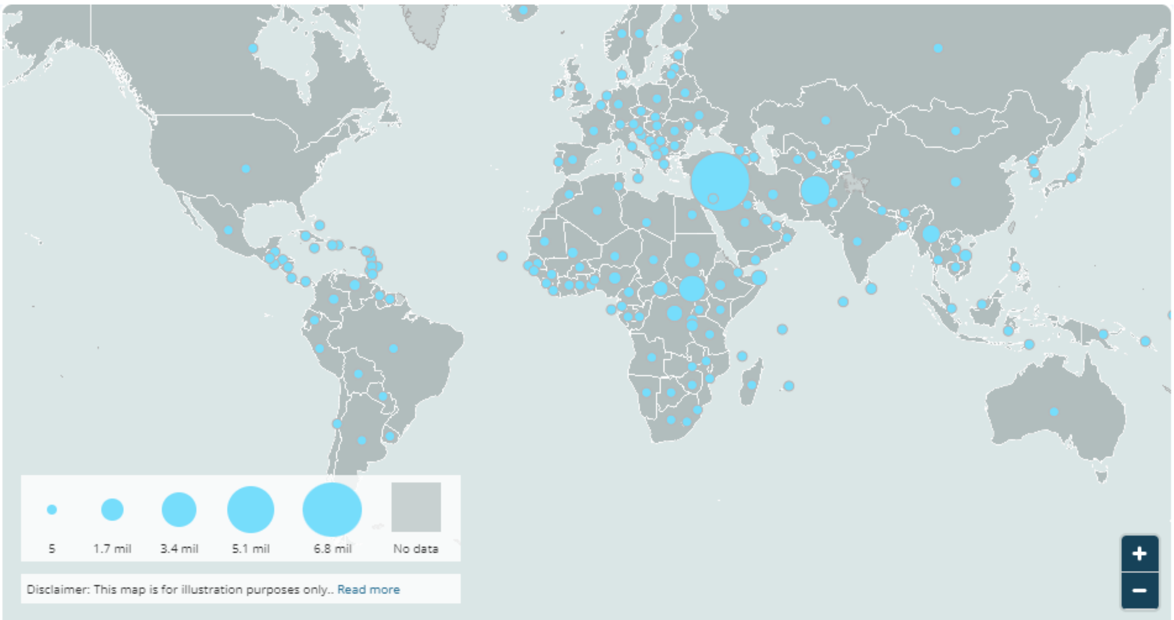


Figure 1.9 Total number of refugees by country of origin, end of 2021[21]

1.5.3 Data Collection

Extant literature argued that Twitter could be 'a proxy for reality' in the context of refugees and migrants (Aswad & Menezes, 2018). Hadgu et al. (2016) observed that the "actual situation would influence the intensity and polarity of discussions" on social media platforms, and subsequently, Twitter also captures "how the discussion on refugees

[20] Available at https://www.migrationdataportal.org/international-data?i=inflow_total&t=2019, Accessed on July 1, 2022
[21] Available at https://www.migrationdataportal.org/international-data?i=refug_origin&t=2021, Accessed on July 1, 2022

changed over time” (p. 1). Nerghes & Lee (2019) also pointed out that “unlike the mainstream media, Twitter offered an alternative and multi-faceted narrative, not bound by geo-politics, raising awareness and calling for solidarity and empathy towards those affected” (p. 275). Thus, we have employed the Twitter search API for collecting Twitter data from May 2020 to September 2020. For our initial crawling, we have considered a set of migrant-related keywords, such as *migrant*, *refugee*, *UNCHR*, *asylum*, *deport*, etc. API-based crawling has yielded around 1.8 million tweets. We have discarded duplicate tweet-ids as well as removed tweets with similar text. Our initial analysis suggests that concise tweets might not be insightful from an opinion-mining perspective. Hence, we have considered long tweets (i.e., more than 90 characters). Next, we have considered tweets that have mentioned either of these four keywords: *‘migrants’*, *‘refugee’*, *‘immigrant’*, and ‘*immigration’*. This step has helped us to discard junk tweets and reduced our corpus to 0.3 million tweets.

Table 1.2 Sample Tweets for Pilot Study

Salient Themes	Representative tweets *(678 tweets)*
Safety Concerns *(142 tweets)*	An Afghan migrant and his two sons have been arrested in the brutal stabbing murder of a man aboard a bus in Kiruna, Sweden, according to reports
Economic Conditions *(117 tweets)*	This is shocking - poor conditions of migrant #strawberry pickers in Spain - it is critical those who harvest our food are paid a fair living wage @FFC_Commission
Employment Opportunities *(89 tweets)*	For all #migrant #entrepreneurs & migrant-led organisations that support migrant and #refugee entrepreneurs in Europe to set up their business: WE NEED YOU. The #EMENproject is building a list of such programmes Complete this FORM by …
Healthcare Support *(154 tweets)*	Sara is an Iranian kid who is been living in Turkey as a refugee with her family since 2015. She is suffering from a rare autoimmune disease called Evans Lupus syndrome. Sara can’t get medical care she needs in Turkey, please help her and her family @ICHRI @SaveSaraLife
Inequality & Discrimination *(176 tweets)*	The media generate high levels of anxiety about immigration, resulting in negative migrant stereotypes. The public debate in EU has been influenced by populist politicians and biased media coverage with a background of subconscious postcolonial legacy

We find a significant portion of these 0.3 million tweets are related to the USA context and revolve around Trump’s migration-related policy. Hence, these tweets are not relevant to the European context. We have checked the linguistic content of tweets and considered tweets that have categorically mentioned keywords such as ‘European Union’, ‘Europe’, ‘Germany’, ‘Greece’, ‘France’, and other European countries. We also find that the volume of English tweets is significantly higher than other major European languages, such as French, Spanish, and German. So, we considered English tweets. This process has resulted in a corpus of around 24,000 unique European migration-related English tweets.

1.5.4 Identification of Salient Themes on Twitter

Our brief review in the previous section reveals that prior studies primarily provided an in-depth understanding of focused issues, such as deserving vis-à-vis undeserving migrants or security-related concerns vis-à-vis humanitarian-related concerns or critical vis-à-vis positive tweets (Guidry et al. 2018; Hadgu et al., 2016; Nerghes & Lee, 2018). However, migrant-related concerns are multi-faceted. Thus, it is essential to probe whether online content can capture these diverse issues, ranging from health concerns to job prospects of migrants, OR whether immigrants face discrimination and, subsequently, engage in violent activities. As noted, the objective of this pilot study is to explore a diverse range of migrant-related concerns. Hence, we referred to the migration literature to understand various concerns and tried to map these concerns with our tweet corpus.

We have classified our tweet corpus into five categories: *safety concerns*, *economic conditions of migrants*, *employment opportunities*, *healthcare support*, and *inequality & discrimination*. For example, we have considered bigotry, societal inequity, hatred, and prejudice toward migrants as *discrimination-related concerns*. Similarly, we have considered violence by migrants and subsequent police action, such as arrest, as *safety concerns*. Table 1.2 reports a few sample tweets for these themes. We have prepared a balanced sample (i.e., the numbers of annotated tweets across classes are broadly in the same range) and considered 678 annotated tweets for our final analysis.

1.5.5 Classification of Salient Themes

Table 1.3 Accuracies of CNN and LSTM models

CNN	LSTM	BS	DR
0.8319	0.7058	16	0.3
0.7341	0.6985	16	0.5
0.8088	0.5955	32	0.3
0.6029	0.7058	32	0.5

Prior studies found that DL models are superior to the traditional bag-of-words approach for this classification task (Conneau et al., 2017; Kalchbrenner et al., 2014). Hence, we have employed CNN and LSTM as our baseline models. We have also considered two transformer-based models: BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019). We have considered the HuggingFace python library (Wolf et al., 2019), and it includes pre-trained models and allows fine-tuning of hyperparameters. For our classification task, we have employed *'BertForSequenceClassification'*. Subsequently, we have considered the BERT-Base-Uncased model comprised of 12-layers and 12-heads with a total of 110M parameters. We have considered *max_seq_length* of 256 for our analysis.

Table 1.4 Accuracies of BERT and RoBERTa models

BERT	RoBERTa	BS	LR
0.9144	**0.9265**	16	2e-5
0.8877	0.9191	16	3e-5
0.9037	0.9191	16	5e-5
0.9037	**0.9265**	32	2e-5
0.9091	0.8676	32	3e-5
0.9091	0.8750	32	5e-5

Table 1.3 reports the classification accuracies of CNN and LSTM models for multiple BS (16 and 32) and DRs (0.3 and 0.5). The hidden layer for all these models was 256. We have considered *SoftMax activation* in our final classification layer to predict the final class. We use 'rmsprop' as our optimizer. Classification accuracies for CNN models are mostly higher than for LSTM models, and the highest accuracy is 83.2%.

Similarly, Table 1.4 reports the findings of the BERT and RoBERTa models. We have considered the following combinations: BS (16, 32) and LRs (2e-5, 3e-5, 5e-5). We have used a maximum sequence length of 256 and employed 4 epochs for our BERT-Base-Uncased models. Our open-source implementation, pre-trained weights, and full hyperparameter values are in accordance with the HuggingFace transformer library[22] (Wolf et al., 2019). Overall, the accuracies of transformer-based models are better than our base models, and we find that accuracies are mostly more than 90%. We also observe that RoBERTa models have performed marginally better than BERT models.

1.5.6 Takeaways from the Pilot Study

Following the extant literature, our pilot study revealed two diametrically opposite societal opinions:

- Themes, such as the economic conditions of migrants or employment opportunities, capture sympathy *toward migrants*.
- Contrarily, a theme like safety concerns captures the *apprehensive view of society*.

[22] Available at https://huggingface.co/docs/transformers/model_doc/bert, Accessed on September 22, 2023

The findings of our pilot study are in accordance with Guidry et al. (2018) which explored the visual and textual posts on Instagram and Pinterest platforms to probe the 2011 Syrian refugee crisis. They also found two types of behavioral concerns: *security-related and humanitarian-related*.

Broadly, this dissertation, specifically Chapters 3 and 5, argues that these antipathies or apprehensions toward migrants can potentially trigger hate speech toward migrants. This pilot study also elucidated some of the inherent limitations of social media data. First, social media data is voluminous but unstructured. Hence, AI-based studies need a nuanced understanding of the context to classify an unstructured social media corpus into specific themes or concerns. It is crucial to appreciate that migration, as a social concern, has been probed by multiple scholars from multiple domains over the last many decades. Hence, a nuanced understanding of social media data may require an interdisciplinary approach, and Chapter 2 elaborates on the need for this approach.

Second, social media data allows opinion mining, but only a minuscule portion of society is vocal on platforms like Twitter. Addressing this second concern is beyond the scope of this dissertation. However, we have made a humble attempt to analyze the endorsement pattern of social media users in Chapter 5. Endorsement analysis may capture the opinion of a relatively passive or silent user, who might not comment, but implicitly express her opinion by endorsing a particular comment. As aforementioned, this dissertation humbly attempts to understand societal opinions toward migration and migrants through user-generated social media data. Hopefully, this dissertation has taken a baby step in the direction of employing *AI for Social Good* (AI4SG). The following section briefly discusses the subsequent chapters of this dissertation.

1.6 An Overview of the Dissertation

1.6.1 Ethical Declaration

At the outset, we want to clarify that this dissertation cites some tweets or YouTube comments that are offensive and vulgar toward migrants and refugees. However, we do not endorse the views expressed in these tweets or comments but quote them only for academic purposes. These offensive quotes do not reflect our opinions, and we strongly condemn offensive language on social media. Existing literature has labeled these offensive tweets as hate speech (Davidson et al., 2017; Waseem & Hovy, 2016). These hate tweets and the propagation of hatred toward migrants and other vulnerable classes on platforms like Twitter and YouTube elucidate the darker side of social media –a concern for society. Effective implementation of some of our policy-level recommendations based on the findings of Chapter 3 can potentially reduce the animosity and abusive attitude toward migrants.

1.6.2 Bibliometric Review of Migration research using Social Media Data

Chapter 2 reviews the migration literature based on social media data. A bibliometric analysis is conducted on around 400 articles in the Scopus databases. Overall, our review indicates an increased interest in migration-related research in recent times from the computer science and social science domains. Chapter 2 also identifies prolific researchers in this domain and how they relate to other researchers in the field. We find that research is scattered at best, and we see multiple small groups of researchers. In addition, we have also performed a content analysis of the keywords of these articles and tried to identify various research clusters and how they have evolved over the years.

1.6.3 Beyond Simplistic Opinion Mining

The pilot study (in this chapter) helped us to identify various latent themes of migrant-related social media discourse. These latent themes indicate that one portion of the social media deliberations is sympathetic toward migrants and concerned about their job opportunities or discrimination towards them. Contrarily, the security-related concerns of host nations reflect the apprehension toward migrants. To probe these issues further, **Chapter 3** draws insights from the social psychology literature to investigate two facets of Twitter deliberations about migrants, i.e., perceptions about migrants and behaviors toward migrants. Our theoretical anchoring helped in identifying two prevailing perceptions (i.e., *sympathy* and *antipathy*) and two dominant behaviors (i.e., *solidarity* and *animosity*) of social media users toward migrants. Chapter 3 indicates that AI-based approaches, such as unsupervised and supervised NLP models, can efficiently identify these perceptions and behaviors. Chapter 3 also argues that tweets conveying antipathy or animosity can be broadly considered hate speech toward migrants, but they are not the same. Broadly, Chapter 3 has fine-tuned the binary hate speech detection task by highlighting the granular differences between the perceptual and behavioral aspects of hate speech - especially in the context of migrants.

1.6.4 The Journey from Home to Host Nations

Extant literature has primarily examined social media deliberations to probe societal opinions, but not the journey of migrants or refugees from their home country to the host country. Thus, **Chapter 4** attempts to identify the various stages of their journey, i.e., challenges and constraints from displacement to emplacement in the host nation. This chapter draws insights from Gennep's seminal anthropological work, i.e., *Les Rites de*

Passage, to identify four phases of the migrant journey: *Arrival of Refugees*, *Temporal stay at Asylums*, *Rehabilitation*, and *Integration of Refugees* into the host nation. We have employed a fusion architecture for classifying our multimodal tweets on the methodological front. As anticipated, we find that a combination of transformer-based language models and state-of-the-art image recognition models can outperform unimodal models. Subsequently, Chapter 4 has also considered multimodal tweets during the 2022 Ukrainian crisis to test the practical implication of this proposed framework in real time. This subsequent analysis strongly indicates the generalizability of our proposed framework.

1.6.5 Migrants: Cultural Us or Cultural Others?

Existing literature has employed chiefly a supervised approach to manually annotated social media data. This stream of studies offers a nuanced understanding of various issues, but a limited sample of annotated text inputs does not provide insights into the prevailing societal-level sentiments. Extant literature has noted that migrant-related deliberations on social media platforms are primarily associated with negative emotions. However, the literature has rarely probed – whether these negative sentiments get endorsed by other users. If yes, does it depend on who the migrants are – especially if they are cultural others? The 2022 Ukrainian refugee crisis offers a natural research setting to probe these intricate issues. Hence, **Chapter 5** has analyzed migrant-related social media discourse before the 2022 crisis and during this Ukrainian refugee crisis. Specifically, this chapter investigates the relationship between user endorsement and sentiments of comments on the YouTube Platform. On the methodology front, we have used logistic regressions. Chapter 5 suggests that users endorse comments with positive sentiments and reveal a negative propensity to endorse comments that use swear words. However, the analysis of the Ukrainian dataset reveals a negative propensity to endorse comments with negative sentiments, but the earlier dataset indicates a positive propensity. Thus, endorsing comments with negative sentiments may depend on who the migrants are!

1.6.6 Struggles in the Host Nations

Chapter 6 explores the challenges or barriers to the social integration of migrants. In other words, we probe the struggle of migrants to settle down in the host economy. We note that the extant literature has primarily analyzed the perceptions and behavior of host nations – not the (unheard) voices of migrants. Rarely prior studies, to the best of our knowledge, have investigated social media data to identify the voices of migrants and analyze their concerns. Our analysis indicates that an insignificant portion of tweets capture the voices of migrants. Thus, in addition to non-migrant voices, we have classified the voices of migrants into three themes: their generic views, initial struggles, and subsequent settlement in the host country. Overall, the first part of Chapter 6 indicates that AI-based approaches can efficiently identify the (unheard) voices of migrants.

When we probe the initial struggles of migrants, we realize that one of the major struggles to settle down is getting employment in the host country. Similarly, our pilot study (in Chapter 1) also indicates that one of the latent themes of migrant-related discourse is related to their job opportunities. Thus, we probe the potential of social media platforms for recommending the right job for skilled migrants. Extant literature suggests that an efficient job recommendation framework needs to recommend an appropriate job seeker to a recruiter and vice versa. Prior studies have mostly considered datasets from commercial job portals such as LinkedIn or CareerBuilder. However, these datasets are proprietary and not publicly available. Moreover, these portals charge their clients for offering customized services. Hence, job recommendations using paid portals for financially not-so-stable migrants might not be a feasible option. Therefore, we explore whether publicly available social media data can be a viable alternative to commercial job portals. Thus, the second part of Chapter 6 has developed a generic job recommendation framework that can be easily adapted for migrants.

1.6.7 The Way Ahead

AI-based approaches should not be implemented as a black box. Interpretability of sophisticated models, such as transformer-based models, is crucial to knowing this approach's limitations, such as inherent biases. Understanding biases has gained momentum in the AI domain, but we rarely came across any studies in the context of migration. So, **Chapter 7**, i.e., the concluding chapter of this dissertation, reports some preliminary analysis in this regard. We have employed the *BertViz* library to explore the interpretability of transformer-based models - especially probing the role of individual layers and attention heads to gauge these models' sensitivity. The scope of our dissertation does not allow us to examine it in a full-fledged manner. The objective of this Chapter is to identify the way ahead for future studies.

1.7 List of Dissertation-related Publications[23]

1. **Chapter 1:** Khatua, A., & Nejdl, W. (2021a). Analyzing European Migrant-related Twitter Deliberations. In *Companion Proceedings of the Web Conference 2021* (pp. 166-170).

2. **Chapter 3:** Khatua, A., & Nejdl, W. (2022c). Unraveling Social Perceptions & Behaviors towards Migrants on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 16, pp. 512-523).

3. **Chapter 4:** Khatua, A., & Nejdl, W. (2022b). *Rites de Passage*: Elucidating Displacement to Emplacement of Refugees on Twitter. In *Proceedings of the 33rd ACM Conference on Hypertext and Social Media* (pp. 214-219).

4. **Chapter 5:** Khatua, A., & Nejdl, W. (2022a). Endorsement Analysis of Migrant-related Deliberations on YouTube: Prior to and During 2022 Ukrainian crisis. In *Open Challenges in Online Social Networks* (pp. 31-38). Co-located with the 33rd ACM Conference on Hypertext and Social Media

5. **Chapter 6:** Khatua, A., & Nejdl, W. (2021b). Struggle to Settle down! Examining the Voices of Migrants and Refugees on Twitter Platform. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 95-98).

6. **Chapter 6:** Khatua, A., & Nejdl, W. (2020). Matching recruiters and jobseekers on Twitter. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 266-269). IEEE.

[23] List of other publications is available at Google Scholar or DBLP profile.

Chapter 2: Social Media and Migration: A Bibliometric Review

Social media can be considered the mirror of society, and pressing global and local issues get discussed and debated freely on social media platforms. On the one hand, social media has evolved rapidly over the last few years and is facilitating information dissemination in real-time, but on the other hand, deliberations on social media platforms are influencing our thinking. We gather information, both factually correct and factually incorrect and biased, and based on that, we form our opinions. Accordingly, some netizens may form an opinion about migration based on social media deliberations, and this is one of the pressing issues across countries for various geopolitical reasons. Thus, there has been an increased interest in probing user-generated data to understand prevailing migrant-related issues. Despite this trend, there is still a lack of understanding of the diverse migrant-related literature. For instance, computer science and social science researchers are probing migration-related social media data – commonly known as computational social science research. We note that computer science researchers mainly use statistical or ML techniques on large-scale social media data. Contrarily, social science literature primarily employs qualitative approaches and does an in-depth study of small-scale data. Mostly, there is no amalgamation between these two streams of research. Interestingly, some of the underlying themes from these two streams of research are more closely related than we realize. For instance, social science research regarding xenophobia is closely associated with AI-based research on hate speech toward migrants. Hence, the purpose of this chapter is to use bibliometric analysis tools to evaluate the discourse of migration-related research using social media data.

2.1 Methodology

To understand migrant-related research using social media data, we conducted a bibliometric study of conference papers and journal articles published from 2009 to 2022. A bibliometric analysis employs quantitative statistical techniques to analyze the overall trend of a discipline, productive authors and their collaborations, leading sources of publications, prominent research themes, and their evolution. For our analysis, we considered the Scopus database - which is regarded as one of the most significant abstract indexing databases. The logical steps of our search query were as follows:

- Step 1: Presence of the words (*Twitter* OR *Facebook* OR *Instagram* OR *YouTube* OR *WhatsApp*) AND (*Refugee* OR *Migrants* OR *Immigrants*) in (the *TITLE* of the Article OR the *ABSTRACT* of the Article OR the *KEYWORDS* of the Article).

– Step 2: Next, we limited our search to Subject Area: *Computer Science* and *Social Science*
– Step 3: Also, we limited our search to document types: *Article* and *Conference papers* published in the *English* language
– The search was performed on July 31, 2022, resulting in 498 articles.

We have not considered a specific time period, but we note that the first migrant-related paper using social media data was in 2009. Thus, effectively, our study period became 2009 to 2022 (refer to Table 2.1). Additionally, the search query logic and selection of keywords for our search evolved through an iterative process. For instance, our initial analysis reveals that some leading information science researchers (e.g., Emilio Zagheni and Ingmar Weber) published their migrant-related research in interdisciplinary venues. This cross-fertilization of knowledge is essential for an issue like migration. Hence, if we only consider *Computer Science* as a subject, these interdisciplinary papers will be dropped from our review. Consequently, we had to consider both *Computer Science* and *Social Science*. Similarly, we have not used the word 'social media' in our search query because the Scopus user interface treats 'social' and 'media' as two keywords. Therefore, the query extracted many migration studies that did not consider social media data, but the word 'social' was present in the abstract, title, or keywords. Hence, we replaced 'social media' with an exhaustive list of social media platforms. Similarly, we have not used the word 'migration' because the migration keyword will extract not only a few articles on 'bird migration' but also a voluminous study on 'software or platform migration'. However, this was not the concern for the keyword 'migrant'. Despite all this finetuning, we note that our query still extracted articles that are not relevant for our review. For example, a recent stream of research is about 'digital migrants' or 'digital immigrants'; most of these studies consider the behavior of digital migrants on social media platforms. So, our search query identified many papers on these themes. Thus, we have performed a content analysis of titles, abstracts, and keywords (and referred to the full-length article for a few instances) of all 498 articles to test the suitability of an article for our bibliometric review. In addition to digital migrant literature, we have identified many papers where the migrant is not the core issue. For example, studies on political discourse in the USA context referred to the proposed immigration policy of Trump. Here, migration was not the core issue. So, we dropped these articles. This process has reduced the size of our corpus for bibliometric analysis to 395 articles.

2.2 Descriptive Analysis

Table 2.1 A Snapshot of Articles

Descriptive Features of Articles	#
Timespan	2009 to 2022
Documents	395
Annual Growth Rate %	33.79
Average citations per doc	8.02
Single-authored articles	135
Co-Authors per article	2.31
International co-authorships %	14.94

Table 2.1 reports various descriptive features of our final sample of 395 articles –255 were published in conference proceedings, and the rest were published as journal articles. Additionally, we note that 135 articles are single-authored, but on average, the co-authors per article is 2.31. Interestingly, only 15% of articles are international collaborations between researchers, and 85% of articles are by co-authors from the same country. The average citation per article is around 8, but the citation distribution is highly skewed.

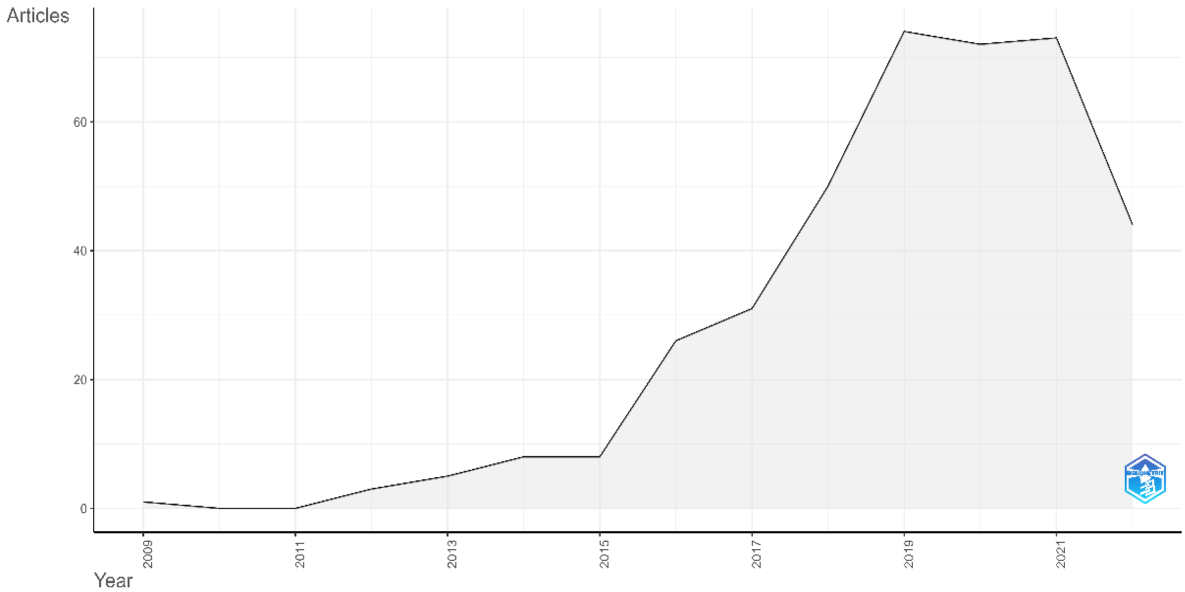


Figure 2.1 Annual Publication in this Domain

Figure 2.1 reports the annual publication trend of this domain. The yearly publication was less than ten until 2015, but it rapidly gained momentum in the next three years. It might be worth noting that the unfortunate drowning of the Syrian kid, Alan Kurdi, happened on September 2, 2015. The image became viral, and netizens primarily expressed their empathy and concern toward the refugee crisis. Probably, social media-based research was influenced by these societal dynamics. For the last three years (except the year 2022), this domain has become vibrant and produced more than 70 articles per year. The first 7 months of 2022 also produced 40-odd articles. Thus, the annual growth rate was around 33.79% (refer to Table 2.1).

2.3 Leading Conferences and Journals

Figure 2.2 reports the leading venues or sources that publish most migration-related studies using social media data, and Figure 2.3 reports the most impactful sources. Both these Figures indicate that Task 5 of the 13th International Workshop on Semantic Evaluation 2019 (commonly referred to as SemEval-2019), collocated with the Annual Conference of the *North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL-HLT 2019), was the most momentous event. SemEval 2019 Task 5 was *Multilingual detection of hate speech against immigrants and women in Twitter*. This task had two subtasks[24] as follows:

- *Task A - Hate Speech Detection against Immigrants and Women:* a two-class (or binary) classification task to predict whether a tweet was hateful or not.
- *Task B - Aggressive behavior and Target Classification:* classifying hate tweets as aggressive or not aggressive and identifying the hate speech target, whether individual (i.e., single human) or generic (i.e., group)

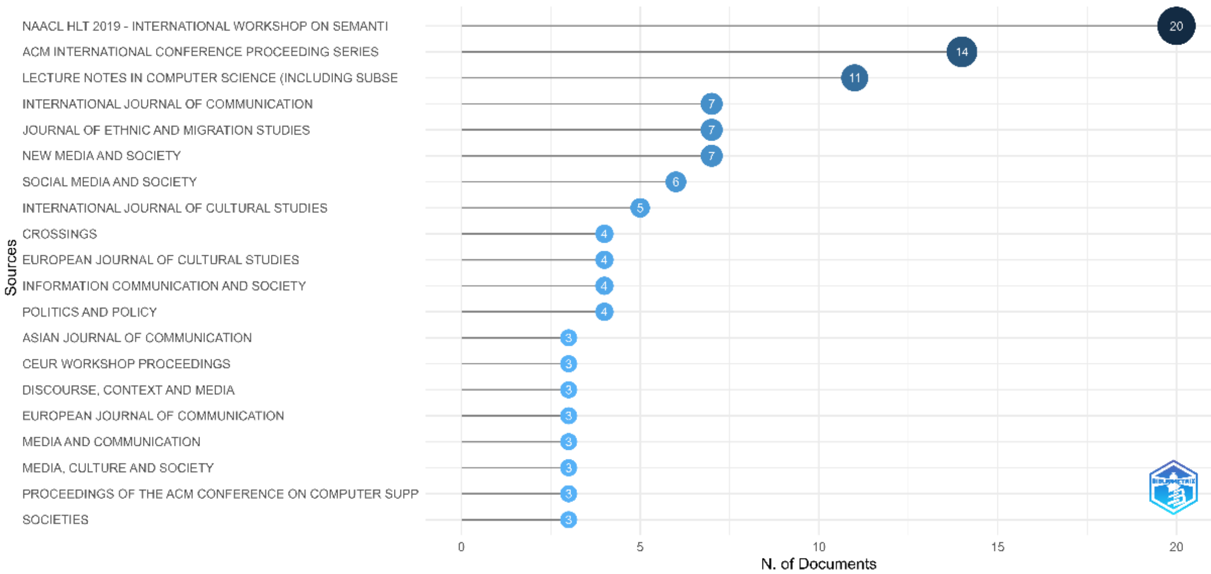


Figure 2.2 Relevant Sources for this Domain

[24] SemEval 2019 Task 5. Available at https://competitions.codalab.org/competitions/19935, Accessed on July 1, 2022

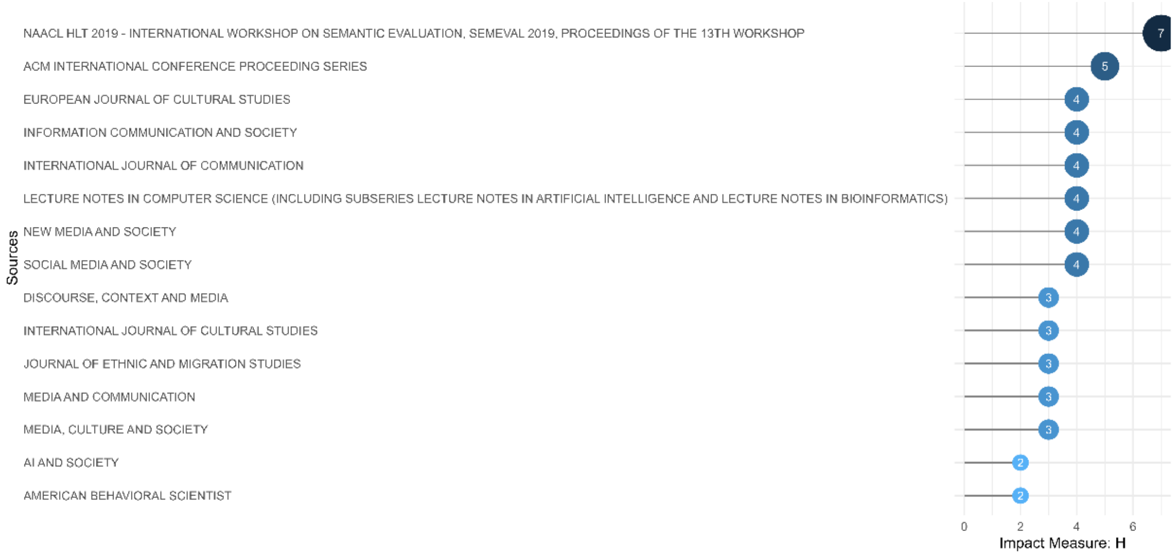


Figure 2.3 Impactful Sources for this Domain

Subsequently, many AI-based hate speech-related studies considered migrants or refugees as one of the use cases. After the SemEval 2019 Task 5, the following relevant sources are the *ACM International Conference Proceedings Series* and *Lecture Notes in Computer Science* - many prestigious information science conferences come under these two sources. Then, we have a few journals in these two lists (refer to Figures 2.2 and 2.3). Some of these journals, such as *Social Media and Society*, *New Media and Society*, and *AI and Society*, are well-known venues for computational social science research. Next, we explore - who are the most productive and impactful authors in this domain (Figures 2.4 and 2.6). Also, we probe the temporal trend of these authors (Figure 2.5).

2.4 Influential Authors and Collaboration Network

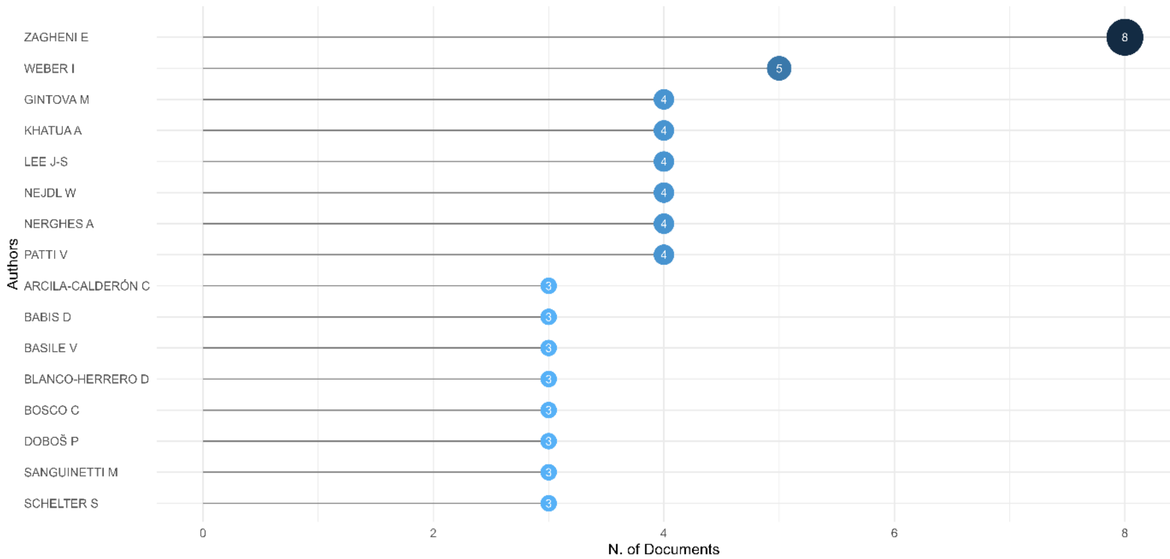


Figure 2.4 Productive Authors of this Domain

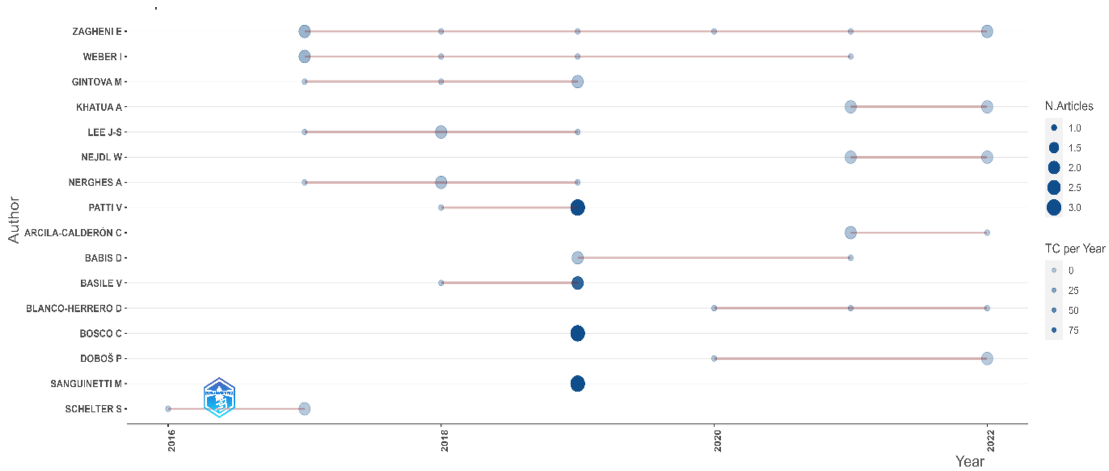


Figure 2.5: Temporal Trend of Productive Authors

The two most influential authors in this domain are Emilio Zagheni (8 papers) and Ingmar Weber (5 papers). However, our author network analysis highlights that some of these papers are jointly co-authored by them and their colleagues. These two authors are also the most impactful authors (i.e., high local h-index) compared to others. It is worth noting that their local h-indexes (e.g., Zagheni -6 and Weber – 4) are significantly lower than their Google Scholar h-indexes (e.g., Zagheni[25] - 31 and Weber[26] – 54). Some of the other prominent authors, both in terms of productivity and impact, in this domain are Ju-Sung Lee and Adina Nerghes. They have jointly written all four papers. Next, we explore the author's production over time (refer to Figure 2.5). We note that Zagheni and Weber have consistently published since 2017, while Lee and Nerghes have published all four papers from 2017 to 2019 – not in the last few years.

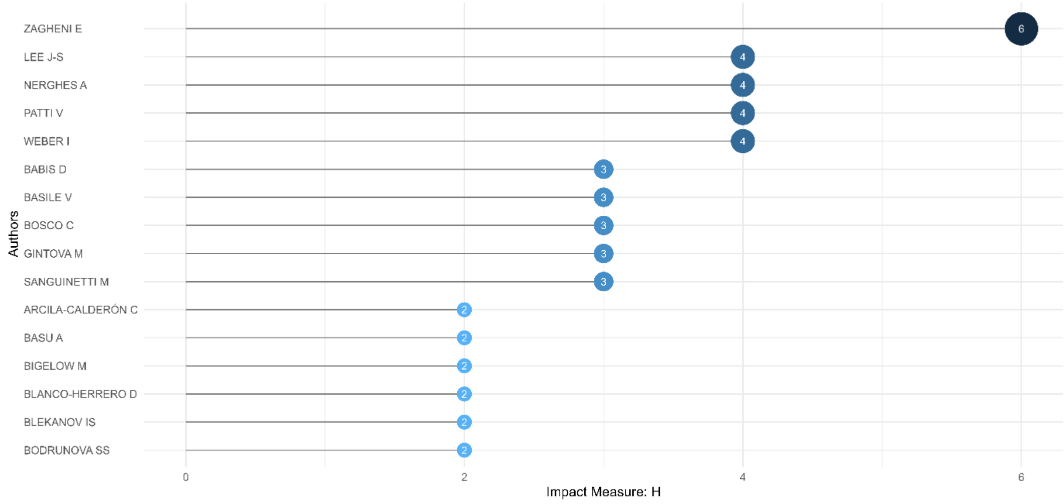


Figure 2.6 Impactful Authors based on local H-index

[25] https://scholar.google.com/citations?user=jNcPX0UAAAAJ&hl=en , Accessed on November 13, 2022
[26] https://scholar.google.com/citations?user=3YDUbP0AAAAJ&hl=en , Accessed on November 13, 2022

Figure 2.7 reports the corresponding author's country. First, we note that the instances of single-country papers (SCP) are significantly higher than multiple-country papers (MCP). Second, we find a strong correlation between this country-wise distribution and prominent host nations (Figures 1.3, 1.7, and 1.8 of Chapter 1). As aforementioned, from the perspective of migrants, the most preferred host nation is the USA, and more than 50% of international migrants live in North America and Europe (World Migration Report, 2020). Accordingly, most of our corresponding authors are from the USA, and leading European countries followed it. The presence of Canada and Indonesia (after the Rohingya crisis) in this list is also not counterintuitive. Hence, Figure 2.7 implicitly indicates that computational social science research reflects the concerns of society.

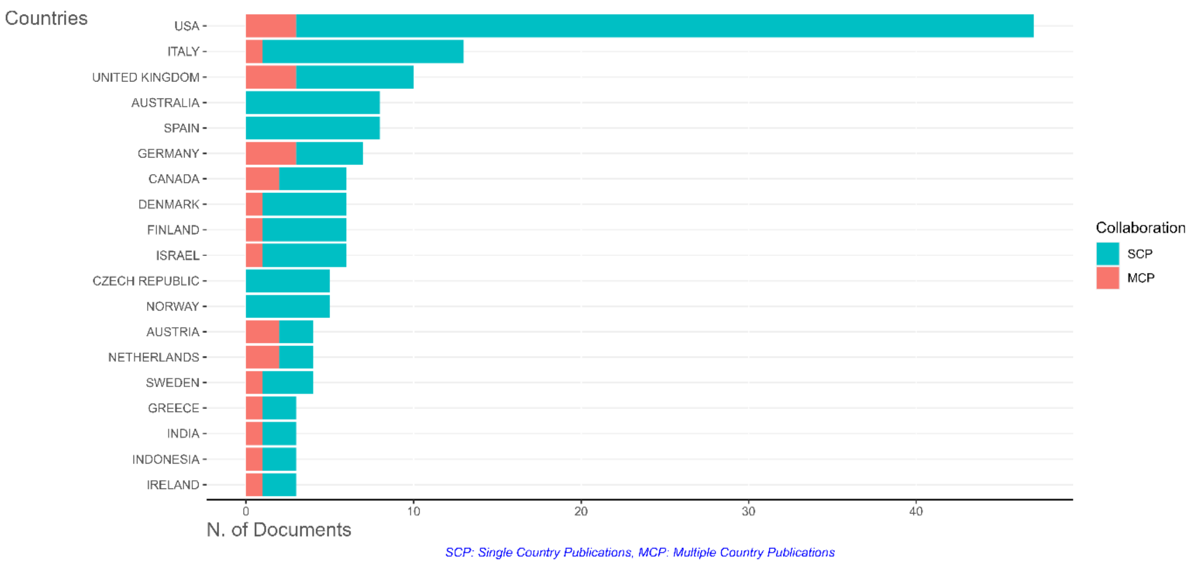


Figure 2.7 Corresponding Author's Country

Figure 2.8 reports author collaboration networks, revealing a few prominent clusters. We find that these clusters are in accordance with previous author-related analyses (Figures 2.4 and 2.6). For example, the leading authors of the prominent violet cluster are Emilio Zagheni and Ingmar Weber. Similarly, the top authors of the pink cluster are Viviana Patti, Manuela Sanguinetti, Valerio Basile, and Cristina Bosco. This team deserves the credit for designing the impactful SemEval 2019 Task 5. As noted, this SemEval 2019 Task 5 was the trigger point for migrant-related research, specifically hate speech-related research toward migrants. Another significant cluster, rather a dyad, was formed by Ju-Sung Lee and Adina Nerghes. The red cluster, probably the biggest in author counts, comprises leading authors like David Blanco-Herrero and Carlos Arcila Calderón. Last but not least, the research papers of this dissertation are also cropped up in this diagram, i.e., the green dyad comprises Aparup Khatua and Wolfgang Nejdl. Next, we briefly probe whether the research themes of these leading clusters are broadly the same or not.

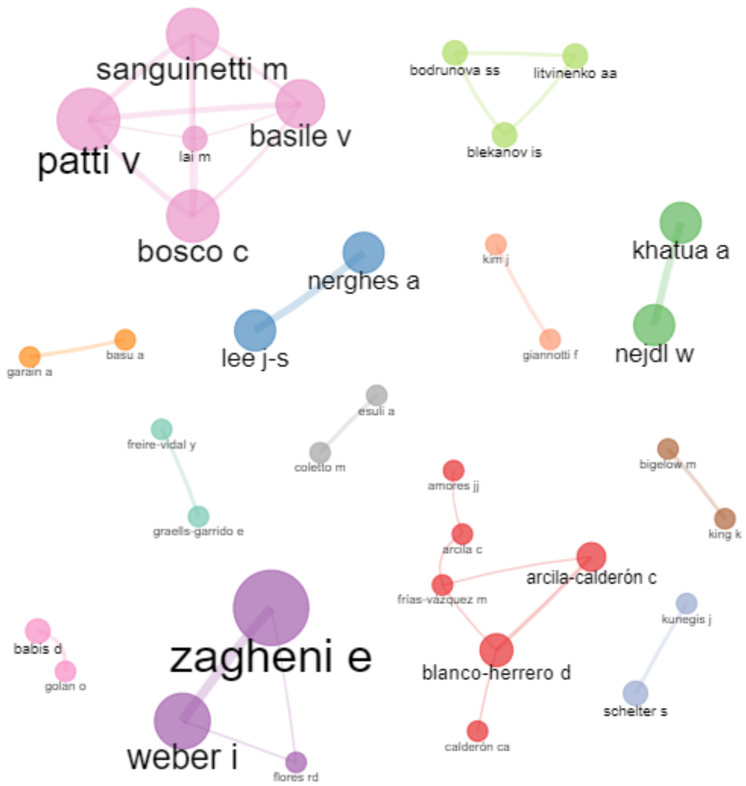


Figure 2.8 Author Collaboration Network

Some of the exciting papers from the **violet cluster** are Dubois et al. (2018), Fiorio et al. (2017), Rampazzo et al. (2021), Stewart et al. (2019), and Zagheni et al. (2017). Zagheni et al. (2017) tried to address the United Nations 2030 Agenda for SDGs. It pointed out that the lack of migrant-related data is one of the significant constraints for estimating stocks of migrants. Hence, this paper argued that the Facebook advertising platform could allow us to perform a digital census. Subsequently, Rampazzo et al. (2021) explored the Facebook advertising platform in the context of the UK for estimating migrant stock. This paper compared the Labour Force Survey and Facebook advertising data. Fiorio et al. (2017), another fascinating study by this group, explored geo-referenced tweets of around 62,000 users from 2010 to 2016 to analyze the US internal migration flows. This study demonstrated the viability of social media data as an alternative to survey data for probing short-term mobility and long-term migration. In addition to estimating migrant stock, this group also explored the social integration or cultural assimilation of migrants. For instance, Dubois et al. (2018) investigated Facebook data for advertisers, i.e., interests expressed on online platforms, to explore the assimilation of Arabic-speaking migrants in Germany and found demographic profiles, such as language or country of origin, are influential factors for assimilation. Similarly, Stewart et al. (2019) also explored Facebook data to investigate the cultural assimilation of Mexican immigrants in the U.S. This paper considers musical tests or interests as a proxy to investigate cultural convergence. Thus, the overall research themes of this cluster are estimating migrant stocks and understanding the social integration of migrants.

As aforementioned, the most influential paper of the **pink cluster** is the SemEval-2019 Task 5 (Basile et al., 2019). Subsequently, Sanguinetti et al. (2018) developed an annotated corpus of about 6,000 Italian tweets in a follow-up project. This paper has refined the hate speech detection task against immigrants by identifying more refined categories: aggressiveness, offensiveness, irony, and stereotype. Additionally, they have labeled five degrees of intensity on an experimental basis. Poletto et al. (2021), a recent paper by this group, performed a systematic review of various resources and benchmark corpora for hate speech detection. Thus, the dominant research theme of this group revolves around hate speech toward migrants.

The research theme of the **red cluster** is broadly similar to the pink cluster. The influential papers of this cluster, such as Calderón et al. (2020a; 2020b; 2021; 2022), are also around hate speech and social acceptance of migrants and refugees. For example, Calderón et al. (2021) probed the Boat Aquarius event. This boat "carried 630 migrants and arrived at the port of Valencia in June 2018 after being rejected by other European countries. The announcement by the Spanish government on 11 June 2018 that the country would welcome the drifting boat was the first high-profile political move of the new Socialist executive, who had reached power 10 days after a successful motion of censure in the national Parliament" (Calderón et al., 2021, p.1). This paper analyzed 24,254 Spanish tweets around this announcement and observed that "a significant part of messages expressed rejection or hate—often supported by stereotypes and lies—towards refugees and migrants and towards politicians" (Calderón et al., 2021, p.1). Similarly, Calderón et al. (2020a) also explored Spanish tweets to understand hate speech, specifically verbal rejections, toward migrants and refugees. They found that "rejection toward migrants was significantly bigger than over refugees" (Calderón et al., 2020a, p.40). In a similar line, Calderón et al. (2022) also explored hate speech and social acceptance of migrants in the European context. This study considered 847,978 geolocated tweets and employed a longitudinal analysis from 2015 to 2020. Interestingly, this study found that a higher share of immigrants in a region positively affects the support for immigrants. Notably, this study also pointed out that "regions with greater support recorded a lower level of hate speech on Twitter" (Calderón et al., 2022, p.33).

The **blue cluster** has produced four influential papers: Lee & Nerghes (2017, 2018) and Nerghes & Lee (2018, 2019). Lee & Nerghes (2018) explored YouTube comments (i.e., 46,313 comments posted in response to a sympathetic YouTube video "The European Refugee Crisis and Syria Explained" and 13,871 comments to the antipathetic video "What Pisses Me Off About the European Migrant Crisis") to understand the impact of labels such as migrant crisis vis-à-vis refugee crisis. To explore opinions toward such labels, the authors have performed topic modeling and sentiment analysis and concluded that the "use of such labels has the potential to dictate the ways in which displaced people are received and perceived" (Lee & Nerghes, 2018, p.1). On a similar line, Nerghes & Lee (2018) also explored "deserving" refugee versus the "undeserving" migrant dichotomy. To probe this, the authors considered 369,485 tweets posted immediately after the unfortunate drowning of Alan Kurdi, the three-year-old Syrian boy, and employed network and sentiment analysis. Regression-based analysis confirms their proposed hypothesis, i.e., "refugee-related hashtags would carry more positivity and less negativity than migrant-related hashtags" (Nerghes & Lee, 2018, p. 279). Lee & Nerghes (2017) also emphasized the importance of labels. They noted that labels "such as 'migrant' or 'immigrant', are embedded in less sympathetic comments than those labels indicating a need to escape war-torn regions or persecution (e.g., asylum seeker or refugee)" (Lee & Nerghes, 2017, p.1). In another follow-up study, Nerghes & Lee (2019) compared the perspectives and narratives of mainstream media and Twitter in response to the tragic death of Alan Kurdi. They have employed topic modeling, especially Latent Dirichlet allocation (LDA), to extract the themes from their corpus. They found that mainstream media and social media are complementary because social media users share news items on the Twitter platform. Overall, mainstream media was broadly neutral, but they politicized the crisis and were geo-politically orientated. Contrarily, Twitter deliberations were relatively sympathetic. Overall, the research theme of this cluster is opinion mining and offers a nuanced understanding of how labels influence our opinions and sentiments. They have explored Twitter, YouTube, and mainstream media to probe this issue.

To sum up, prior studies primarily probed the Twitter platform. Also, a few studies examined Facebook, YouTube, and mainstream media. Primarily, these studies can be categorized into two domains: opinion mining, specifically hate speech toward migrants and refugees, and social integration. Opinion mining studies also explored the effects of labels, i.e., refugees vis-à-vis migrants, on our perceptions or sentiments. A few studies also considered user-generated data to predict migrant stocks.

2.5 Influential Articles

This section briefly reviews the influential papers. Figure 2.9 reports the list of most cited articles, and we note that the citations are highly skewed. Obviously, the highest cited paper was the SemEval 2019 Task 5 by Basile et al. (2019). Notably, Basile et al. (2019) pointed out that this was "one of the most popular tasks in SemEval-2019 with a total of 108 submitted runs for Subtask A and 70 runs for Subtask B, from a total of 74 different teams" (pg. 54). They concluded that this high number of participations "confirms the growing interest of the community around abusive language in social media and hate speech detection in particular" toward the marginalized section of the society (e.g., migrants) (pg. 62). Consequently, migration-related research gained momentum in 2019 compared to the previous years. In addition to Basile et al. (2019), we have also discussed Sanguinetti et al. (2018) and Zagheni et al. (2017) in the preceding section. So, we are not repeating them here for the sake of brevity.

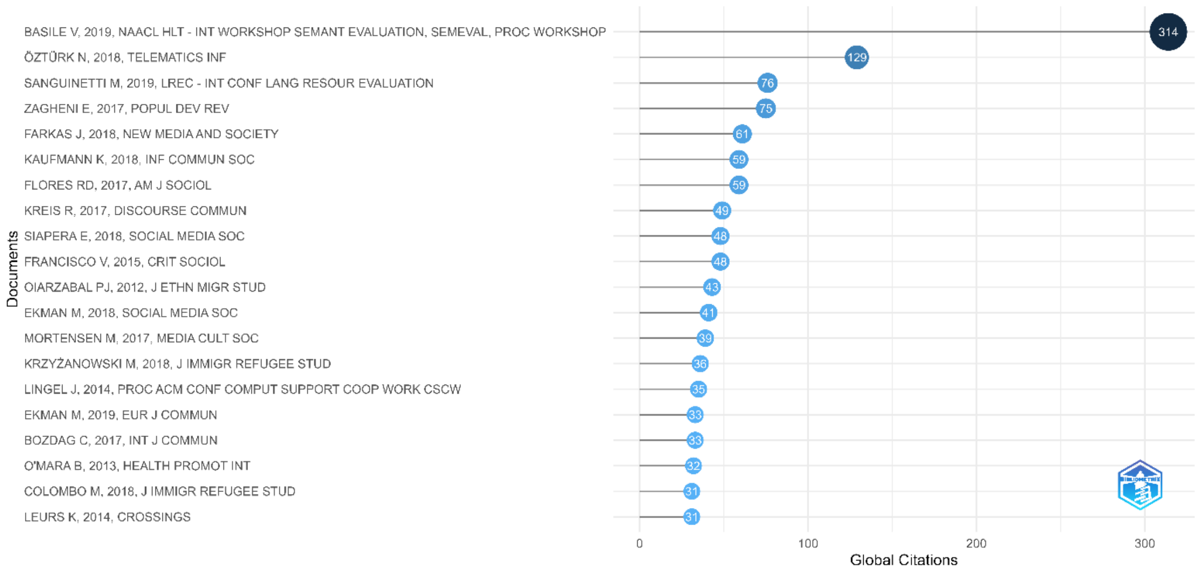


Figure 2.8 Most Cited Articles in this Domain

Öztürka & Ayvazb (2018) investigated the public opinions and sentiments toward the Syrian refugees on the Twitter platform. This bilingual study noted that Turkish tweets (35% positive) displayed more positive sentiments than English tweets (12% positive tweets). English tweets were primarily neutral (48%) or expressed negative sentiment (40%). It is worth noting that during this crisis, Turkey welcomed many Syrian refugees. This pattern is in accordance with the prior study by Calderón et al. (2022), i.e., "regions with greater support recorded a lower level of hate speech on Twitter" (pg. 33). Similarly, Kreis (2017) analyzed 100 tweets with the hashtag *#refugeesnotwelcome,* and probed the reasons, i.e., arguments and strategies to justify the apathy toward migrants or refugees.

This study finds that the "argumentation lies on sharing and recirculating events, stories, articles, or images where refugees and immigrants are depicted as criminals and exploiters" and the social media deliberations "discursively construct a national or European identity on the one hand, and migrants as 'criminals' and 'out-group members' on the other, which underscores racist and ethnocentric ideologies and discourses circulated by nationalist-conservative Europeans" (Kreis, 2017, p.511). The large-scale study by Siapera et al. (2018) echoed this view. This study explored around 7.5 million refugee crisis-related tweets during the period from October 2015 to May 2016 and noted the deliberation or dominant frames are "revolving around security and safety on one hand and humanitarianism on the other" (Siapera et al., 2018, p.1). Security and safety concerns are triggered not only due to the depiction of refugees and immigrants as criminals but also due to violence by refugees or migrants, such as attacks in Köln on New Year's Eve 2015 or the terrorist attacks in Paris in 2015. In this context, Ekman (2018) probed the emergence of vigilante gangs, who "claim to protect citizens from alleged violent and sexual attacks by refugees" in Europe, and these "racist actors use social media to mobilize and organize street politics targeting refugees/immigrants" (pg.1). Overall, Ekman (2018) concluded that social media communication aids these racists actors in anti-refugee mobilization, but generally they lack public support and media framing is negative about these vigilant gangs.

In addition to the above negative feelings, Siapera et al. (2018) also noted a humanitarian framing. Accordingly, Mortensen (2017, p. 1142) referred to the extant literature on "icons and appropriations to develop a theoretical framework for how appropriations construct, confirm, and contest icons and how personification constitutes the main link between icons and their appropriations" in the context of iconic imagery of Alan Kurdi. Mortensen (2017) explored the appropriations of the Kurdi imagery with the hashtag *#humanitywashedashore* and identified two modes of appropriations: *Decontextualization* (i.e., isolating "the figure of the drowned child and includes appropriations within the genres of realistic drawings") and *Recontextualization* (i.e., "inserts the figure into new contexts") (pg. 1150). This study concludes that "Decontextualizations raise the question of how audiences are to cope with the traumatic reality represented so bluntly in the Kurdi imagery and deal with the interplay between esthetic expression and the tragic death of a refugee child. Recontextualizations point to the moral implications of focusing on one particular victim out of many and ask how we are to assume political responsibility on individual and collective levels in the face of the humanitarian catastrophe" (Mortensen, 2017, p.1158). Similarly, Bozdag & Smets (2017) qualitatively analyzed 961 tweets from Turkey and Belgium as a response to the iconic image of Alan Kurdi. Interestingly, this study did not notice a radical shift in the discourses and representations, but “references to Kurdi were incorporated into preexisting discourses on and representations of refugees, thus offering different actors in the public debate on refugees with new symbols and motifs to construct meaning” (Bozdag & Smets, 2017, p.4046).

An exciting study by Farkas et al. (2018) employed an online ethnographic case study-based approach to explore the propagation of fake Islamist propaganda on the Facebook platform by using cloaked pages – in other words, radical Islamists did not create these counterfeit pages. Farkas et al. (2018) examined Danish Islamist Facebook pages that used to post aggressive and hostile comments toward Danes and Denmark, and consequently, these posts used to trigger anti-Muslim sentiments, specifically toward refugees and immigrants. The authors pointed out that it is difficult to verify whether a profile page is credible or not “given Facebook’s design (which provides almost unlimited anonymity and security to page owners)” (Farkas et al., 2018, p.1863). Another study by Flores (2017) explored the effects of policies on public attitude and behavior in the context of a stringent anti-immigration policy like Arizona’s SB 1070. Analysis of 250,000 tweets reveals two interesting patterns: first, this law “had a negative impact on the average sentiment of tweets regarding immigrants, Mexicans, and Hispanics, but not on those about Asians or blacks”; and second, the policy had more influence on the behavior compared to the attitude of social media users (Flores, 2017, p. 333).

Kaufmann (2018) explored the struggles of Syrian refugees in Vienna. This qualitative study revealed the importance of “smartphones to cope with everyday challenges ... in the contexts of: place-making and geographical orientation; information access and self-help; language learning and translation; and ‘doing family’. Hence, refugees are both emotionally attached to and technically dependent on their devices” (pg. 882). In short, this paper argues that smartphones help refugees to settle in the host country. Similarly, Fracisco (2015) also pointed out the role of communication channels like Skype or Facebook for undocumented immigrants in the USA, especially when the “economic necessity of working abroad and legal conditions deter family reunification” (pg. 173). Along similar lines, Lingel et al. (2014) interviewed 26 transnational migrants in New York City and noted that social media platforms, such as Facebook, allow these migrants to be in “touch with friends and family abroad and documenting everyday urban life” (pg. 1502). Similarly, Leurs (2014) also emphasized the role of digital connectivity for young Somalis stranded in Ethiopia. Additionally, O’Mara (2013) argued that social media platforms, such as YouTube, Vimeo, and Flickr, facilitate digital video production and promote health and well-being in the context of culturally and linguistically diverse refugee and migrant communities of Australia. To sum up, on the one hand, social media can propagate misinformation and racial views to trigger hatred toward migrants and refugees. However, on the other hand, social media platforms can play a positive role for migrants by providing digital connections with their families and friends or helping them to learn new languages or access information about the host nations.

2.6 Evolution of the Discipline

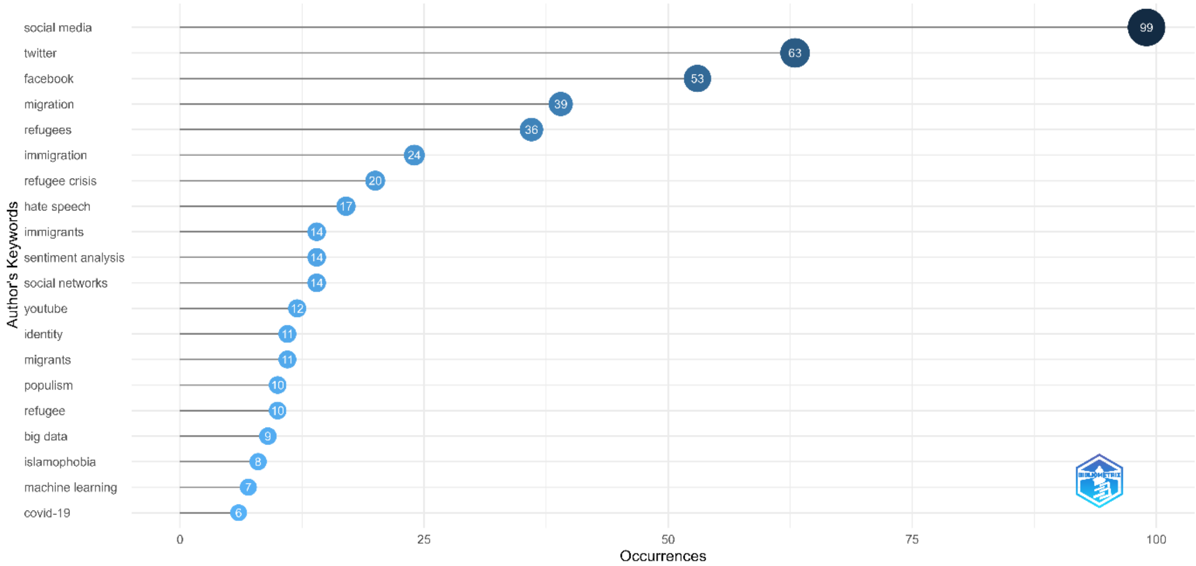


Figure 2.9 Prominent Keywords Used by Authors

Figure 2.9 lists the main keywords of 395 articles, and Figure 2.10 reports the co-occurrence network of keywords. Keywords in Figure 2. 10 can be classified into four distinct categories as follows: *migrant-specific words* (migration, refugees, immigration, refugee crisis, and immigrants), *platforms* (social media, social networks, big data, Twitter, Facebook, and YouTube), *methodology* (sentiment analysis and ML), and *specific research themes* (hate speech, islamophobia, populism, identity, and covid-19). The co-occurrence network reveals three main clusters (i.e., red, blue, and green) and two independent but critical nodes/clusters, i.e., *xenophobia* and *disinformation*. The red cluster associates *social media* with *immigration*, *immigrants,* and *refugee crises*. In this cluster, some interesting smaller nodes are Islamophobia, Covid-19, and YouTube. The Blue cluster associates *Twitter* with *refugee(s)* and *asylum seekers*. It also features *framing*, *sentiment analysis*, and *hate speech*. Finally, the green cluster connects *social networking sites*, like *Facebook,* with *racism* and *identity*. Overall, Figures 2.9 and 2.10 indicate that *Twitter* is one of the prominent social media platforms for migrant-related research, and prior studies primarily used an ML-based approach to perform *sentiment analysis* of tweet feeds. Some research themes, such as *hate speech* or *Islamophobia*, portray xenophobia toward migrants.

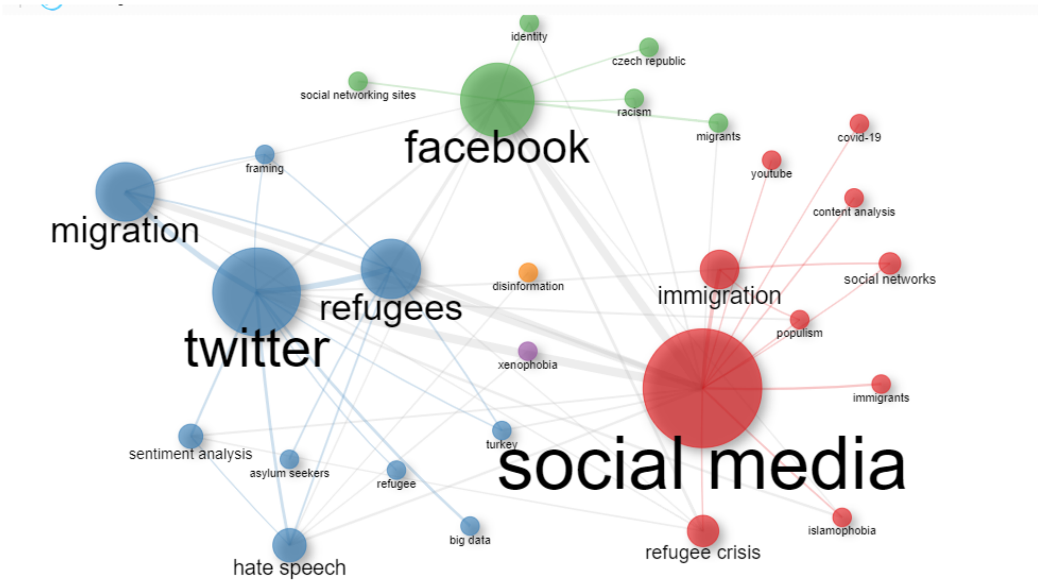


Figure 2.10 Co-occurrence Network

Finally, we employed the thematic map to identify this domain's different research clusters. Callon et al. (1991) developed the strategic diagram, commonly known as a thematic map, to identify research clusters using centrality and density. Callon et al. (1991, p.164) measured the Centrality of a cluster as "the intensity of its links with other clusters. The more numerous and stronger are these links, the more this cluster designates a set of research problems considered crucial by the scientific or technological community". Callon's centrality (c) can be measured as $c = 10 \times \sum e_{kh}$, where k is a keyword belonging to the theme, and h is a keyword belonging to other themes. Accordingly, "*Density* characterizes the strength of the links that tie the words making up the cluster together. The stronger these links are, the more the research problems corresponding to the cluster constitute a coherent and integrated whole. It could be said that density provides a good representation of the cluster's capacity to maintain itself and to develop over the course of time in the field under consideration" (Callon et al., 1991, p.165). Callon's density (d) can be measured as $d = 100 \times \sum \frac{e_{ij}}{w}$, where i and j are keywords belonging to the theme and w is the number of keywords in the theme.

Callon et al. (1991) argued that we could categorize the literature into four quadrants/clusters using these two dimensions as follows:

- *Quadrant 1 (high centrality and high density):* This cluster is "central to the general network (they are strongly connected to other clusters) and have intense internal links (they display a high degree of development)" (Callon et al., 1991, p.166). In other words, this quadrant is the core theme of a particular research domain. Cobo et al. (2011) label it **Motor Themes**.
- *Quadrant 2 (high centrality and low density):* As defined, this quadrant is central, i.e., "strongly connected to other clusters, but the density of their internal links is relatively low" (Callon et al., 1991, p.166). So, this quadrant is central or essential but underdeveloped because density or internal links are not developed, and the graphical interface labels it as **Basic Themes** (Cobo et al. 2011)
- *Quadrant 3 (low centrality and high density):* This quadrant is peripheral (i.e., not central) but developed. In other words, these specialized clusters are well-developed or well-researched but not well-connected with other clusters or sub-networks of the research domain. Callon et al. (1991, p.166) suggested that these might be "clusters which at an earlier time were central, but which - while remaining the object of significant investments (it is not so difficult to explain such permanence) - have been progressively marginalized, generating less and less interest". Cobo et al. (2011) label it as **Niche Themes**.
- *Quadrant 4 (low centrality and low density):* This quadrant is peripheral (i.e., not central), as well as underdeveloped (i.e., internal links within this cluster are weak). Callon et al. (1991) considered this as the *margins of the network*. Cobo et al. (2011) argued that this quadrant comprises **Emerging** *or* **Declining Themes**. Hence, Callon et al. (1991, p.166) suggested that "only a dynamic analysis (the evolution of a network over several periods) or a comparative one (the relationship of the network with other networks) allows us to determine their contribution to the field."

Following the suggestion of Callon et al. (1991), we have employed a dynamic approach. Based on the publication trend (i.e., Figure 2.1), we have categorized the corpus into three phases as follows: 2009 to 2017 (i.e., mostly this phase was stagnant with 10 to 20 papers annually), 2018 to 2019 (i.e., the literature gained momentum from 20 to 30 papers in a year to around 70 papers in a year) and 2020 to 2022 (i.e., the publication trend gained stability and the world faced on the worst-ever pandemics during these years). Figures 2.11, 2.12, and 2.13 reveal some interesting insights.

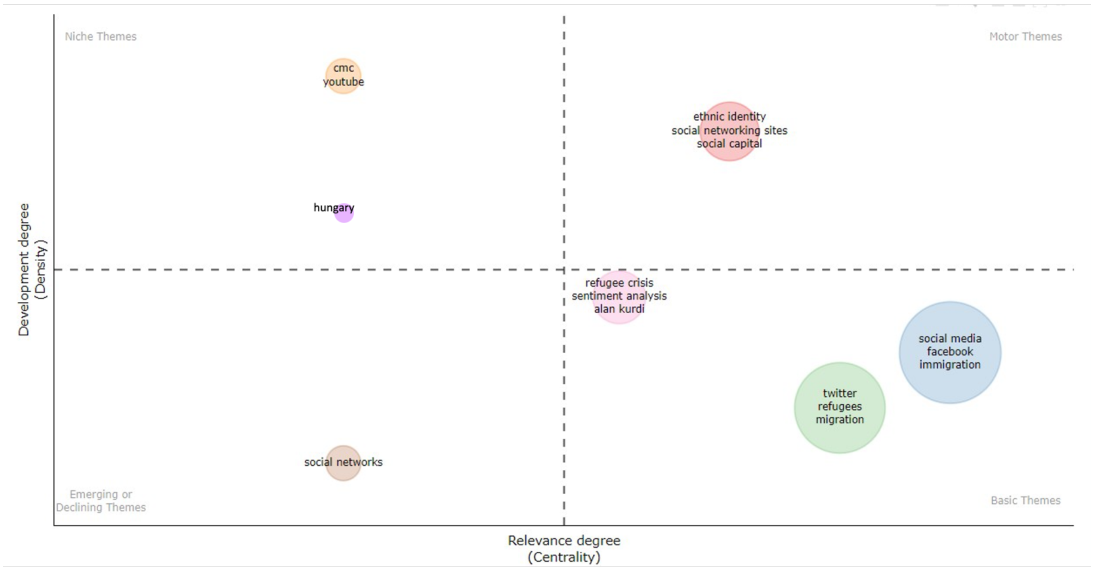


Figure 2.11 Thematic Map 2009 to 2017

First, *Motor Themes*, i.e., *Quadrant 1*, have evolved significantly over time. Initially, the dominant themes used *social networking site* data to probe *ethnic identity*. However, in the middle phase, specifically after the unfortunate Alan Kurdi incident in 2015, some significant themes were the *refugee crisis*, *hate speech*, and *forced migration*. Intuitively, the genesis of the hate speech cluster can be traced back to the SemEval 2019 Task 5 by Basile et al. (2019). *Supervised machine learning* on *big data* gained prominence in this phase. In the last phase, *Facebook* and *YouTube* emerged as social media platforms. The European refugee crisis also gets captured through terms like *Europe*, *Turkey*, *South Africa*, and *Nationalism*. Methodologically, we see the coexistence of *Bayesian methods* and *qualitative methods* like *content analysis* or *discourse analysis*.

Second, for *Basic Themes*, i.e., *Quadrant 2*, we didn't notice any significant changes over time. We find platform-related clusters (i.e., *social media*, *Twitter*, or *Facebook*) are consistent across all the phases. Following Callon et al. (1991) and Cobo et al. (2011), it can be concluded that these platforms are central but underdeveloped. Thus, the consistent presence of these themes in Quadrant 2 also justifies the importance of this dissertation. Alan Kurdi was also featured in the initial phase, but *Alan Kurdi* was a symbolic representation of the struggles of refugees from the Middle East and became a synonym for the *refugee crisis*. *Covid-19* also emerged as a significant cluster during this phase because refugees and migrants became economically vulnerable due to lockdowns across nations and faced racial abuses. For example, even after the World Health Organization clarifications, Trump, the former US President, has repeatedly referred to the Covid-19 virus as the Chinese virus[27]. Thus, in the initial phase, Chinese migrants, especially for their eating habits and wet markets in China, were blamed for this Covid-19 crisis[28]. Additionally, we find that core concerns such as *solidarity*, *social justice*, and *refugee voices* are underdeveloped. Therefore, this dissertation has attempted to address underdeveloped concerns like solidarity (in Chapter 3) or refugee voice (in Chapter 6).

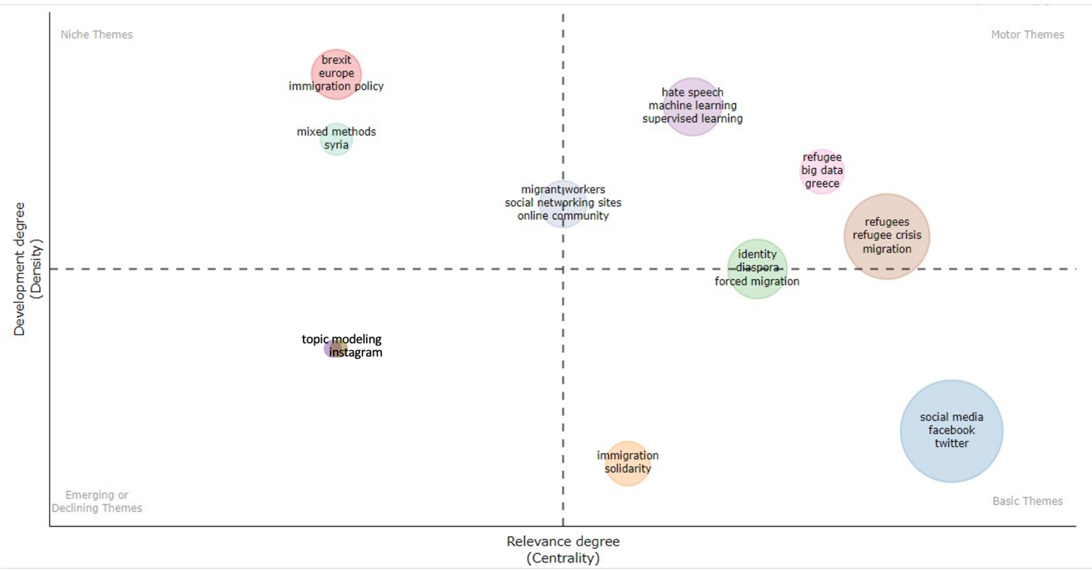


Figure 2.12 Thematic Map 2018 to 2019

Third, we find an interesting pattern for *Niche Themes*, i.e., *Quadrant 3*. For example, the number of niche themes has increased over the years. Initially i.e., from 2009 to 2017, there were just a few niche themes like *CMC* (i.e., computer-mediated communication) and *YouTube*. Interestingly, YouTube became a motor theme in phase 3 (i.e., 2020 to 2022). We noticed niche themes like *Brexit*, *Europe*, *immigration policy*, and *Syria* in the middle phase. As aforementioned, Alan Kurdi, the unfortunate *Syrian boy*, was a basic theme in the previous thematic map. In the final phase, i.e., 2020 to 2022, we find a plethora of niche themes like the *European migrant crisis*, *polish immigrants*, *anti-immigration* policies, concerns for *transnational families,* and consequently, *social movement* for these migrants. The researcher also employed *critical discourse analysis* and probed platforms like *WhatsApp*. This pattern implicitly hints that these issues are well-researched but not well-connected with other clusters or research domains, such as AI-based research. Our review indicates that the struggles of migrants and migrant

[27] The New York Times: Trump Defends Using 'Chinese Virus' Label, Ignoring Growing Criticism, Available at https://www.nytimes.com/2020/03/18/us/politics/china-virus.html, Accessed on July 1, 2022
[28] BBC News: Sinophobia: How a virus reveals the many ways China is feared, Available at https://www.bbc.com/news/world-asia-51456056, Accessed on July 1, 2022

families are well-researched topics from the perspective of the social science domain, but it seems that AI-based research is more concerned with model accuracies than the actual societal impact. Intuitively, a change in approach, i.e., from AI to AI for Social Good, can make these topics more *central*, i.e., well-connected to other clusters.

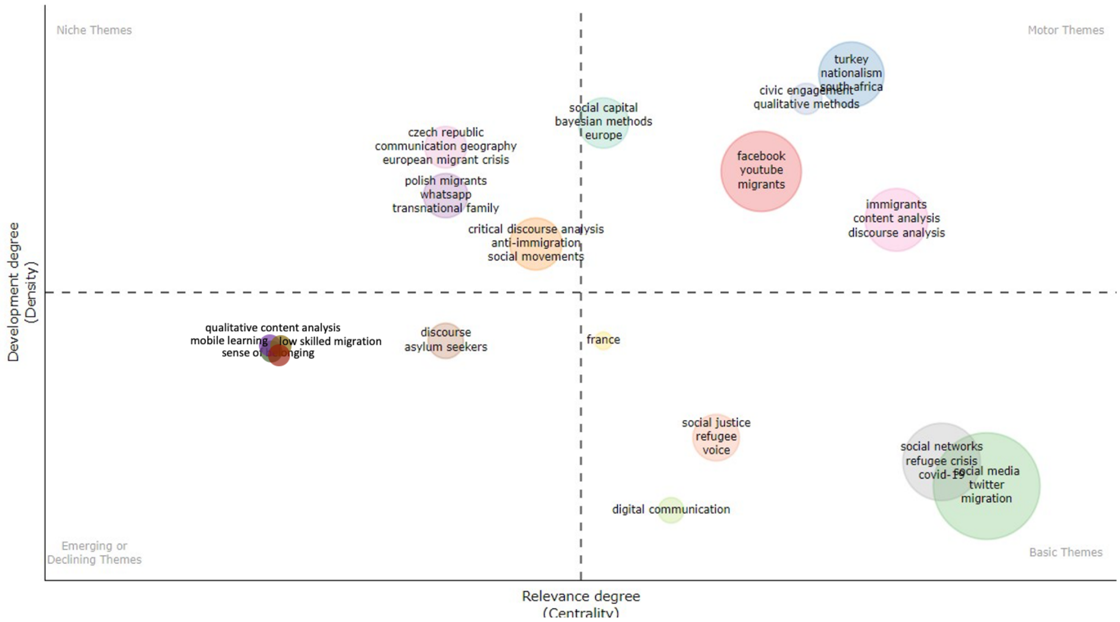


Figure 2.13 Thematic Map 2020 to 2022

Lastly, *Quadrant 4* can have both *Emerging and Declining themes*. For example, *social networks* in Phase 1 was an emerging theme, but *topic modeling* in Phase 2 was declining (because information science researchers started preferring advanced, unsupervised models over traditional approaches like topic modeling). In phase 3, we find multiple scattered themes like *mobile learning*, *low-skill migration*, *sense of belonging*, *asylum seekers*, and methodological approaches (e.g., *qualitative content analysis* of *discourse*). However, it will be difficult to conclusively comment on whether research clusters like the *sense of belonging* of *asylum seekers* will be an emerging trend. Probably, the answer will be affirmative if information science researchers use AI for Social Good in the coming days.

Figure 2.14 graphically reports the temporal evolution and connectivity (based on the occurrences of each keyword appearing in a theme) of research themes of Figures 2.11, 2.12, and 2.13 over the years through a *Sankey Diagram*. By considering the occurrences of keyword(s) in a theme, the connectivity might appear intuitive, but a nuanced analysis indicates the underlying complexity. For example, the apparent keyword for connectivity between 'social media 2009-2017', 'social media 2018-2019', and 'social media 2020-2022' is *social media*, but occurrences of keywords like *Facebook* and *Twitter* also played a role. Surprisingly, *Islamophobia* is one of the connecting keywords between 'social media 2018-2019' and 'social media 2020-2022'. Interestingly, *machine learning* is the connecting keyword between 'social media 2009-2017' and 'hate speech 2018-2019'. However, *xenophobia* is one of the connecting keywords between 'hate speech 2018-2019' and 'social media 2020-2022'. Overall, these connecting keywords can be broadly classified into four classes: migrant-related keywords (e.g., *refugee*, *refugee crisis*, *immigration*, and *immigration policy*), platform-related keywords (e.g., *social networking sites*, *social media*, *Twitter*, *Facebook*, and *big data*), and methodology-related keywords (e.g., *machine learning*, *sentiment analysis*, *discourse analysis*, and *framing*) and research-related themes (e.g., *hate speech*, *xenophobia*, *Islamophobia*, *identity*, and *populism*). The pattern is similar to Figures 2.9 and 2.10.

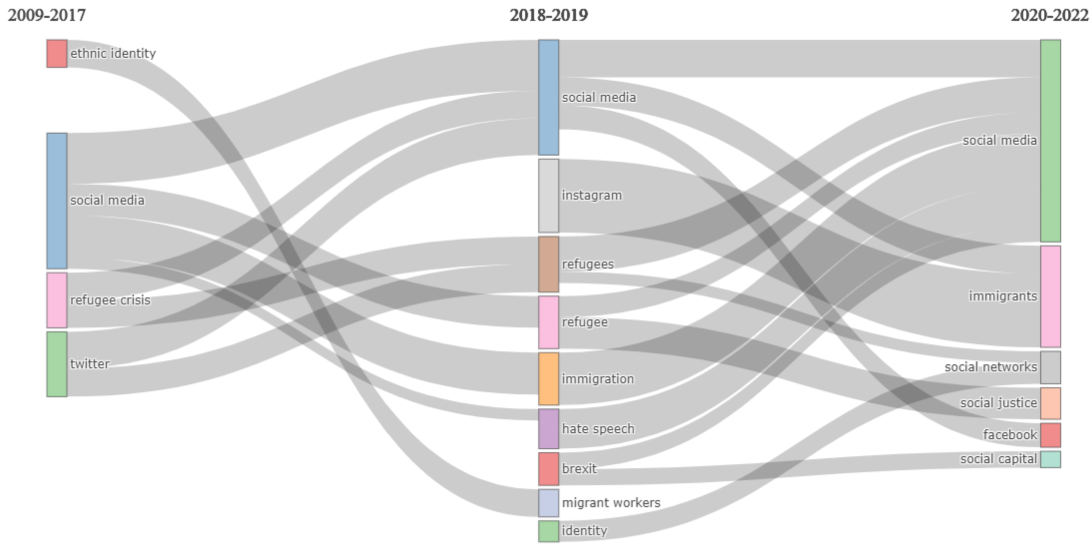


Figure 2.14 Evolution of the Domain

2.7 Conclusion

Migration has been a pressing issue for over a decade and has attracted the attention of the academic community, policymakers, social media, and mainstream media. Interestingly, the international migrant stock as a percentage of the world population has not increased drastically, but the absolute number has increased significantly – specifically in a few developed nations like the USA and European countries. Consequently, migration-related deliberations are also cropping up on various social media platforms. This allowed the academic community to probe user-generated data to explore societal opinion, estimate migrant stocks, or analyze the struggles of migrants –

this stream of research has gained impetus in recent times. However, it still offers various unexplored and underdeveloped terrains to probe. This chapter reviewed the articles published on migrants from 2009 to 2022 through a bibliometric approach. Also, this chapter attempted to explore the evolution of research themes over this period and to identify undeveloped themes for potential future research avenues.

We find that this domain was not so vibrant for nearly a decade (i.e., 2009 to 2017). However, it gained momentum only after 2017. The migrant-related papers got published in leading computer science conferences (e.g., NAACL workshop, ACM Proceedings, etc.), migration-specific social science journals (e.g., Journal of Ethnic and Migration Studies, Politics and Policy, European Journal of Cultural Studies, etc.) as well as interdisciplinary journals (e.g., Social Media and Society, New media and Society, AI and Society, etc.). Subsequently, our author-based analysis identified some influential and productive researchers in this domain, such as Emilio Zagheni, Ingmar Weber, Viviana Patti, Manuela Sanguinetti, Valerio Basile, Ju-Sung Lee, Adina Nerghes, and Carlos Arcila Calderón. Interestingly, most of these researchers are from the information science domain; however, this is not the case for influential articles. Influential articles are from both the computer science and social science domains. Our analysis of the author's collaboration network suggests that linkages within clusters are significantly dense, but connections between clusters are missing. This pattern might restrict the potential for cross-fertilization of ideas, and this is essential for the evolution and maturity of a research domain. Finally, we explored the development of thematic maps, revealing some interesting patterns. For example, the core themes (motor themes) have evolved significantly. Migration, as an issue, is intricately connected with other socio-economic issues. Accordingly, we note that research clusters related to Brexit or Covid-19 cropped up in our thematic maps. Notably, social media-related themes were mainly in Quadrant 2 (i.e., basic themes), indicating high relevance but low development. Hence, this observation justifies the overarching theme of this dissertation, i.e., social media mining of migration discourse.

Our review doesn't reveal a strong symbiotic relationship between ML-based big data analysis by computer science researchers and the qualitative approach of small-scale data analysis by social science researchers. For instance, following the SemEval-2019 Task 5, a plethora of papers employed machine-learning-based approaches to identify hate speech toward migrants on social media platforms. Similarly, social science researchers are theoretically probing the genesis of xenophobia. Ideally, AI-based research should draw insights from these theoretical works to interpret their findings, but this amalgamation is mostly missing. We also note that social media mining by AI-based research primarily probed societal opinion (and a few studies used user-generated data to estimate migrant stock or probed cultural assimilation). This stream of research generally noted the prevalence of negative sentiment toward migrants. Contrarily, social science researchers are considering small-scale social media data and mainly trying to understand the challenges faced by migrants in the host nation (and a few studies also probed societal opinion formation like - how cloaked Facebook pages can trigger racial biases toward refugees and migrants). Social science studies were mostly interested in understanding the mental trauma of refugees and migrants (due to prolonged separation from friends and families and hostile new environment), harassment they face in asylums, and struggles they go through in the host nation – ranging from gaining social acceptance to learning a new language. Social science studies also noted how social media platforms, like Facebook, WhatsApp, or Skype, allow them to connect for emotional support and to share information about the host nation. Interestingly, AI-based research rarely probed these issues.

Finally, this chapter has some limitations due to its bibliometric approach. First, we have considered computer science and social science domains. Due to this selection criteria, we might have missed some influential articles. A possible solution is to consider multiple domains, but the scope of this chapter did not allow us to do the same. Second, another potential limitation of our thematic analysis is the consideration of only authors' keywords. We have assumed that authors' keywords could capture the essence of an article most accurately. However, there might be interesting and insightful information in each article's title and abstract. Third, our selection of keywords for search queries was stringent to avoid non-relevant articles, and this might lead to the potential omission of relevant articles. However, a lenient approach will return voluminous, non-relevant articles. Hence, there will always be a trade-off between a rigorous and lenient process. Despite these limitations, the analysis presented in this chapter provided significant insights into migrant literature and the evolution of literature over the years.

Chapter 3: Social Perceptions & Behaviors Toward Migrants

3.1 Societal Opinions: Do we need a Deep Dive?

As we have noted in Chapter 1, the number of international migrants has reached 272 million in 2020 from 150 million in 2010 (International Organization for Migration & United Nations, 2000). Generally, these migrants come from economically weaker or politically disturbed countries, assuming they would lead a better life in the host nations. However, migrants not only change the demographic fabric of the host nation but also impact the host nation's politics, law enforcement, economic, and labor market conditions (Aswad & Menezes, 2018). Consequently, a specific segment of the host nations can be apprehensive about these international migrants. Recent political discourses during the 2016 and 2020 USA Presidential elections or Brexit referendum reveal an apprehensive view toward migrants (Khatua & Khatua, 2016; Ogan et al., 2018; Waldinger, 2018). On the contrary, the other segment of society can be sympathetic toward migrants. This segment is concerned about the inequality and discrimination toward migrants.

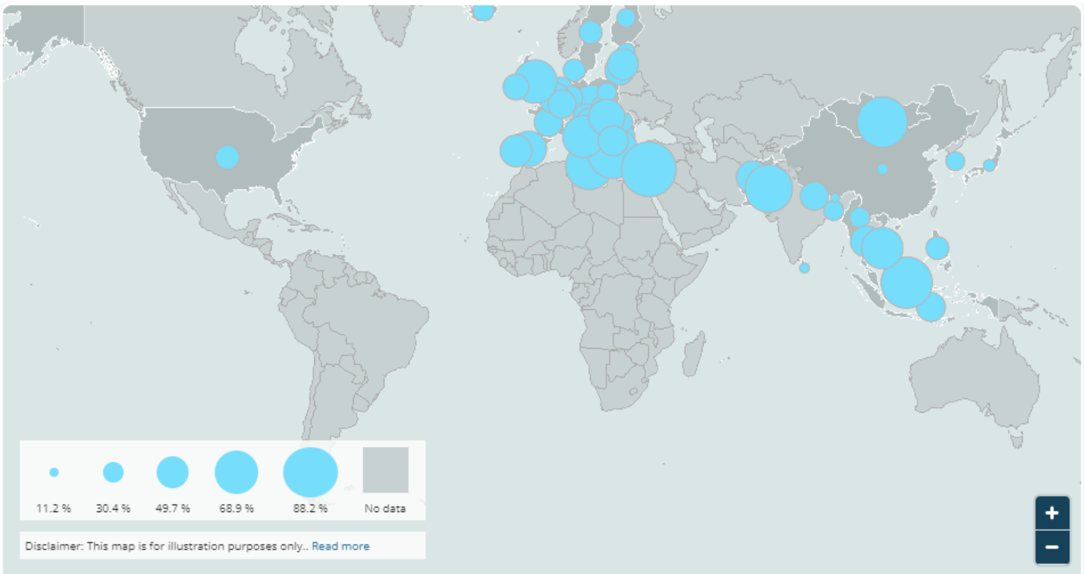


Figure 3.1 Public Opinion on Immigration Level[29]

Figures 3.1 and 3.2 effectively capture these aspects. Figure 3.1 reports the apprehension towards migrants, i.e., percentage of adult respondents who would like to see a decrease in immigration levels (Source: Gallup, 2016, Based on the latest data available on 26 March 2019). This apprehension is significantly high in some European countries like Greece (81.9%), Malta (75.5%), Italy (66.5%), Czechia (59.4%), Hungary

[29] Available at https://www.migrationdataportal.org/international-data?i=co_immig_yr&t=2013, Accessed on July 1, 2022

(57%), and Spain (55.5%). Also, this indicator is high in countries like Cyprus (88.2%), the UK (68.8%), and the USA (36.3%). A high level of apprehension toward migrants in countries like Greece and Italy is intuitive because inflows from the Middle East to Europe mostly happen through Greece. Similarly, inflows from North Africa to Europe occur through Italy.

Contrary to these negative opinions, Figure 3.2 reports the acceptance of migrants, i.e., the percentage of adult respondents who think that the city or area where they live is a good place to live for immigrants from other countries (0 = disagree; 100 = agree) (Source: Gallup, 2016, Based on the latest data available on 26 March 2019). We note that migrant acceptance level was very high in most advanced economies like Canada (85.5%), Spain (84.2%), Ireland (83.1%), the UK (82.4%), Netherlands (81.9%), Germany (79.7%), Denmark (78.2%), Belgium (77.1%), the USA (76.4%). At first glance, Figures 3.1 and 3.2 contradict each other. Thus, it is worth noting that these surveys were not conducted simultaneously (Figure 3.1 in 2013 and Figure 3.2 in 2015). In retrospect, it seems that the unfortunate death of Alan Kurdi, the two-year-old Syrian kid, in September 2015 might have affected the conscience of respondents in Figure 3.2. Literature suggests that this unfortunate event caused a dramatic upturn in the refugee narratives. This can be a plausible explanation for the contrast between Figures 3.1 and 3.2.

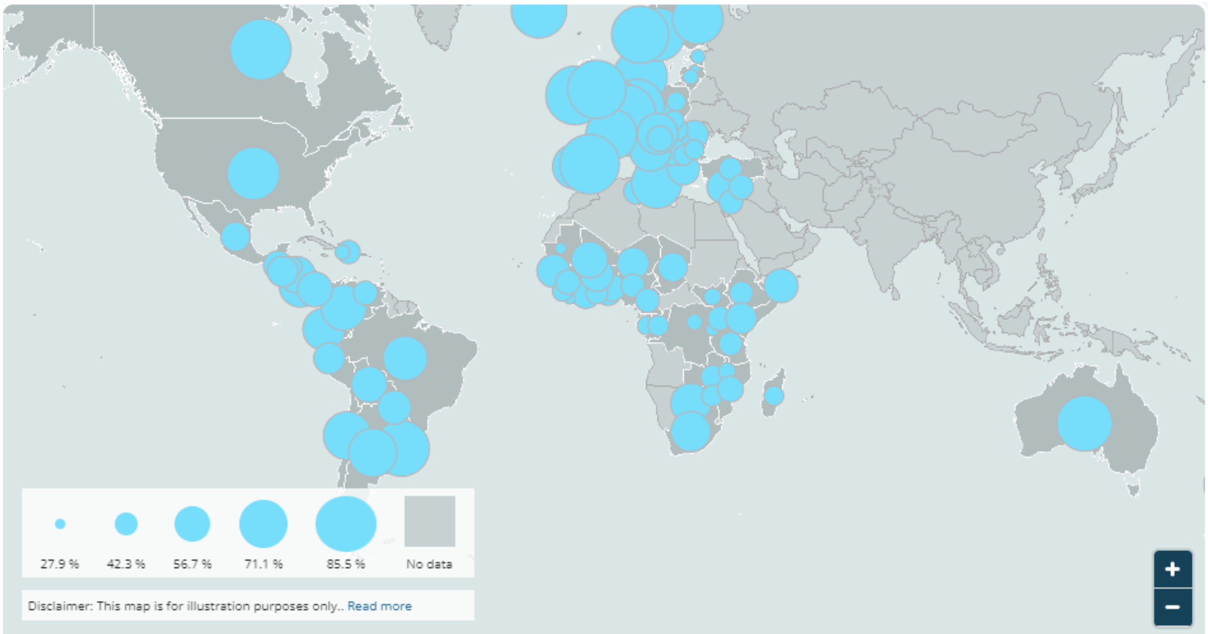


Figure 3.2 Migrant Acceptance[30]

[30] Available at https://www.migrationdataportal.org/international-data?i=co_diversity_yr&t=2015, Accessed on July 1, 2022

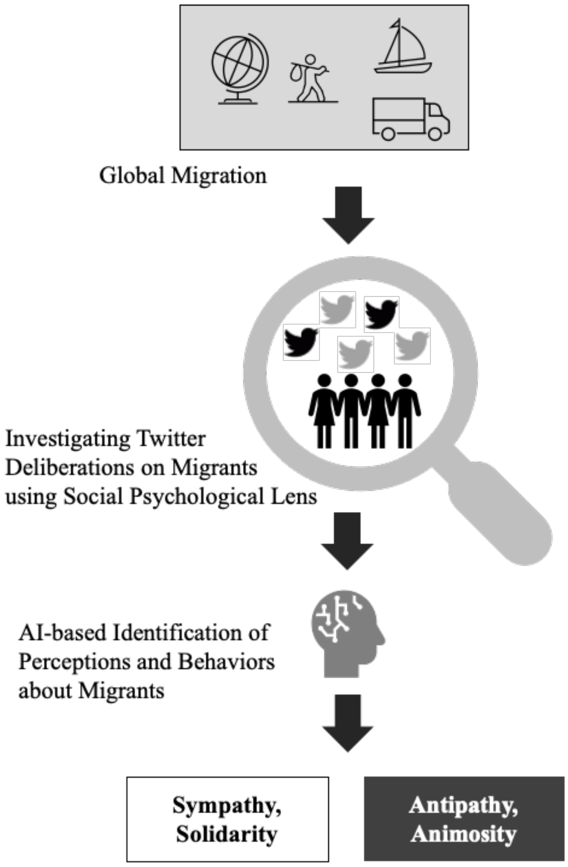


Figure 3.3 Framework to Identify Perceptions and Behaviors

To probe societal opinions about migrants, we draw insights from the social psychology literature that argues *perception* mostly leads to *behavior* (Dijksterhuis & Bargh, 2001). According to this theory, apprehensiveness or antipathy toward migrants may lead to animosity or xenophobic behaviors. Similarly, a sympathetic view toward migrants may lead to solidarity. The scope of our dissertation did not allow us to investigate this causality. Hence, Chapter 3 attempts to understand the diverse and diametrically opposite migrant-related societal perceptions and behaviors on the Twitter platform. Based on the social psychology literature, we argue that concerns regarding the sad state of affairs in asylums or the discrimination faced by migrants indicate a sympathetic *perception* of the user, and getting involved in fundraising or support activities is a solidarity *behavior*. Similarly, assuming or believing that migrants are often involved in illegal activities is a negative *perception,* and demanding their deportation is negative *behavior*. We have considered 0.8 million migration-related tweets (after pre-processing) from May 2020 to Sept 2020 to probe perceptions and behaviors toward migrants.

As mentioned in the introductory chapter, opinion mining using social media data, especially in the context of migration, is a challenging task. To the best of our knowledge, none of the prior studies have analyzed the granular differences between perceptions and behavior toward migrants on social media platforms. Thus, in the domain of applied NLP, this chapter attempts to address this gap. We refer to interdisciplinary literature to conceptualize and identify perceptions and behaviors towards migrants on the Twitter platform. Figure 3.3 reports our overall research framework.

As noted in Chapter 2, a series of prior studies probed the apprehensiveness toward migrants (often expressed through swear words or offensive language). Literature has conceptualized this as a hate speech detection task (Davidson et al., 2017; Waseem & Hovy, 2016). However, the literature mostly ignored the delicate nuances between perceptual and behavioral aspects of hate speech. We argue that both types of tweets can be anti-migrant in their orientation, but they are not the same. Thus, we have reconceptualized the binary hate speech detection task into a fine-grained task of detecting the perceptual and behavioral aspects of hate speech. This is another contribution of our study in the domain of applied NLP. Interestingly, we also note that a tweet that is supportive of migrants can also use 'swear words' against discrimination.

On the methodology front, this chapter has employed unsupervised and supervised models for analyzing the corpus. We have employed three unsupervised, i.e., ZSLMs. Next, we consider CNN and Bi-LSTM models with fastText embedding. Finally, we employ transformer-based models: BERT and RoBERTa. Our proposed BERT + CNN architecture has outperformed other models and reported an F1-weighted score of 0.76 for this complex perception-behavior identification task.

3.2 Migration on Twitter: What Do We Know?

Migration has attracted the attention of researchers from multiple disciplines that range from sociology (Crawley & Skleparis, 2018) to communication (Sajir & Aouragh, 2019), and psychology (Goodman et al., 2017; Volkan, 2018) to information science (Aswad & Menezes, 2018; Urchs et al., 2019; Vázquez & Pérez, 2019). Migration-related issues were probed in the context of France (Siapera et al., 2018), Germany (Riyadi & Widhiasti, 2020; Siapera et al., 2018), Italy (Capozzi et al., 2020; Kim et al., 2020b), Korea (Kim et al., 2020a), Netherlands (Udwan et al., 2020), Spain (Calderón et al., 2020a, 2020b; Vázquez & Pérez, 2019), Syria (Dekker et al., 2018; Rettberg & Gajjala, 2016; Reel et al., 2018; Öztürk & Ayvaz, 2018; Udwan et al., 2020), Turkey (Bozdag & Smets, 2017; Özerim & Tolay, 2020), the UK (Coletto et al., 2016), and the USA (Zagheni et al., 2018). Information science researchers mostly analyzed online contents, such as Facebook (Capozzi et al., 2020; Hrdina, 2016; Zagheni et al., 2018), Instagram (Guidry et al., 2018), Pinterest (Guidry et al., 2018), YouTube (Lee & Nerghes, 2018), Twitter (Alcántara-Plá & Ruiz-Sánchez, 2018; Aswad & Menezes, 2018; Calderón et al., 2020; Gualda & Rebollo, 2016; Kim et al., 2020a, 2020b; Nerghes & Lee, 2018; Pope & Griffith, 2016; Vázquez & Pérez, 2019) as well as mainstream media (Nerghes & Lee, 2019).

These studies have employed various NLP tools such as topic modeling (Calderón et al., 2020; Guidry et al., 2018), sentiment analysis (Nerghes & Lee, 2018; Öztürk & Ayvaz, 2018; Pope & Griffith, 2016), hashtag analysis (Özerim & Tolay, 2020; Kreis, 2017; Riyadi & Widhiasti, 2020), and network analysis (Himelboim et al., 2017; Nerghes & Lee, 2018; 2019).

The Twitter platform can be a real-time source of migration issues (Aswad & Menezes, 2018). Hence, Twitter was widely employed by prior studies – as we noted in Chapter 2. Accordingly, extant literature probed Twitter data to analyze migration movement (Mazzoli et al., 2020; Urchs et al., 2019; Zagheni et al., 2014). For example, Urchs et al. (2019) investigated the movement of migrants during 2015 in three European countries. They have identified 583 tweets with details regarding the number of migrants moving from one country to another. Geo-tagged tweets were also used for analyzing migration movements (Mazzoli et al., 2020; Zagheni et al., 2014). Kim et al. (2020b) analyzed location information to identify immigrants and emigrants. Mazzoli et al. (2020) demonstrated that Twitter-based prediction of migration flow is consistent with official statistics. Similarly, Coletto et al. (2016) considered spatial, temporal, and sentiment dimensions of their corpus and argued that Twitter provides real-time spatial information.

As we noted in Chapter 2, Twitter data was also used for sentiment analysis and opinion mining in the context of migration (Lee & Nerghes, 2018; Reel et al., 2018). A multilingual study (i.e., German and English) considered two specific refugee-related events and performed a sentiment analysis of Twitter discussions around these two events (Pope and Griffith, 2016). Similarly, Siapera et al. (2018) have analyzed various hashtags to study the network evolution as a response to three refugee-related specific events. This study argues that an event can have two predominant framings. First, a humanitarian frame where the discussion revolves around how an organization can help refugees. Some of the prominent hashtags of this first frame were *#safepassage*, *#humanrights*, *#refugeesupport*. Second, a far-right perspective where refugees are framed as terrorists or criminals, and subsequently, these create security and safety concerns in the host nation. These apprehensions toward migrants were also observed by other studies – especially in the context of Syrian refugees (Özerim & Tolay, 2020; Öztürk & Ayvaz, 2018; Reel et al., 2018). Reel et al. (2018) have proposed a random forest-based classifier to extract and identify tweets about Syrian refugees. Later Özerim & Tolay (2020) explored Turkish tweets, especially against Syrian Refugees, and this study observed the presence of echo chambers on the microblogging platform. Similarly, Kreis (2017) also analyzed negative perceptions about Syrian refugees through a hashtag-based analysis. These studies found a strong apathy toward refugees and identified nationalist hashtags such as *#EuropeforEuropeans*. However, Coletto et al. (2016) have found that positive and negative sentiments are not uniform in European Unions and emphasized the opinion dynamics. Similarly, our pilot study of Chapter 1 also probed Twitter deliberations in the European context. Chapter 1 has identified five themes, namely economic conditions, employment opportunities, healthcare support for migrants, discrimination against migrants, and safety concerns of the host nations.

Intuitively, nationalist ideologies may lead to abusive behaviors toward migrants. Davidson et al. (2019) found substantial racial biases in multiple hate speech and abusive language detection datasets. Detecting abusive language on social media platforms is a challenging task (Davidson et al., 2017; Waseem & Hovy, 2016). A series of prior studies probed hate speech in the context of migration. As noted in Chapter 2, the SemEval 2019 task tried to detect hate speech against immigrants (and women) on the Twitter platform (Basile et al., 2019). This task had two components: to detect the target of hate speech (generic or individual) and the presence (or absence) of aggressiveness. Similarly, in the context of immigrants, Sanguinetti et al. (2018) have also prepared an Italian tweet corpus with binary labels for hate speech, stereotyping, and irony (i.e., yes or no); multiple classes for aggressiveness and offensiveness (i.e., no, weak, and strong), and a five-point scale for intensity analysis. Hrdina (2016) also analyzed publicly visible pages and profiles on the Facebook platform. This study has found that hate speech against migrants was aggravated by disparate Facebook users, extremist groups' propaganda, and news media. This study has observed that frequent hate speech producers are primarily middle-aged and middle-class males and noted a significant under-representation of elderly and young Facebook users. Calderón et al. (2020) also considered 1469 tweets to analyze the reasons behind the perception of rejecting migrants and refugees. This study argued that apathy toward foreigners is mainly driven by the economic burden of the host nations, security threat, invasion threat, identity threat, social prejudice, and explicit rejection. They also find that this rejection of foreigners is often due to multiple reasons from the above list.

To sum up, this brief review reconfirms the findings of Chapter 2, i.e., the extant literature probed Twitter data for understanding latent opinions. Sentiment analysis, especially negative sentiment, reveals a xenophobic discourse on the social media platform. A handful of studies also employed an opinion-mining approach to understanding the genesis of this xenophobic discourse. We did not come across an article that holistically investigated the diverse range of perceptions and behaviors toward migrants. Thus, this chapter has attempted to address this gap, and we refer to social psychology literature to analyze the perceptions and behaviors toward migrants.

3.3 Perceptions and Behaviors toward Migrants

Social psychology literature argues that "we perceive because we want to know what is going on around us … perception is essential for us to comprehend our environment, but that does not mean that this understanding is an end in itself. Rather, understanding is a means by which we act effectively" (Dijksterhuis & Bargh 2001, p. 3). This literature also assumes a "shared representational systems for perception and action" because "people have a natural tendency to imitate" (Dijksterhuis & Bargh 2001, p. 8). In other words, our perceptions and behaviors converge at the societal level due to our tendency to mimic others. Hence, societal perception is the cumulative outcome of individual perceptions. For example, the following anti-immigrant tweet from a specific user might be representing her personal view about immigrants.

– *We need to get rid of the Human Rights Act and political correctness. Immediately, we need to deport all illegal immigrants – the potential terrorists, plus those migrants convicted in criminal cases.*

Probably, like-minded social media users may propagate the above view by retweeting, and xenophobic behavior would gain momentum because we tend to mimic.

This perception-behavior theory has identified three trigger points of social perceptions. The first trigger point is *observables* – it is easy to understand because "it involves behavior that we can literally perceive" (Dijksterhuis & Bargh, 2001, p. 9). Next, we develop *trait inferences* based on the behaviors of others. Interestingly, we generate trait inferences "without being aware of it" (Dijksterhuis & Bargh,2001, p. 9). Lastly, "social perceivers also go beyond the information actually present in the current environment through the activation of *social stereotypes* (emphasis added) based on easily detectable identifying features of social groups" (Dijksterhuis & Bargh, 2001, p. 9). Cumulatively, we may perceive more than reality. For example, the following tweet captures a negative perception of migrants.

– *There is a high probability that a migrant has already committed a crime*

Intuitively, this perception was triggered either by *social stereotyping* or *trait interference*. The word *probability* indicates that *observable* was not the trigger point for this perception. This *trait interference* or *social stereotyping*-based perception formation is crucial because without knowing the actual context, a specific segment of the society may develop an inappropriate perception of migrants. Our study investigates Twitter deliberations to unravel societal perceptions toward migrants.

Theoretically, social psychology literature suggests that our perceptions about migrants might influence our behavior toward them (Ferguson & Bargh, 2004). For instance, if we have sympathy toward migrants, then there is a high propensity that our behavior will express solidarity. On the contrary, antipathy may lead to animosity. However, it is worth noting that the scope of our research did not allow us to investigate this causality between perception and behavior. At the user level, chronologically tracking an individual's tweets and analyzing her perception and behavior for a sensitive issue like migration, even for academic purposes, can have ethical concerns. Testing the causality, even at the societal level, using Twitter data will be a challenging task. For instance, considering the evolution of Twitter deliberations over a longer time horizon might be an option, but this may broadly capture the composition of pro- versus anti-migrant users on the Twitter platform instead of the causality.

3.4 Methodology

3.4.1 Twitter Data

To explore perceptions and behaviors toward migrants, we have considered Twitter data and employed the Twitter search API (version Standard v1.1) for crawling. Twitter API-based search allows retrieval of up to 1% of all the tweets on the Twitter platform. Morstatter et al. (2013, p. 406) compared API-based data collection with Twitter's Firehose and found this API-based crawling "is a sufficient representation of activity on Twitter as a whole." Our initial crawling has considered keywords as follows: '*migrants*', '*refugee*', '*immigration*', and so on. We have regularly crawled data from May 2020 to September 2020. Prior migration-related studies observe that English tweets are predominant compared to other languages (Khatua and Nejdl, 2021a; Kim et al., 2020b). Accordingly, we also considered English tweets and crawled 1.2 million English tweets. We found that a significant portion of our initial corpus was biased toward popular tweets. Hence, we have removed tweets/retweets with similar content and duplicate tweet IDs. The corpus size became 0.8 million tweets after removing these popular tweets. Thus, 33%

of our initial corpus (i.e., 0.4 million tweets out of a total of 1.2 million tweets) was repetitive tweets with similar content.

3.4.2 Geographical focus of our data

The locational/country data is available for a small portion of these 0.8 million tweets. Based on this sub-sample, our corpus comprises tweets from 130 countries, but the distribution was skewed. Table 3.1 reports the geographical focus of our corpus. Around 75% of our corpus is from the USA, the UK, and Canada. One probable reason is - English is the most commonly used language in these countries.

Table 3.1 Geographical focus of our corpus

#	Country	Tweet	#	Country	Tweet
1	USA	53.6%	5	Australia	2.0%
2	UK	18.4%	6	Nigeria	1.9%
3	Canada	4.1%	7	Others	16.1%
4	India	3.9%		Total	100.0%

3.4.3 Identification of Themes and Aspects

Labeling a huge tweet corpus is a challenging task. Hence, Hedderich et al. (2020) suggest a distant and weak supervision approach for a new dataset. Here, a domain expert uses her tacit knowledge to design a set of rules using contextual (external) knowledge sources and heuristics (Ratner et al., 2017; Rijhwani et al., 2020). However, this semi-automatic supervision approach, which is essentially syntactical, can lower the performance of classifiers (Fang & Cohn, 2016). Hence, prior studies suggest combining distant supervision with noise-handling techniques (Hedderich et al., 2020). Following this stream of research, we initially employed distant supervision and manual annotation later. For designing our distant supervision rules, we have juxtaposed two threads of literature: perception-behavior literature and migration literature.

Perception is lexically defined as an idea or a belief you have based on how you see or understand something. Similarly, the lexical[31] definition of behavior is "the way that somebody behaves, especially toward other people." Subsequently, for distant supervision, we need to prepare an exclusive corpus of keywords related to our categories of perceptions and behaviors. To prepare this corpus, we have referred to multiple reports

[31] Available at https://www.oxfordlearnersdictionaries.com/definition/english/behaviour, Accessed on July 1, 2022

and scholarly articles on migration. We also went through various UNHCR policy documents to understand the context and identified the aspects[32] accordingly.

Based on our understanding of the interdisciplinary literature, our *Sympathy Perception* comprises tweets expressing concerns about inequality, discrimination, and injustice toward migrants. Some of these tweets deliberate about discrimination in terms of low wages or inadequate facilities in the asylums. Thus, to identify sympathy in our distant supervision approach, we have considered the following aspects: *vulnerable economic conditions*, *discrimination against migration*, *human rights violations*, *poor living conditions in asylums*, *lack of job opportunities*, and *inadequate access to health/education facilities*. On the contrary, *Antipathy Perception* considers tweets that assume migrants are getting preferential treatment compared to citizens of host nations or that most of them are involved in criminal or violent activities. Accordingly, for antipathy, we consider aspects: *migrants entering illegally*, an *economic burden in host nations* (because they might destroy job opportunities for citizens of the host nations), and *safety concerns* (because migrants can be violent out of desperation).

Our *Solidarity Behavior* tries to capture various support activities to rehabilitate the migrants. It ranges from fundraising activities to awareness campaigns. Thus, if social media users organize a donation drive and share the same on the Twitter platform, we consider it a solidarity behavior toward migrants. Aspects such as *support migrants/immigrants*, *donate for migrants*, *safety of migrant women*, and *help refugee entrepreneurs* were considered under the solidarity category. On the contrary, the *Animosity Behavior* captures the dislike and hatred toward migrants. Hence, for animosity, we have considered the following aspects: *migrants not in our country*, *no refugees*, *go back*, *deport migrants*, and *take back control*.

Some migrant-related tweets do not belong to the above four categories. These tweets are as follows: tweets that refer to migration superficially (where migration is not the dominant theme or core issue, but it might have an opinion about migrant issues) and tweets by migrants or refugees where they share their personal experiences (e.g., Chapter 6). We label them as *Generic* tweets. The intersection between each category corpus and the tweet was computed to label a tweet in this weakly-supervised approach. We have considered this labeling based on the distant supervision approach as our silver standard (Ménard & Mougeot, 2019). Subsequently, human annotators use their contextual understanding and domain knowledge to prepare the final gold standard by tackling the

[32] In this dissertation, we have employed the terms 'class,' 'category,' 'theme', 'aspect' and 'label' interchangeably.

noisy data from the silver standard (Ménard & Mougeot, 2019). The Cohen's Kappa coefficient was 82.8% for our inter-rater reliability.

3.4.4 Complexity of our Task

A fine-grained analysis of our corpus has elucidated the challenges associated with identifying perceptions and behaviors on the Twitter platform. We did not find much overlap between pro-migrant and anti-migrant categories, but we do observe overlaps within them – especially for anti-migrant tweets (i.e., antipathy and animosity categories). It is worth noting that social media platforms allow a user to express her views/voices. Hence, the puzzling question is - whether a voice is a perception or behavior. For example, let us consider two tweets as follows:

– *I find migrants to be a frightful lot, so different from us.*

– *Hello, migrants! Go back to your own country.*

Both the above tweets are anti-migrant tweets, but the first tweet is less provocative. Our annotation process has considered that the first tweet is “a belief or opinion” and labeled it as an antipathy toward migrants. We felt that the second tweet is closer to “the way that somebody behaves, especially toward other people” – hence, we considered it an animosity behavior.

However, the counterargument can be that migrants should go back to their own country – this can be ‘a belief or opinion’, and a significant portion of the society might have the same opinion. Accordingly, the second tweet can also be considered as a perception toward migrants. In other words, perception and behaviors are not dichotomous in a stringent sense, and we acknowledge this fluidity. In the context of anti-migrant tweets, perceptions are tweets that are less opinionated or less provocative.

We note that some tweets convey concerns from more than one of our four categories. Theoretically, a tweet can be a combination of any two or three (or even four) perceptions and behaviors. For instance, a tweet can be simultaneously pro- and anti-migrant as follows:

– *We must improve the living conditions of legal and needy migrants in our government asylums, but illegal migrants don’t you f***king dare to enter my country… you are criminal because you are entering illegally.*

To tackle these types of tweets, we need to frame our problem as a multi-label classification problem. However, our corpus doesn’t have enough data points like the above tweet to train our neural network models. Hence, we ignored these tweets in our analysis, but future studies can probe these tweets. Additionally, some of the joint categories (e.g., sympathy + animosity) are rare.

Table 3.2 Representative tweets from our corpus

Class	Sample Tweets
Sympathy *(SYM)*	- Let us end stigma and discrimination against migrant workers and their children - An asylum seeker is not an illegal immigrant ... You f***ing idiot go buy a dictionary
Antipathy *(ANT)*	- It doesn't matter even though we were born here and pay for the healthcare. Just be a migrant and suddenly, it is a human rights violation - We must stop the immigrants coming to our country they are crossing our borders in increasing numbers and putting the strain on our facilities
Solidarity *(SOL)*	- We believe that everyone deserves a fair chance to become an #entrepreneur. Therefore, we support #migrant entrepreneurs! - Support our campaign today to help those in <location> facing all the ongoing humanitarian crises including the forgotten <location> refugees
Animosity *(ANM)*	- F**king illegal immigrants are not welcome in <location> F**k off you pr**k. - You don't need no f**king answers. You are an immigrant and part of the problem. Just go back!
Generic *(GEN)*	- I am an immigrant and a citizen ... I have paid taxes for 25 years and I care about this country - Our data shows most <members of a political party> agree both that discrimination against whites has become as much of a problem as discrimination against immigrants

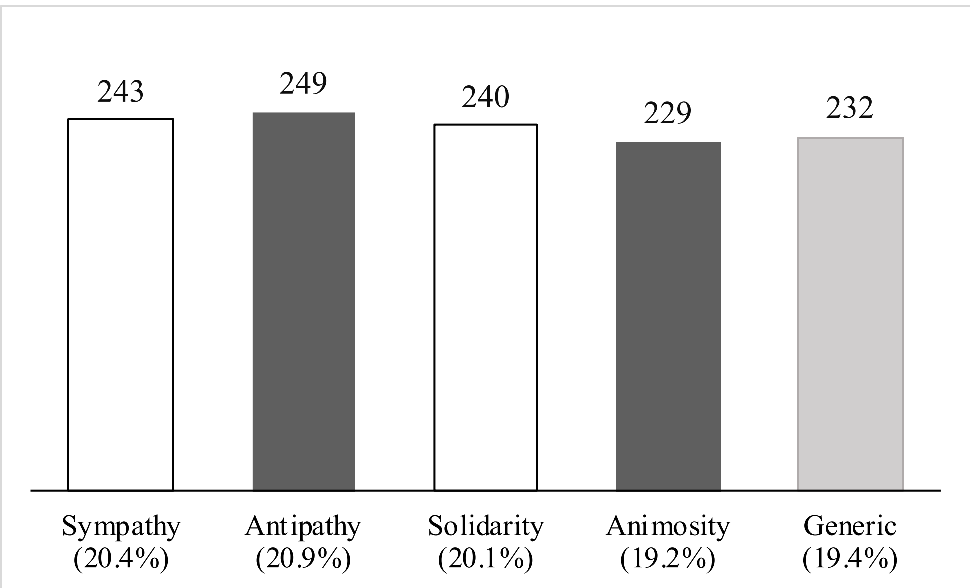


Figure 3.4 Distribution of Annotated Tweets

3.4.5 Annotated Data

Figure 3.4 reports the distribution of 1193 annotated tweets, and Table 3.2 provides a few sample tweets from perception (i.e., sympathy & antipathy), behavior (i.e., solidarity & animosity), and generic categories. Considering the sensitivity of these tweets, we paraphrased all quoted tweets in Chapter 3 to maintain user anonymity. We randomly split our 1193 annotated tweets into 85%, as training dataset, and 15%, as test dataset, for our subsequent analysis. Additionally, we also employed 5-folds cross-validation for our analysis.

3.4.6 Comparison with prior Hate Speech Corpora

Anti-migrant tweets, which capture antipathy and animosity toward migrants, mostly use offensive language or swear words. Thus, these tweets can be labeled as hate speech, but they are not the same. Hence, these two classes deserve comparison with prior studies on hate speech. A few prior studies considered voluminous annotated data, but some datasets were also smaller in size. For example, Ross et al. (2017) annotated 541 German tweets with key hashtags on the refugee crisis that could be offensive – this is comparable to our study. In comparison to prior studies, we find that our dataset is significantly complicated and balanced. For example, the dataset prepared by Davidson et al. (2017) contains 24,802 English tweets in English. However, only 5.77% of tweets were hate speech, 77.43% were offensive, and 16.80% were neither in these two categories. Similarly, Waseem & Hovy (2016) have considered 16,914 English tweets. They have annotated this corpus into three classes: 12% of tweets on racism, 20% of tweets on sexism, and 68% of tweets do not belong to either of these two classes. Madukwe et al. (2020) pointed out that these datasets were not balanced, and it can inappropriately improve the classification accuracy. Unlike these prior studies, our dataset is balanced (refer to Figure 3.4).

Nowadays, Twitter allows its users to post 280 characters compared to the previous restriction of 140 characters. Hence, we find that the average length of annotated tweets of Davidson et al. (2017) is significantly shorter than our annotated tweets. For example, the average word count of their corpus is 14 without pre-processing, whereas the average word count of our corpus is 30 after pre-processing. Some of the annotated tweets from the corpus of Davidson et al. (2017) are as follows:

- *Bad b**ches is the only thing that I like.*
- *Foreign chick, no lie … Man, that b**ch beautiful.*

Intuitively, a less complex syntactic approach can correctly classify these shorter texts by considering context-specific swear words. However, longer tweets are more complex. For instance, a tweet from our corpus says:

– *We are bringing in thousands of migrants every year and they call us racist. No matter what we do or how much we give them these as*h***s will always view themselves as the oppressed.*

The above tweet indicates that the social media user has developed a negative perception and is using offensive words toward migrants (probably) based on his experience. Another tweet from our corpus says:

– *If white Americans say, 'Take America back!' or tell an immigrant to 'Go back to your home country!', I am going to chuck a f**king history textbook in their face. White people originally came from Europe! Your f***ing ancestors were illegal immigrants!*

As we pointed out earlier, this tweet uses offensive words and argues that ancestors of present American citizens had moved from Europe to America. Hence, according to the above tweet, the legal citizens of the USA are historically immigrants. Therefore, the above tweet is sympathetic to today's immigrants, but a syntactic approach (by considering offensive words) will not be able to decipher it appropriately.

3.5 Analysis

3.5.1 Zero-Shot Learning

Building a rich training corpus is a time-consuming and resource-intensive task. Unsupervised models, such as ZSLMs, do not need this training corpus. Since ZSLMs can perform the task without the training corpus, these models are emerging as an alternate option in combination with large pre-trained models like BART (Lewis et al., 2020) and XLM-RoBERTa (Conneau et al., 2020). Hence, we consider unsupervised ZSLMs to predict unseen classes in the context of migration using the natural language inference (NLI) method. Yin et al. (2019) argue that pre-trained NLI models can perform the classification task without training. This approach trains a model to interpret the relationship (i.e., entailment, contradiction, or neutral) between two text streams. Next, it returns the probabilities of different classes according to their text content. However, the performances of ZSLMs are lower than those of supervised models. For instance, Nie et al. (2020) cautioned that non-expert annotators could successfully find the weakness of unsupervised models.

We have considered three pre-trained ZSLMs as follows: BART-Large-MNLI (Lewis et al., 2020)[33], XLM-RoBERTa-Large-XNLI (Conneau et al., 2020)[34], and XLM-RoBERTa-Large-XNLI-ANLI (Nie et al., 2020)[35]. BART-Large-MNLI considers a conventional seq2seq/machine translation architecture with a bidirectional encoder and a left-to-right decoder (Lewis et al., 2020). Using the MNLI dataset, this pre-trained model has shuffled the order of the original texts and employed an in-filling approach where a single mask token has replaced the spans of texts.

Table 3.3 Performance of Unsupervised Models

	PR	RC	F1	AUC	Tag
BART-Large-MNLI	0.35	0.34	0.27	0.58	Single
	0.32	0.25	0.25	0.53	Double
	0.43	0.41	**0.38**	0.63	Multi
XLM-RoBERTa-Large-XNLI	0.28	0.32	0.28	0.58	Single
	0.25	0.28	0.25	0.55	Double
	0.31	0.33	**0.31**	0.58	Multi
XLM-RoBERTa-Large-XNLI-ANLI	0.31	0.33	0.27	0.58	Single
	0.32	0.30	0.30	0.56	Double
	0.36	0.35	**0.31**	0.59	Multi

The training of XLM-RoBERTa-Large-XNLI has considered larger datasets, a more extensive vocabulary, and longer sequences with larger batches (Conneau et al., 2020). This model has considered 2.5 TB of newly created clean CommonCrawl data. This model is a combination of XLM and RoBERTa architecture. The approach makes full use of the entire content of the sentence to extract relevant semantic features. This model is the multilingual variant of RoBERTa, and it has considered multilingual MLM for training. However, it also performs well for monolingual language tasks. Finally, XLM-RoBERTa-Large-XNLI-ANLI, took XLM-RoBERTa-Large as a base model and fine-tuned it by combining NLI data with XNLI and ANLI across multiple languages (Nie et al., 2020). Recently, Nie et al. (2020) considered a new large-scale NLI benchmark dataset that was collected through an iterative, adversarial human-and-model-in-the-loop procedure.

[33] https://huggingface.co/facebook/bart-large-mnli, Accessed on September 22, 2023
[34] https://huggingface.co/joeddav/xlm-roberta-large-xnli, Accessed on September 22, 2023
[35] https://huggingface.co/vicgalle/xlm-roberta-large-xnli-anli, Accessed on September 22, 2023

Table 3.4 Details of our tagging approach for ZSLMs

Single tag (Column 1)	**Double tags** (Column 2)	**Multi-tags** (Column 3)
Sympathy	Sympathy + Humanitarian	Empathy, Inequality
Antipathy	Antipathy + Xenophobic	Hatred, Disgust, Illegal
Solidarity	Solidarity + Consensus	Unity, Support
Animosity	Animosity + Bitterness	Deport, Hostility
Generic	Generic + Experiential	Impartial, Nondiscriminatory

Results: Table 3.3 reports the performance of transformer-based unsupervised models. We find that the weighted F1 score of unsupervised models for the single tag is significantly low, i.e., less than 0.30. Extant literature says that the NLI approach investigates the semantic similarity for predicting unseen classes. Hence, the tag word(s) specific to a class can play a crucial role in the correct prediction. Pushp & Srivastava (2017) argued that multiple tag words could improve the accuracies of ZSLMs. Hence, we also followed a similar approach (refer to Table 3.4 for details). Our single tag approach considers only one keyword for each class (i.e., the word in Column 1 of Table 3.4). Double tags consider both the words of Column 2.

Interestingly, double tags have lowered the model performance, but multiple tags (i.e., Column 2 keywords + Column 3 keywords) have slightly improved model performance. Probably, double tags have created confusion (for closely resembling conceptual classes), whereas multiple tags have enhanced the interpretation of ZSLMs. The performances of these models are not impressive for our complex classification task, and these models failed to tease out the differences between broadly similar classes, such as antipathy and animosity. Additionally, ZSLMs wrongly classified some of the Generic class tweets into other classes. Our low AUC scores also confirm the same. We have found that BART-Large-MNLI has reported the best F1-weighted score of 0.38. In Table 3.3, a few F1 scores are lower than PR or RC values because we considered a weighted F1-score.

3.5.2 Neural Models with Embedding

Extant literature found that DL-based classification models are superior to traditional bag-of-words models or n-gram models (Conneau et al., 2017; Kalchbrenner et al., 2014;

Young et al., 2018). For instance, the CNN model embeds words into low-dimensional vectors (Kim 2014). Next, convolutional filters slide over the word embedding matrix. These filters play a crucial role in task-specific performance. Finally, the max-pooling function provides a fixed dimension output for the desired classification task. In addition to CNN models, RNNs were also used by prior studies for the classification task. However, RNNs cannot capture long-term dependencies of very long sequences. However, Bi-LSTM, which is a variation of RNN models, addresses this concern. Pre-trained embeddings improve the performance of these models. Hence, we have considered CNN and Bi-LSTM with pre-trained embeddings from fastText - wiki-news300d-1M, built using a web-based corpus and statmt.org news dataset (Joulin et al., 2016).

Results: We have considered different hyperparameters, such as multiple BSs (BS: 16 and 32) and DRs (DR: 0.4 and 0.5), for our CNN and Bi-LSTM models. The hidden layer for all these models was 128. We have considered SoftMax activation in our final classification layer to predict the final class. We have used 'adam' as our optimizer for the modeling. Table 3.5 reports the model performances. Performances of CNN + fastText models are better than Bi-LSTM + fastText models, and the highest F1-weighted Score for CNN is 0.72 (BS - 32, DR - 0.4).

Table 3.5 Performance of Deep Learning Models

	PR	**RC**	**F1**	**AUC**	**DR**	**BS**
CNN + fastText	0.71	0.69	0.70	0.81	0.4	16
	0.75	0.68	0.70	0.82	0.5	16
	0.75	0.71	**0.72**	0.83	0.4	32
	0.71	0.69	0.70	0.81	0.5	32
Bi-LSTM + fastText	0.69	0.59	0.59	0.75	0.4	16
	0.62	0.60	0.59	0.75	0.5	16
	0.62	0.56	0.56	0.74	0.4	32
	0.69	0.60	**0.61**	0.76	0.5	32

3.5.3 Transformer-based Neural Models

Next, we have considered two transformer-based models: BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019) for the classification task. Transformer-based models work reasonably well for text classification tasks because transformers are pre-trained on a diverse and large corpus. Core aspects of these models are their multi-head self-attention mechanism to extract the input tokens' semantic aspects for contextual representation with multiple layers. Unlike RNNs, these models can handle long-term dependency problems. BERT has successfully performed numerous NLP-related tasks – including the classification task. BERT is a bidirectional unsupervised pre-trained model. Devlin et al. (2019) have considered BooksCorpus and English Wikipedia (16GB) for training purposes. BERT was introduced in 2018. However, within a year, BERT's performance was further improved by adding more training corpus and incorporating minor adaptations to the training process (Liu et al., 2019). This advanced version of BERT is known as RoBERTa. In addition to the pre-training corpus of BERT, RoBERTa also used an additional corpus from CC-News (76 GB), Open Web Text (38 GB), and Storie's dataset (31 GB) for training.

Prior embedding approaches, such as word2vec (Mikolov et al., 2013) or GloVe (Pennington et al., 2014), have considered a single-word embedding representation for each word without considering the context of that specific word. Therefore, these language representations failed to capture the context. In contrast, BERT considers the context of a particular word from both directions - both from the left and right direction. As we noted earlier, BERT and RoBERTa are pre-trained on a diverse and large corpus. This allows these models to effectively understand most of the words used in online content compared to word2vec or GloVe. To sum up, BERT's fundamental principle is to employ bidirectional transformers for the feature extraction layer to extract the contextual meaning of the words. Following prior studies, we have fine-tuned our transformer-based models. BERT requires input data to be in a specific format. Thus, the [CLS] special token was used to indicate the beginning, and for the separation or the end of the sentence, the [SEP] was used. The next step was to tokenize the text corpus and extract tokens that match BERT's vocabulary. For this task, we have used the HuggingFace Python library (Wolf et al., 2019). This library includes pre-trained models and allows fine-tuning for the classification task. We have used '*BertForSequenceClassification*' for our classification task. We have considered the *BERT-Base-Uncased* model comprised of 12-layers and 12-heads with a total of 110M parameters. We have considered max_seq_length of 256.

As we mentioned earlier, the convolutional filters of CNN models play a crucial role in classification tasks. Thus, in combination with BERT, CNN can outperform $BERT_{Base}$ or BERT + LSTM (Dong et al., 2020; He et al., 2019; Mozafari et al., 2020). Accordingly, we consider the outputs from all individual layers of BERT architecture, and the outputs of each layer of the transformer are concatenated for the final result. We perform the convolutional operation with a window size – 3, hidden size of BERT – 768, and apply the max pooling on the convolution output from each transformer layer. Lastly, we concatenate these values, which is the input of the fully connected layer, before SoftMax performs the final classification task.

Additionally, we propose another BERT + CNN model that considers only the final layer of the BERT transformer for CNN-based classification. The DR is 0.2 for all supervised models. For robustness, we have considered the following combinations of BSs (BS: 16, 32) and LRs (LR: 1e - 5, 2e - 5, 3e - 5). Like our DL models, we have used 'adam' as our optimizer. Following prior studies, we have considered ten epochs for our *BERT-Base-Uncased* models. The open-source implementation, pre-trained weights, full hyperparameter values, and experimental details are in accordance with the HuggingFace transformer library (Wolf et al., 2019).

Table 3.6 Performance of Transformer-based Models

	PR	RC	F1	AUC	LR	BS
BERT	0.65	0.65	0.65	0.78	1e - 5	16
	0.64	0.64	0.64	0.78	2e - 5	16
	0.66	0.66	0.66	0.79	3e - 5	16
	0.67	0.67	**0.67**	0.79	1e - 5	32
	0.67	0.66	**0.67**	0.79	2e - 5	32
	0.66	0.66	0.66	0.79	3e - 5	32
RoBERTa	0.73	0.71	0.71	0.82	1e - 5	16
	0.74	0.73	**0.73**	0.83	2e - 5	16
	0.74	0.73	**0.73**	0.83	3e - 5	16
	0.70	0.68	0.68	0.80	1e - 5	32
	0.72	0.72	0.71	0.82	2e - 5	32
	0.74	0.74	**0.73**	0.83	3e - 5	32
Layer-wise BERT + CNN	0.73	0.70	**0.71**	0.81	1e - 5	16
	0.69	0.69	0.69	0.81	2e - 5	16
	0.71	0.71	**0.71**	0.82	3e - 5	16
	0.65	0.65	0.64	0.78	1e - 5	32
	0.69	0.67	0.68	0.80	2e - 5	32
	0.67	0.67	0.67	0.80	3e - 5	32
Final Layer of BERT + CNN (Proposed)	0.74	0.74	0.74	0.84	1e - 5	16
	0.79	0.76	**0.76**	0.85	2e - 5	16
	0.77	0.74	0.75	0.84	3e - 5	16
	0.72	0.70	0.71	0.82	1e - 5	32
	0.73	0.73	0.72	0.83	2e - 5	32
	0.75	0.73	0.72	0.83	3e - 5	32

Results: Table 3.6 reports the performances of transformer-based models. Our proposed BERT (final layer) + CNN architecture has outperformed other models. F1-weighted scores for some of the top-performing models are as follows: 0.74 (95% confidence interval 0.67 – 0.81), 0.75 (95% confidence interval 0.68 – 0.82), and 0.76 (95% confidence interval 0.70 – 0.82). These are significantly high performances considering the complexity of our task due to closely resembling classes.

To test the efficiency of our proposed approach, we have also considered the publicly available dataset of Davidson et al. (2017). Interestingly, the F1-Score of our proposed BERT + CNN architecture is around 0.91 for the Davidson et al. (2017) dataset – as noted, this dataset is not balanced and comprises shorter texts with offensive words. This significant gap between 0.76, for our corpus, and 0.91, for Davidson et al. (2017) corpus, strongly indicates the complexity of our classification task.

Table 3.7 Analysis of Wrongly Classified Tweets

#	Tweets from Test Dataset	ORI	PRE
1	Calling migrants as criminals when statistically ***illegal immigrants commit*** less ***violent crime*** than native born Americans is unfair!	SYM	ANT
2	I didn’t get ***health benefits*** because I am a veteran and not an illegal immigrant. F*** you and f*** <location>.	ANM	GEN
3	F*** everyone in that parliament for turning a blind eye to illegal immigration, ***corruption, inequality, racism, genocide and trafficking***.	ANT	GEN
4	Hey, spineless <political party>, where is the social justice for the victims of illegal immigrant's crime. The same criminals you hide in your <location>	ANM	ANT
5	Let's do our part to prevent the gentrification <fast-food chains> from ***infiltrating*** <location> and other corridors. Support immigrant-owned businesses in <location> & <location>	SOL	ANT
6	True, ***I am an immigrant***, but this immigration numbers are outrages. Why do we need these many migrants? We don’t have enough jobs in <location>	ANT	GEN
7	Try to understand the science of climate change - an environmental refugee ***is also a refugee*** in a broader sense!	GEN	SYM

3.5.4 Error Analysis

Table 3.7 reports a few wrongly classified tweets. We have highlighted (in ***bold and italic font***) selected portions of these tweets to analyze why our model failed to classify these tweets correctly. For instance, syntactically, tweet #1 resembles an antipathy tweet, but semantically it is sympathetic. Our model has wrongly labeled tweet #5 due to the presence of an aspect such as ‘infiltrating’. Similarly, the phrase ‘I am an immigrant’ in tweet #6 misguided our model. Our model labeled tweets #2 and #3 as generic due to the presence of multiple issues, such as health, corruption, genocide, and so on, beyond migration. In tweet #4, annotators felt the word ‘criminal’ conveys animosity. Similarly, tweet #7 used the word ‘refugee’ in a different context. In brief, the analysis of these wrongly classified tweets reconfirms the complexity of our task.

3.6 Conclusion

Prior AI-based studies have probed social media data to examine migration-related issues. Our literature review reveals that this literature has primarily investigated a focused issue or a specific topic. To the best of our knowledge, none of the prior studies holistically analyzed the social media deliberations. To address this gap, we draw insights from social psychology literature to identify various implicit perceptions and behaviors toward migrants. Figure 3.3 graphically presents our overall research framework. Our perception-behavior conceptualization of Twitter data is a contribution to the applied domain of NLP literature. Also, this perception-behavior approach can potentially fine-tune future hate speech detection studies. Our proposed transformer-based supervised model, i.e., BERT + CNN architecture, has outperformed other models.

3.6.1 Potential Policy Implications

Interestingly, social psychology literature pointed out that perception mostly leads to action in animals, but there can be deviations for humans. Some perceptions may require an *additional facilitating mechanism,* and “sometimes the facilitator is present, sometimes it is absent; hence, sometimes perception leads to action whereas on other occasions it does not” (Dijksterhuis & Bargh, 2001, p. 5). The “default option is that perception does lead to action (as in fish or frogs), but under some circumstances a ‘stop-sign’ is given in order to block the impulse from resulting in overt behavior” (Dijksterhuis & Bargh,2001. P. 5). Figure 3.5 presents the same graphically.

We argue that the above two possible roads of flexibility, either by using additional facilitating mechanisms or stop-sign, allow regulators to influence societal behaviors. For example, after identifying sympathetic perceptions, regulators can promote these perceptions by incorporating an ‘*additional facilitating’* mechanism such as endorsing or appreciating those sympathetic tweets. This will reinforce solidarity activities at the societal level. On the contrary, regulators can identify the negative (and mostly inaccurate) perceptions and use a ‘stop sign’ to weaken the antipathy-animosity link. For instance, a common misperception is that – the inflow of migrants increases the labor supply. It lowers the wages and job opportunities – especially for the low-skilled employees of the host nations. However, in their recent book, Nobel laureates Banerjee & Duflo (2019), argued that migrants do not always lead to lower wages because these migrants ‘spend money: they go to restaurants, get haircuts, and go shopping. This creates jobs, and mostly jobs for other low-skilled people’. Thus, regulators can also use social media to debunk the myth and counter inappropriate negative perceptions. This will reduce the animosity toward migrants and subsequently reduce xenophobic behaviors at the societal level. Twitter has a stringent policy against discrimination. Hence, regulators can also collaborate with Twitter to identify these inappropriate perceptions and label these tweets as disputed claims. Our research offers an AI-based framework to identify these societal perceptions. However, the implementation of these interventions is not within the scope of our study. Researchers from the communication domain need to perform psychological experiments to design appropriate and effective intervention mechanisms. Overall, this chapter is an attempt to apply AI for social good, and hopefully, this is an incremental step toward an egalitarian society.

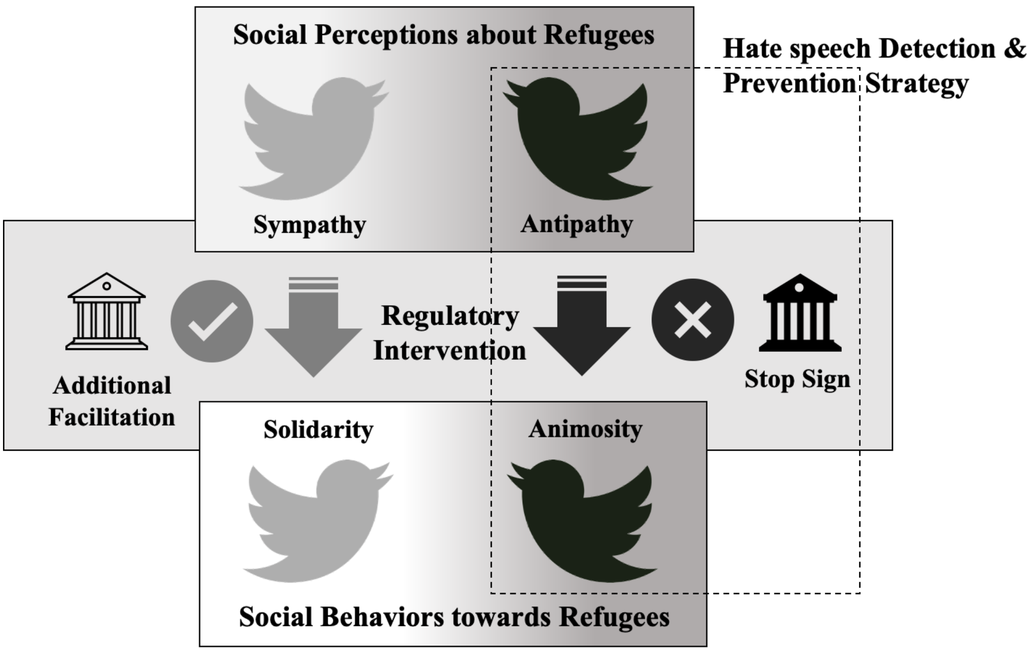


Figure 3.5 Takeaway for Policymakers

Chapter 4: Refugee Journeys from Displacement to Emplacement

4.1 The Uncertain Voyage

> *34,361 deaths recorded. Not all the deaths occur at sea, but also in detention blocks, asylum units ... more than 27,000 deaths by drowning since 1993, often hundreds at a time when large ships capsize ... Some entries have a name and a story, but the majority are anonymous data points – just over 1,000 are named ... Some 400 have taken their own lives; more than 600 have died violently at the hands of others.*
>
> *- The Guardian*[36]*, June 20, 2018*

More than 80 million human beings worldwide have experienced forced displacement by the mid-2020s, and 34 million of them are children below 18 years of age. 26.3 million are refugees, and 4.2 million are stateless people (United Nations Commission on Human Rights Report, 2021). The conventional myth is that refugee deaths occur only at sea, but the above excerpt from *The Guardian* debunks this myth. Refugees face adverse environments not only in the Mediterranean Sea but also after crossing the Mediterranean Sea – if they are lucky. The sufferings of refugees continue at detention blocks or asylums – these sufferings and subsequent deaths are social concerns. Thus, to explore their journey, we draw insights from the seminal work '*Les Rites de Passage*' (1909; The Rites of Passage) by Arnold van Gennep – the French ethnographer and folklorist (Van Gennep, 1960). Gennep systematically studied the passage of individuals from one social or religious status to another. He has identified three distinct stages of this journey: separation, transition, and reincorporation. A few studies, mainly from the social science domain, such as Castle and Diarra (2003), and Monsutti (2007), have employed this framework in the context of refugees or migrants. However, to the best of our knowledge, none of the prior AI-based studies in the domain of refugee and forced migration studies have employed this framework on social media data to explore the transition of refugees from their home country to the host nation.

A plethora of studies probed refugee-related issues using microblogging data, such as Twitter (Adler-Nissen et al., 2020; Kreis, 2017; Pope & Griffith, 2016). While the relevance of multimodal content for extracting relevant and actionable information was widely acknowledged (Alam et al., 2018; Gui et al., 2019), refugee-related studies mostly focused on syntactic text processing (Kreis, 2017; Pope & Griffith, 2016; Siapera et al., 2018) and did not probe the richness of multimodal data. A handful of studies also explored the

[36] Available at https://www.theguardian.com/world/2018/jun/20/the-list-europe-migrant-bodycount, Accessed on July 1, 2022

emotional responses to refugee-related visual content (Ibrahim, 2018; Olesen, 2018). Primarily, the extant literature probed refugee-related deliberations around a specific crisis or event (Pope & Griffith, 2016; Siapera et al., 2018). This does not offer a holistic view. BenEzer & Zetter (2015, p. 297) pointed out that "the refugee journey is the defining feature of the exilic process: it is a profoundly formative and transformative experience and a 'lens' on the newcomers' social condition." Thus, we employ Gennep's framework to study their journey.

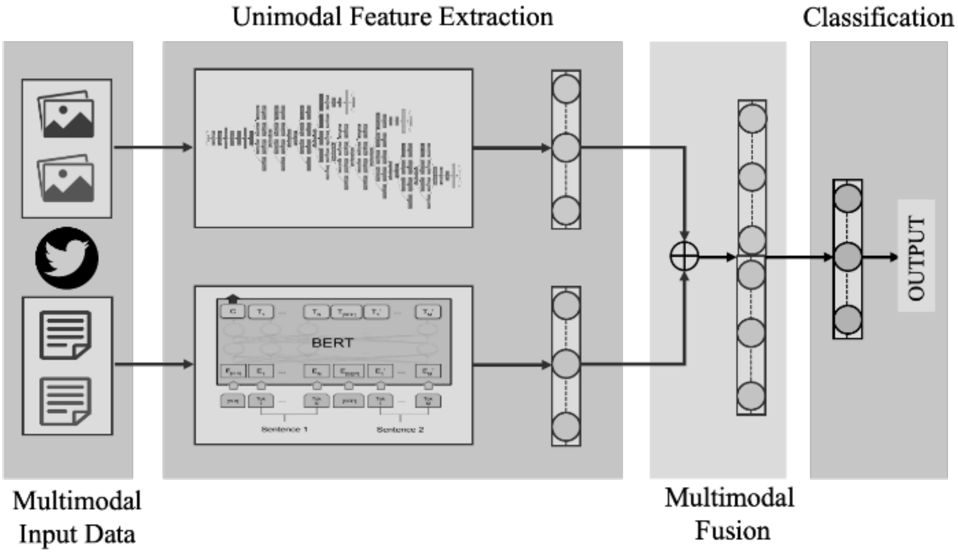


Figure 4.1 Proposed Multimodal Framework

We collected 0.23 million multimodal refugee-related tweets from April 2020 to March 2021 and manually annotated 1722 multimodal tweets for the analysis. On the methodology front, the refugee journeys can be conceptualized as a classification task where the common subject for our image inputs across all four phases are human beings, i.e., mostly refugees and border forces or NGO staff – starkly different and challenging from conventional image datasets, such as CIFAR-10 that comprises of distinctly different objects like flowers, animals, and buildings. For our classification task, we have initially considered unimodal models such as BERT (Devlin et al., 2019) + LSTM (for text inputs); and InceptionV3 (Szegedy et al., 2016), VGG19 (Simonyan & Zisserman, 2014), and ResNet (He et al., 2016) (for visual inputs). Subsequently, we have employed an early fusion of these unimodal models. Figure 4.1 presents our multimodal framework. Our multimodal models have outperformed unimodal models. The BERT+ LSTM + InceptionV4 (Szegedy et al., 2017) model has reported an accuracy of 80.93%. Subsequently, we collected 10,000 multimodal tweets related to the 2022 Ukrainian refugee crisis and tested the generalizability of our proposed framework. Our findings, i.e., an F1-score of 71.88 % for this 2022 dataset, strongly indicate that the phases of the refugee journey from displacement to emplacement were identical. Additionally, we have briefly analyzed the emotional aspects of our annotated corpus using LIWC (Tausczik & Pennebaker, 2010).

4.2 Refugee Issues on Social Media Platforms

Refugee journeys can have a wide range of consequences. The significant presence of refugees can impact the economic conditions of the host country (Alloush et al., 2017; Stark, 2004). Similarly, a challenging environment in the host country can have adverse consequences on the psychological well-being of refugees (Khawaja et al., 2017). Thus, scholars from multiple disciplines, such as anthropology (Cabot, 2016; 2019), economics (Alloush et al. 2017; Stark 2004), management (Karsu et al., 2019), psychology (Brown-Bowers et al., 2015; Khawaja et al., 2017; Maclachlan & McAuliffe, 1993), and sociology (FitzGerald & Arar, 2018; Karakayali, 2018), have probed refugee-related issues. The subsequent sections briefly review this literature.

4.2.1 Refugee Studies Using Text Data

As pointed out in Chapter 2, a handful of prior studies analyzed refugee-related text contents from social media platforms. These studies explored the opinions (i.e., perception and behavior) of social media users mainly from the host nation's perspective. For example, Siapera et al. (2018) investigated network formation on the Twitter platform in the context of three refugee-related events, namely the terrorist attack in Paris, France; sexual assault in Cologne, Germany; and the EU-Turkey deal for refugees. As mentioned, they identified two dominant perspectives - an apprehensive far-right perspective where refugees were considered as terrorists or rapists, and it leads to security and safety concerns in the host countries. Alternatively, a sympathetic and humanitarian view where the discourse revolves around possible ways to help refugees. Similarly, Pope & Griffith (2016) have considered a multi-lingual approach and employed sentiment and emotion analysis on 28,866 English and 24,469 German tweets to probe the Twitter deliberations related to Paris and Cologne events. Findings indicate a prevalent presence of negative sentiments - associated with emotions like anger and anxiety. Özerim & Tolay (2020) also explored Turkish tweets about Syrian Refugees to understand the effect of echo chamber formation on social media. They found that the Twitter discussions mainly revolved around xenophobic and nationalistic themes. An anti-Syrian hashtag, *#ülkemdesuriyeliistemiyorum* (i.e., *#IdontwantSyriansinmycountry* in English) became viral, and it got tweeted and retweeted nearly 50,000 times on a single day in 2016. Similarly, Kreis (2017) analyzed 100 tweets with the hashtag *#refugeesnotwelcome* and found that the discourse of racism against refugees is mainly propounded by the nationalist-conservative and xenophobic right-wing political parties in the European context. Also, they noted hashtags like *#EuropeforEuropeans* by far-right groups in response to the refugee crisis due to the Syrian Civil War. However, in addition to this widespread antipathy and animosity toward migrants and refugees, the dissertation also noted sympathy and solidarity activities by a certain section of the host nations (e.g., Chapter 3).

Extant literature also probed Twitter text contents for analyzing migration movements across nations. For instance, Urchs et al. (2019) tried to extract locational (to understand "where the refugees/migrants are headed") and quantitative (to understand "how many migrants/refugees") information from tweets in the European context. This study has identified 583 tweets about refugees who crossed the border to Hungary, Austria, and Germany in 2015. Similarly, in the context of OECD countries, Zagheni et al. (2014) considered geolocated Twitter data to understand the relationships between internal displacement and international migration.

4.2.2 Refugee Studies Using Image Data

Images can invoke emotions. Hence, existing literature probed the emotional responses of mainstream print and social media to the death of the three-year-old Alan Kurdi in the Mediterranean Sea - one of the most unfortunate refugee-related events in recent times (Adler-Nissen et al., 2020; Bozdag, 2017; Ibrahim, 2018; Olesen, 2018). The shocking image of the drowned Syrian kid made global headlines on September 2, 2015, and immediately, netizens started sharing it on various social media platforms. This image appeared on the screens of 20 million people worldwide in less than 12 hours after its first release, and it became viral with more than 50,000 tweets per hour (Vis & Goriunova, 2015). Adler-Nissen et al. (2020) probed the relationship between images, emotions, and international politics – specifically, how the above image influenced emotional responses and, subsequently, the impact of cumulative emotional responses on international relationships and foreign policy deliberations. They observed that overloaded emotional reactions to the tragic incident of Alan Kurdi had changed the discourse from an open-door approach to stopping refugees from arriving. Similarly, Bozdag (2017) qualitatively analyzed 961 tweets in the context of Turkey and Belgium to investigate the framing and perceptions of refugees before and after the release of Alan Kurdi's image. They did not find any radical shift in the discourse, but the public interest in refugees gained momentum after this tragic incident.

4.2.3 Why Multimodal Data?

Multimodal data was widely used in the domain of food items (Wang et al., 2015), e-commerce business (Bi et al., 2020), depression detection (Gui et al., 2019), meme detection (Kiela et al., 2020), and disaster management (Alam et al., 2018) like floods (Ahmad et al., 2019; de Bruijn et al., 2020; Quan et al., 2020). In addition to social media data, prior multimodal studies also considered satellite images in the context of hydrological data (de Bruijn et al., 2020). Multimodal approaches mostly outperform unimodal approaches (Blandfort et al., 2018; Gallo et al., 2020). For instance, hateful meme detection, where opinions are expressed sarcastically, is challenging because it needs to consider both the contextual knowledge and contents of the meme. Here, multimodal approaches are more efficient than unimodal approaches (Kiela et al., 2020). Multimodal analysis can perform a wide range of tasks. For instance, a food dataset, such as the UPMC Food-101 dataset, which comprises 100,000 recipes for 101 food categories, can be used to analyze and retrieve recipes, dietary assessments, and classification of different food categories (Gallo et al., 2020; Wang et al., 2015). Similarly, in the e-commerce domain, the Rakuten Group has employed multimodal classification to predict each product's type code for defining their catalog (Rychalska & Dąbrowski, 2020). Gui et al. (2019) analyzed depression-related tweets. They observed only text contents might fail to detect the depression accurately, but considering both text and image contents of a tweet improves accuracy. Multimodal Twitter data was widely used for disaster management. For instance, in the context of natural disasters, such as hurricanes, earthquakes, or wildfires, Alam et al. (2018) argue that retrieving real-time information from the text content and the associated image within a tweet can assist in restoration works. Blandfort et al. (2018) have prepared a multimodal dataset of 1851 tweets to investigate gang violence in the US context. They have identified potential psychological antecedents of violence, such as aggression, loss, and substance use, and their multimodal approach outperformed unimodal analysis.

To sum up, prior studies have elucidated the potential of multimodal approaches for depression detection, disaster management, or identifying psychological antecedents of violent activities. However, refugee-related studies mostly considered unimodal data from social media platforms, i.e., either text or visual content. None of the prior studies explored the potential of multimodal data for the societal transition of refugees from their home country to the host nation. Thus, we employ a multimodal fusion approach to explore the refugee journeys from displacement to emplacement in the host country (BenEzer & Zetter 2015).

4.3 Refugee Journey: Les Rites De Passage?

> *"Territorial passages can provide a framework for the discussion of rites of passage … an imaginary line connecting milestones or stakes, is visible – in an exaggerated fashion – only on maps. But not so long ago the passage from one country to another, and, still earlier, even from one manorial domain to another was accompanies by various formalities. These were largely political, legal, and economic, but some were of a magico-religious nature. For instance, Christians, Moslems, and Buddhists were forbidden to enter and stay in portions of the globe which did not adhere to their respective faiths."*
>
> *- Chapter 2 of Les Rites de Passage (p.15)*

> *"…foreigners cannot immediately enter the territory of the tribe or the village; they must prove their intentions from afar and undergo a stage best known in the form of the tedious African palaver. This preliminary stage, whose duration varies, is followed by a transitional period consisting of such events an exchange of gifts, an offer of food by the inhabitants, or the provision of lodging. The ceremony terminates in rites of incorporation – a formal entrance, a meal in common, an exchange of handclasps."*
>
> *- Chapter 3 of Les Rites de Passage (p.28)*

Even after 100 years, the above two excerpts from the English translation of Gennep's antiquarian French work is contemporary and relevant. For instance, there is a striking resemblance between the above century-old "magico-religious" barriers and today's islamophobia in Europe due to the Middle East refugee crisis (Zunes, 2017). Foreigners (i.e., refugees) cannot immediately enter the territory of the host nations. They need to "prove their intentions from afar" to border forces, and there's a lot of "palaver" involved in this process. However, during this period, refugees get food and lodging (i.e., asylums) from inhabitants (i.e., host nations). Finally, rites of incorporation can be equated to the integration of refugees into the society of the host nation. Thus, a handful of prior studies used this theoretical lens in the context of refugee or migration. For instance, Castle & Diarra (2003, p. 2) note that the migration process of young Malians "comprises social and psychological dimensions pertaining to the need to explore new places, experience new settings and accumulate material possessions in order to conform to peer group aspirations." They conclude that the migration of these young Malians "is as much a rite of passage as a financial necessity" (Castle & Diarra, 2003, p. 2).

Gennep's research has identified three distinct phases, namely, *pre-liminal rites* (i.e., 'rites of separation from a previous world'), *liminal (or threshold) rites* (i.e., 'those executed during the transitional stage'), and *post-liminal rites* (i.e., 'the ceremonies of incorporation into the new world') (Van Gennep, 1960). These three stages are commonly referred to as *separation* (getting detached from the previous world and loss of identity), *transition,* or liminal stage (the individual has got detached from her previous world and lost her old identity but not joined the new world), and *incorporation* (getting a new identity after incorporation into the new world). Monsutti (2007) employed this separation-transition-incorporation analogy to explore the journey of young Afghans to Iran. In the initial phase, young Afghans get separated from their families and homes. In the next phase, "they have to prove their capacity to face hardship and to save money … represents a period of liminality", and finally, in the reincorporation phase, they "return to their village of origin … as adult marriageable men" and mostly they continue to commute between Afghanistan and Iran for the rest of their life (Monsutti, 2007, p. 167). Accordingly, we also employed Gennep's framework for probing the refugee journeys. In addition to Gennep's work, we have also referred to refugee-related interdisciplinary research to map these three distinct phases of transitions with the refugee journeys. Refugee journeys mostly start because of conflict and violence in their home countries. In many of these countries, freedom of speech is restricted. For instance, in Afghanistan, Facebook allowed its users to lock their profiles instantly and hide friends lists for security concerns. Moreover, internet penetration is significantly low in many of these nations. Social media data does not allow us to analyze this forced displacement process in their home countries. Thus, in this study, refugee journeys start after they plunge into the ocean. If they are lucky, they arrive at their desired destination.

Phase 1: We have conceptualized *pre-liminal rites* as the *'Arrival of Refugees'*. Refugees are getting separated from their previous world, getting detached from their families, and taking a leap of faith through risky sea routes as anonymous data points after losing their identity. BenEzer & Zetter (2015) pointed out that the "mode of travel may influence the meaning of the journey and its impacts on the individual … For someone who has not … crossed the sea before, the mode of travel will be a highly symbolic part of the experience of the journey. These possible differences, and their meaning, need to be investigated" (p.309). Thus, for this phase, we have considered tweets deliberating or sharing information about the mode of travel or tweets related to border control forces.

Phases 2 & 3: We have considered two interrelated but distinct stages or activities of *liminal (or threshold) rites* as follows: temporal stay at asylums and rehabilitation of refugees. Our second phase is the *'Temporal stay at Asylums.'* According to Turner (2016), asylums are "a place of social dissolution and a place of new beginnings where sociality is remoulded in new ways" (p. 139). Thus, he suggested "explore the precarity of life in the camp … in this *temporary space* (emphasis added)" (p. 139). Accordingly, our tweets in this category capture the details of living conditions in refugee asylums. These tweets also share images of asylums and camps. Our third phase is the *'Rehabilitation of Refugees'*. Khan & Amatya (2017) emphasized the need for health support because most refugees arrive with health problems ranging from infectious diseases to non-communicable musculoskeletal issues. Notably, refugees "face continued disadvantage, poverty, and dependence due to lack of cohesive support in their new country ... This is compounded by language barriers, impoverishment, and lack of familiarity with the local environment and healthcare system" (Khan & Amatya 2017, p. 378). Thus, tweets in this category share information and images of various support activities like arrangements of medical aids, donations of food items or garments etc.

Phase 4: We have conceptualized the *post-liminal rites* as the *'Integration of Refugees'* into the society of the host nation. Charitable organizations arrange various support activities, such as helping them learn a new language. For instance, Abou-Khalil et al. (2019) note that Syrian refugees in Lebanon focus on learning English. In contrast, Syrian refugees in Germany try to learn German for better social inclusion. Hence, this category of tweets shares information about the arrangement of the education system for refugee kids or vocational training programs for adult refugees. These skill up-gradation activities help refugees settle down in the host nations (Bellino & Dryden-Peterson, 2019).

Finally, we have also identified one distinct type of multi-modal tweets where text content might be related to the above phases, but image content shares refugee-related statistics or data points through graphical images or charts. These tweets can be crucial for information dissemination in the context of refugee and migrant-related issues. Hence, we have labeled this category of tweets as *'Infographics'*. It is worth noting that this category is not aligned with Gennep's framework.

4.4 Twitter Data and Annotation Process

Prior studies, such as Alam et al. (2018), Chen & Dredze (2018), and Gui et al. (2019), have considered the Twitter platform for multimodal analysis. Hence, we have performed keyword-based searching using Twitter's advanced search API that allowed us to crawl tweets containing a specific keyword. We have considered a set of keywords like *refugee*, *refugee camps*, *refugee asylums*, *migration*, *immigration*, *immigration policy*, etc. We consider English tweets because prior studies on migration observed that the volume of English tweets was significantly higher than tweets in other languages (e.g., the pilot study of Chapter 1). We have collected 3.98 million refugee-related English tweets from April 2020 to March 2021. Our initial analysis indicates that a significant portion of our corpus does not contain images. For our study, we need to consider tweets with images. Subsequently, we also dropped tweets with similar tweet-ids or similar text contents, and our corpus size became 0.23 million multimodal tweets. Our percentage of multimodal tweets (i.e., 5.7% of 3.98 million) is in accordance with prior studies. For instance, Alam et al. (2018) collected 3.5 million tweets during Hurricane Irma, but they found only 0.17 million images i.e., 4.8% multimodal tweets.

Annotation: For resource constraints, we randomly selected around 2500 tweets from our corpus of 0.23 million tweets. Our manual annotation process requires fine-grained contextual understanding. For instance, to label the first phase of the journeys, i.e., *'arrival of refugees'*, we have considered the following aspects: arrival through sea routes, risk of traveling through sea routes, mode of transport, and activities by border control forces. Similarly, the *'rehabilitation of refugees'* phase has considered activities such as arranging medical aids, charity activities by NGOs, or facilitating donations of essential livelihoods. We have carefully analyzed each tweet based on its text content and the associated image for assigning the final class. In this stage, we discarded tweets with poor images or cryptic short texts. We annotated 1722 tweets with distinct text-image pairs from the above sub-sample with an inter-rater reliability of 0.84. These 1722 tweets are distributed as follows: Phase 1: 398 tweets; Phase 2: 387 tweets; Phase 3: 289 tweets; Phase 4: 343 tweets; and Infographics: 305 tweets. Following Madukwe et al. (2020), we have tried to maintain a balance (in terms of tweet volume or percentage) across our five classes. Alam et al. (2018) prepared Twitter-based multimodal datasets for natural disasters or crises like hurricanes, wildfires, earthquakes, and floods. The final volume of annotated tweets for some of their crises was as follows: 1486, 1239, 499, and 832. Hence, our final sample size for analysis is similar to prior Twitter-based studies. We randomly split our 1722 annotated tweets into 80% (as a training dataset) and 20% (as a test dataset) for our analysis. Table 4.1 reports a few representative tweets from our corpus.

As noted, in contrast to other publicly available image datasets, our image classification is a challenging classification task for computer vision algorithms. Due to the temporal nature of our four phases, common objects across all the categories (except *infographics*) in our corpus are human beings, i.e., refugees or migrants. In other words, we have images of refugees arriving through sea routes, images of refugees with border control forces, images of refugees at asylums, or images of refugees receiving donations or support.

Table 4.1 Representative Multimodal tweets

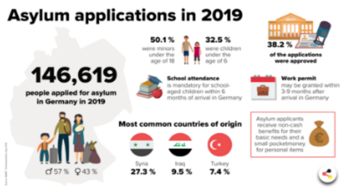
	Visual Input	Language Input
Arrival of Refugees		#Morocco intercepts nearly 200 #migrants trying to reach #Spain: The #Moroccan coast guard intercepted 168 migrants this week who tried to reach Spain using makeshift crafts, including jet-skis and kayaks.
		On August 18, 2020, a total of 167 #migrants were found in the back of a lorry in Samsun, Turkey. The migrants (from Afghanistan and Pakistan) struggled to breathe due to hot weather, and teared the tarpaulin covers during a regular road control by police.
Temporal stay at Asylums		More than 3,000 Syrians living in Idlib's refugee camps have lost their shelter after days of torrential rain and snow. Although some have left their camps, many families have no choice but to live in flooded tents.
		Proposed EU 'Pact on Migration and Asylum' will not help alleviate migration pressure on EU's southern member states. Nadia Petroni – EU's Pact on Migration and Asylum will do little to ease pressure on southern member states. #refugees #EuropeanUnion
Rehabilitation of Refugees		A recent fire destroyed an entire refugee camp on a Greek island and we were able to respond quickly with food. Our monthly donors allow us to respond nimbly. We couldn't do it without our #CTFriends. Set up your monthly gift here.
		UNRWA joined a national polio vaccination campaign that started today. Some16000 Palestine refugee children under 5 will be vaccinated at its health centres in Syria. Child immunization is an important part of primary health care @UNRWA provides to Palestine refugees.

Integration of Refugees

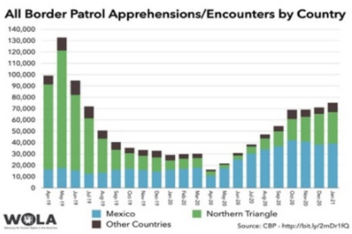
By purchasing the School Enrollment fees for a refugee girl through our online gift shop, you are helping her survive, recover and build a better future.Help improve her chances of escaping poverty for #InternationalWomensDay

#DayoftheGirl is so important. Young women and girls still do not have equal access to human rights, education and health. This needs to change. Today I pledge to keep advocating for refugee girls to make sure they can thrive.

Infographics



At the end of 2019, the number of @Refugees worldwide was 79.5 million. #Germany takes in a great amount of refugees, yet the number of #AsylumSeekers has dropped sharply since the refugee crisis of 2015. We give you an insight into the 2019 #asylum statistics.



Here's undocumented migrants apprehended at the US-Mexico border by Border Patrol since April 2019. The largest increase of these 3 categories, by far, from December to January, was "Other Countries"—47%. Arrivals from Mexico and Central America were up only slightly.

4.5 Methodology and Findings

We have initially considered unimodal models for text, i.e., BERT + LSTM, and visual inputs, i.e., InceptionV3 (Szegedy et al., 2016), VGG19 (Simonyan & Zisserman, 2014), and ResNet (He et al., 2016). Subsequently, we have also employed an early fusion of the above unimodal models. DL-based models are highly efficient for text classification. For instance, RNN models consider previous information for processing the present computation task. In addition to this basic RNN architecture, LSTM consists of three additional gates: input gate, forget gate, and output gate. LSTM calculates the hidden state by considering the combination of these three gates. In LSTM, the input sequence feed is only in the forward direction. However, a BERT-embedded LSTM can consider tokens from both directions.

Language encoders capture the contextual relation between words either context-free or by preserving the contextual information. Earlier approaches, such as Word2vec (Mikolov et al., 2013) or Glove (Pennington et al., 2014), generate embeddings for each word without considering their position within a text and surrounding information. However, bi-directional transformer-based models, like BERT, can capture both directions. Thus, BERT models consider the contextual representation of a word to decipher the difference between two contexts (Devlin et al., 2019). Consequently, LSTM models with BERT embedding are more efficient in capturing the contextual representation than LSTM models without BERT embedding (Minaee et al., 2021). We have considered the English language uncased base version of the BERT model with 12 hidden layers, 768 hidden sizes, and 12 self-attention heads.

For image classification, CNN-based computer vision algorithms are the most efficient. These models contain convolution layers, max pooling, and fully connected layers to solve complex computer vision tasks. One layer's output becomes the subsequent layer's input in the cascading structure. Like language encoders, visual encoders also extract the dominant visual features from an input image and map them into pre-defined categories for downstream tasks like image classification, object detection, or instance segmentation. These encoders extract lower-dimension features using deep neural network-based models.

Our first CNN-based model is VGG-19 (Simonyan & Zisserman, 2014). VGG takes an image size of 224 × 224 × 3 as an input and performs convolution operation using a 3 × 3 filter. We have considered the weights of the VGG19 network based on pre-trained models, but we have trained the output layer using our dataset for the final classification task. Next, we consider ResNet (He et al., 2016) or Residual Networks models. These models skip connection to add the output from a previous layer to a later layer, and they tackle the vanishing gradient problem. For the ResNet model, we have resized all images to 224 × 224 (i.e., rescaled for ease of handling), and the model retains all three RGB channels of the images. We have used ResNet with its ImageNet pre-trained weights to initialize the model, and the outputs are followed by dense layers of 256 units and a SoftMax layer (Deng et al., 2009). Our third CNN architecture-based model is InceptionV3 (Szegedy et al., 2016). This state-of-art third version of Google's initial Inception model explores "ways to scale up net-works in ways that aim at utilizing the added computation as efficiently as possible by suitably factorized convolutions and aggressive regularization" (Szegedy et al., 2016, p.2818). InceptionV3 is known for its efficiency in interpreting images and detecting objects. We have also considered the fourth version of the Inception model for multimodal analysis (Szegedy et al., 2017).

Multimodal fusions leverage the advantages of both modalities to produce a more robust solution for end-to-end system implementation. Thus, multimodal fusion captures the information from content-rich social media data using NLP (for textual inputs) and computer vision techniques (for visual contents) for the downstream task. This approach consists of two parallel deep neural architectures. We employ a transformer-based BERT + LSTM model for textual data and use multiple pre-trained CNN-based models for visual data. In the final layer, we have employed a fusion-based approach to the outputs from these two models to get the final class from a multimodal tweet (refer to Figure 4.1).

Our joint multimodal representation can be expressed as $x_m = f(x_1 \cdot \cdots x_n)$ where x_m is computed using function f that relies on unimodal representations $x_1 \cdot \cdots x_n$. Joint representations are useful for tasks where more than one modality of data is available during the training and inference stages. The simplest example of a joint representation is concatenating different modality features at a low level, also known as an early fusion (Baltrusaitis et al., 2019). Early fusion integrates features directly after they are extracted. Early fusion allows us to perform multimodal representation learning—as it can learn to exploit the correlation and interactions between low-level features of each modality.

Our fusion considers BERT + LSTM for text inputs and ResNet-50, VGG19, InceptionV3, and InceptionV4 for image inputs. We partially freeze the pre-trained weights and add a dense layer with a dropout layer (p = 0.4), followed by a linear layer to extract the latent features. InceptionV3 and InceptionV4 implementation requires an input size of 299 × 299 (i.e., height × width). Hence, we resized the pictures to 299 × 299 for InceptionV3 and InceptionV4. Similarly, we resized the pictures to 224 × 224 for VGG19 implementations. We have standardized the pictures using the original ImageNet training mean and standard deviation. Initially, we performed average pooling of 8 × 8 for InceptionV3 and InceptionV4 and 7 × 7 for VGG. Next, we applied a dropout of 0.4 and flattened it in the next layer. On top of this, we have added a dense unit of 128 before concatenating with the language model stack. In other words, we pass both the visual stack and language stack through a shared dense layer of size 128 and concatenate (i.e., early fusion) the outputs to form a joint vector of length 256. We also apply a dense layer of size 256 before the final SoftMax layer and ReLu activation function. This leads to the final classification layer, i.e., a dense layer with 5 units (for our 5 categories) and SoftMax activation – this will give the predicted class of multimodal inputs. Next, we have used SGD for model optimization. For robustness, we have used different LRs for better learning. We train the model using SGD optimizer, starting with an LR equal to 0.001 and then decreasing it using the Reduce LR on Plateau from keras, and this automatically changes the LR if there is no improvement in training after a certain number of epochs.

We have considered the cross-entropy loss function for our modeling because we have multiple classes in our dataset.

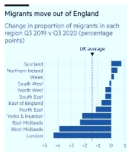
Findings: Table 4.2 reports the accuracies of our unimodal and multimodal models using 5-fold cross-validation. Our BERT + LSTM model for text inputs has reported an F1 score of 58.00%. For brevity, we have not reported the t-SNE plot, but this plot suggests that the weak performance of BERT + LSTM is primarily due to the *Infographics* class, where images differ distinctly from other categories but not the text content. We find VGG19 is our best-performing model for the unimodal visual classification task, and the F1-score is 75.30%. F1-scores of our InceptionV3 and ResNet models are 71.52% and 69.35%, respectively. Like prior studies, such as Blandfort et al. (2018) and Gallo et al. (2020), we observe that our multimodal models have outperformed unimodal models – except BERT + LSTM + ResNet model. Our BERT + LSTM + InceptionV4 has outperformed other models and reported an accuracy of 80.93%.

Table 4.2 Performance of Various Unimodal and Multimodal Models

Models	Modality	F1-Score
BERT+LSTM (Text)	Unimodal	58.00%
VCG19 (Visual)		75.30%
ResNet50 (Visual)		69.35%
Inception V3 (Visual)		71.52%
BERT+LSTM + VGG19	Multimodal	79.69%
BERT+LSTM + ResNet50		70.00%
BERT+LSTM + Inception V3		79.06%
BERT+LSTM + InceptionV4		**80.93%**

Error Analysis: Table 4.3 reports our multimodal models' input text, input image, true label, and predicted label of a few sample tweets. For instance, BERT + LSTM + InceptionV3 has failed to classify 'Example 4' correctly. The text content of Example 4 indicates their stay in the camp house, but the image is generic – not offering much information about the 4 kids. However, the tweet also talks about the upgrade from 'living in open spaces at the border' to 'private camp house to sleep in.' A significant portion of rehabilitation tweets talk about the upgradation of their status. Probably, our model considers this up-gradation in their living condition as an indicator of 'rehabilitation.'

Table 4.3 Error Analysis of Multimodal Tweets

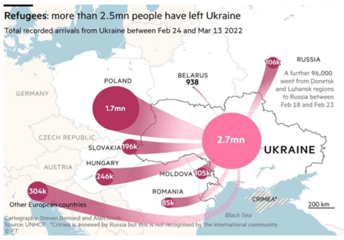
<table>
<tr><th></th><th>Example 1</th><th>Example 2</th></tr>
<tr><td>Input Image</td><td></td><td>
</td></tr>
<tr><td>Input Text</td><td>Help us reach our goal of raising $10,000 (by Dec 10) to support students from refugee and asylum seeker backgrounds at UTS! The funds we raise will help them pay for food, rent, textbooks and cost of living</td><td>The importance of immigration and multicultural society. But, Tory Gov supported by Tory Scots ... and their support for English nationalism, Brexit n anti-immigration policies has destroyed Scotland's economy.</td></tr>
<tr><td>True Label</td><td>Integration of Refugees</td><td>Infographics</td></tr>
<tr><td>Predicted Label</td><td>Integration of Refugees</td><td>Infographics</td></tr>
<tr><td></td><td>Example 3</td><td>Example 4</td></tr>
<tr><td>Input Image</td><td></td><td></td></tr>
<tr><td>Input Text</td><td>"It is essential to ensure healthcare during this humanitarian crisis." -Leonardo, nurse on the border, MedGlobal has supported nurses who provide health screenings; first aid, ensuring Venezuelan migrants;</td><td>... a refugee mother of 4 children from is happy to have a family shelter in ... camp. "My children and I had been living in open spaces at the border for long time. Now I am glad that we finally have our private camp house to sleep in" she said.</td></tr>
<tr><td>True Label</td><td>Rehabilitation of Refugees</td><td>Temporal stay at Asylums</td></tr>
<tr><td>Predicted Label</td><td>Rehabilitation of Refugees</td><td>Rehabilitation of Refugees</td></tr>
</table>

4.6 Ukrainian Refugee Crisis: An Application

The previous section leads to the intriguing question: is our proposed framework of four phases generic or context-specific? To probe the practical relevance of our approach, we have considered the 2022 Ukrainian refugee crisis. More than 4 million Ukrainians had to move to neighboring countries, such as Poland, Romania, Hungary, or Moldova, in a span of one and a half months. Another 6 million were displaced within the country because of the military invasion (UNHCR Data Portal). Ukrainians got displaced from their previous world and left behind all their belongings (and sometimes family members) toward an uncertain future.

Table 4.4 Tweets from 2022 Ukrainian refugee crisis

	Visual Input	Language Input
Arrival of Refugees		A Ukrainian refugee pushes her baby in a pushchair as they arrive at the Medyka border crossing, Poland, Saturday, Feb. 26, 2022.
		Every 2 or 3 minutes, a bus full of refugees is arriving into Korczowa's makeshift reception centre. This is what Europe's fastest-growing refugee crisis since WW2 looks like #Poland #Ukraine
Temporal stay at Asylums		Our camera is the first inside this refugee center in Rzeszow, #Poland 800 beds, food, clothes.. and this place where kids who've come through bullets and bombs can be kids again. #Ukraine
		... People lie on camp beds at a refugee reception centre at the Ukrainian-Polish border crossing in Korczowa, Poland #UkraineRussiaWar #Russia #Ukraine #UkraineWar
Rehabilitation of Refugees		Great to be in @_______ this morning and to see the donations they're collecting to go to Poland, for refugees from #Ukraine Closer to home thanks for all the donations collected for our @_____ refugee foodbank too
		Polish soldier giving a teddy to a refugee child from Ukraine. #Ukraine #UkraineRussianWar #StandWithUkraine #UkraineRussiaWar #UkraineUnderAttack

Integration of Refugees		Ukrainian refugee children fleeing the Russian assault are welcomed with cheers by their Italian peers on their first day at a primary school in Pomigliano d'Arco, Italy, a video shared on social media shows
		One of our Durlston Dads is headed to Poland tomorrow with aid for the Ukraine. We have been busy making Happiness Postcards for him to give to the refugee children - a little symbol of love and unity. #DurlstonFamily
Infographics		The number of Ukrainians fleeing the fighting reached 2.7mn by March 13, the UN's refugee agency reported, amid concerns over the growing refugee crisis. The country taking the highest number of refugees is Poland, with 1.7mn alone.
		Nearly one Ukrainian child becomes a refugee every second, as the UN says an average of 73,000 children a day have escaped Russia's onslaught over the last 20 days

Data: We have collected 0.6 million tweets, out of which around 10,000 tweets were multimodal, from February 24, 2022, to March 15, 2022. Our data collection and pre-processing strategies were similar to the previous analysis, but for data crawling - we have added a few crisis-specific keywords such as *Ukraine*, *Russia*, and so on. Next, we have explored the linguistic and image contents of this corpus. Some of the visible differences are - that refugees mostly took the sea routes in the previous corpus, whereas it was mostly rail or road transport during the 2022 crisis. Earlier, the travel risk was boat capsizing, and in 2022 it was a missile attack. Similarly, we find images of camps and tents in the previous corpus, whereas it was a temporary makeshift arrangement in hotels and other large public buildings during the 2022 crisis. Broadly, the struggles, risks, and traumas of refugees are identical. A few representative tweets in Table 4.4 confirm the same. Thus, the intriguing question is: can we apply our proposed framework during this 2022 crisis? If yes, the follow-up question is – whether the previous corpus collected prior to this 2022 crisis (as training data) can help us to extract actionable and relevant information in real-time during the Ukrainian crisis.

Findings: To test the same, a single annotator (i.e., with prior experience in handling migrant-related tweets) has annotated 234 multimodal tweets (evenly distributed across four phases and the infographics class) and employed our best performing model, i.e., BERT+ LSTM + InceptionV4, on this unseen test data. For a complex unseen dataset like ours, we find that the F1-score came down to 71.88%. As expected, F1-scores of BERT+ LSTM (for unimodal text inputs) and VGG19 (for unimodal visual inputs) models are 50.01% and 59.17%, respectively. Unlike the previous corpus, this 2022 corpus also elucidated the evolution of Twitter deliberations. For instance, initial tweets in February were mainly about the arrival and temporal stays, and in the later period, tweets related to rehabilitation and integration gained momentum.

4.7 Discussion and Future Research

Refugee-related deliberations have gained momentum in recent times. Twitter gives netizens a platform to raise their voice and share their views. Hence, refugee-related issues are also getting deliberated on Twitter. From the AI domain, prior studies have mostly considered a specific refugee-related event or crisis for their analysis. Opinion mining in the context of one particular event, such as the involvement of a migrant in terrorism in France or the unfortunate drowning in the Mediterranean Sea, offers a nuanced understanding of public opinion around that specific event. However, this approach does not provide a holistic view to activists, refugee workers, and policymakers. To the best of our knowledge, none of the prior studies considered multimodal data to probe refugee journeys from displacement to emplacement.

We have employed the *Les Rites de Passage* framework to elucidate the societal-level transitions of refugees from home to host nations. It is worth noting that we have not attempted to track the journey of an individual refugee for ethical concerns but explored the societal-level transitions of refugees using multimodal refugee-related tweets - instead of an unimodal approach. Our study confirms that this framework can be used to extract relevant and actionable phase-wise information from social media platforms. We note, especially in the context of the 2022 Ukrainian crisis, that refugee needs to evolve from essential health support in the initial days to social integration in the later days. In the domain of applied AI-based studies, our study may aid policymakers in understanding phase-wise concerns and taking appropriate actions. We conclude that it hardly matters whether someone is taking the risky sea route to reach the shore of the host nation or opting for the rail route or migrant caravan to enter the host nation. The traumatic journey from displacement to emplacement is the same– irrespective of who they are or wherever they come from!

On the methodology front, our image inputs, which comprise mostly human beings (i.e., refugee kids, adult refugees, health workers, border force, educators, and so on), are significantly challenging in comparison to publicly available image datasets like CIFAR-10 that comprises of distinctly different objects like flowers, animals, or buildings. Thus, in the context of applied AI-based research, our study confirms that multimodal models can outperform unimodal models even for a challenging classification task like ours.

Future Research: We have also identified a few exciting avenues for future research. For instance, prior NLP-based studies widely used opinion mining of Twitter data. Hence, we have examined the text content of our initial corpus, i.e., 1722 annotated tweets, using linguistic inquiry and word count (LIWC) software (Tausczik & Pennebaker, 2010). LIWC is a language tool designed to capture opinions and perceptions by analyzing text contents, and it does not consider multimodal inputs – this is one of the potential limitations. Based on our exploratory unimodal analysis, we observe two distinct patterns in our data.

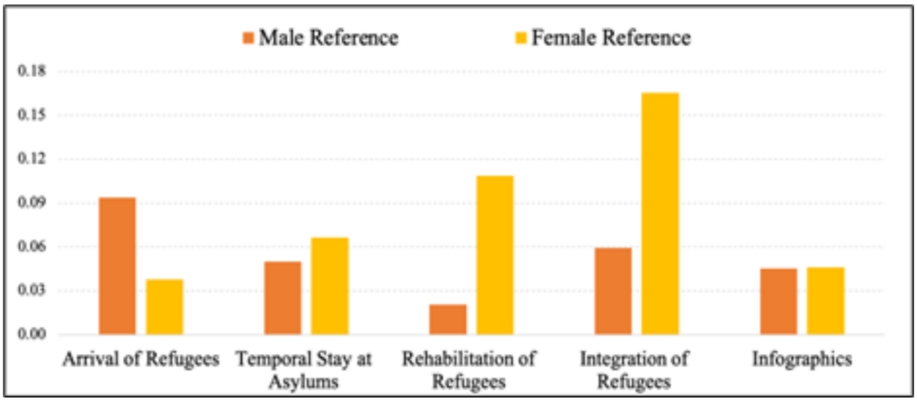


Figure 4.2 Gender Analysis of our corpus using LIWC

First, we have performed LIWC-based male vis-à-vis female reference analysis, and the vertical axis reports the average score of our 5 categories (refer to Figure 4.2). We find a distinct pattern. For example, the arrival of refugee class is more associated with male references than female references, whereas female references are significantly higher for rehabilitation and integration classes. Infographics is a gender-neutral class. Probably, our findings indicate that female refugees are getting more support, whereas safety concerns associated with refugee arrival are more associated with male refugees. Similarly, Rettberg & Gajjala (2016, p.179) explored "the portrayal of male Syrian refugees in a post-9/11 context where the Middle Eastern male is often primarily cast as a potential terrorist … the claim that the Syrian refugees are primarily male is often repeated on *#refugeesnotwelcome* through images of men with text highlighting the absence of women and children."

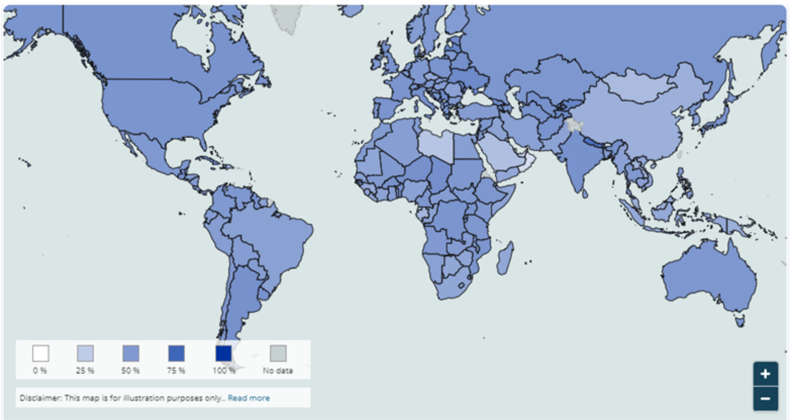


Figure 4.3 Share of female migrants at mid-year 2020

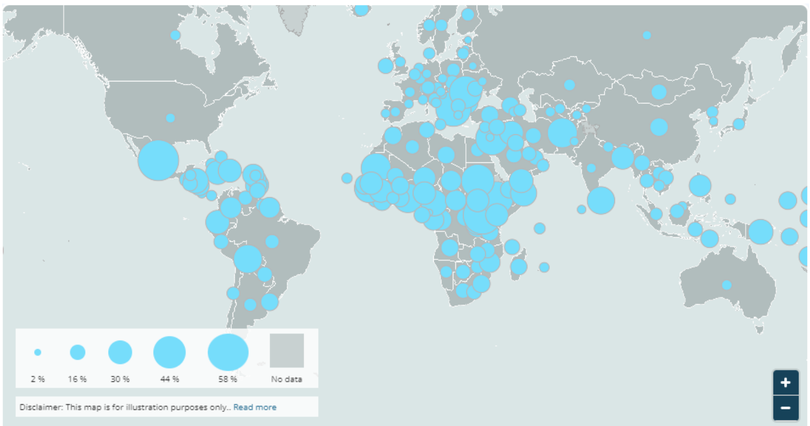


Figure 4.4 Share of international migrants under 18 in 2019

Subsequently, we have referred to the Migration Data Portal to explore the share of female migrants in the international migrant stock in mid-year 2020 (Source: UN DESA, 2020, Based on the latest data available on 20 January 2021). Figure 4.3 reports the same. Globally, 48.1% of migrants are female. So, associating the arrival of refugees with male references reflects the bias of the society. This percentage varies from 45% to 55% in all leading nations such as Canada (52.4%), France (51.5%), Germany (49.9%), Italy (53.6%), the UK (52.3%), and the USA (51.7%). Likewise, in Figure 4.4, the share of international migrants under 18 residing in the country/region in 2019 is also not insignificant (Source: UNICEF, 2020, Based on the latest data available on 11 May 2020). The global average is 12%, and it is mostly consistent in all leading nations. Thus, as we noted in Chapter 3, societal perception might not be based on reality because "social perceivers also go beyond the information actually present in the current environment through the activation of social stereotypes based on easily detectable identifying features of social groups" (Dijksterhuis & Bargh, 2001, p. 9).

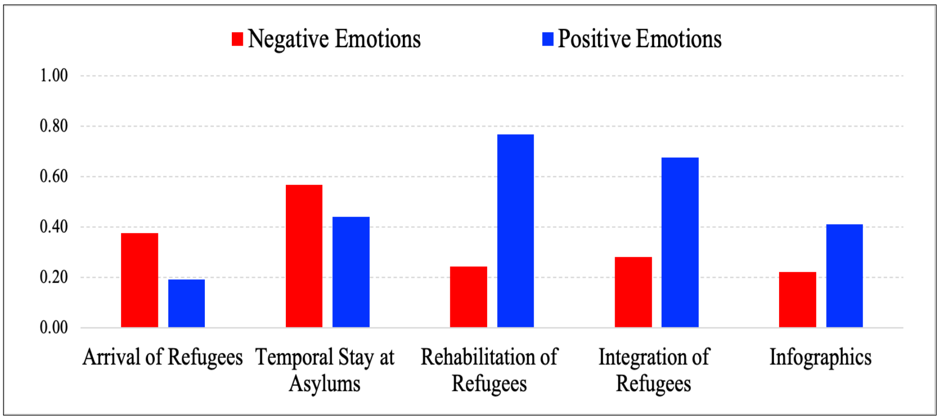


Figure 4.5 Emotion Analysis of our corpus using LIWC

LIWC-based emotion analysis was also used in prior studies (Saha et al., 2019). LIWC[37] manual considers "a number of cognitive strategies, several types of thematic content, and various language composition elements" to capture the positive and negative emotions (p. 8). The vertical axis of Figure 4.5 reports the average score of positive and negative emotions for each category. Once again, we find an interesting pattern. Support activities, such as rehabilitation and integration of refugees, are primarily associated with positive emotions. On the contrary, the first two phases, i.e., arrivals of refugees and temporal stay at asylums, are predominantly displaying negative emotions. The puzzling question is - whether these negative emotions reflect xenophobic mindsets. Are social media users sympathetic to refugee sufferings, and do negative emotions reflect the sadness? Future studies need to probe this. Our study elucidates the need to consider these finer nuances of Twitter deliberations for getting an in-depth understanding of refugee-related issues.

[37] Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Available at https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf?Sequence=3, Accessed on July 1, 2022

Chapter 5: Endorsement Analysis of Migration Discourse

5.1 Do we Endorse Xenophobia?

Xenophobia, i.e., a fear of strangers or foreigners toward migrants, has attracted the attention of scholars in the domain of migration (Peterie & Neil, 2020; Rivera-Pagán, 2012). AI researchers are probing social media data to probe the same (Aguirre & Domahidi, 2021; Basile et al., 2019; Vázquez & Pérez, 2019). Prior studies noted not only instances of negative sentiments in the context of migrant-related discourse (Lee & Nerghes, 2018; Öztürk & Ayvaz, 2018; Pope & Griffith, 2016) but also instances of hate speech (Aguirre & Domahidi, 2021; Basile et al., 2019; Vázquez & Pérez, 2019). However, most of these studies have considered a small sample of manually annotated data (Kreis, 2017; Latorre & Amores, 2021). These studies offer a nuanced understanding of various aspects of migrant-related deliberations, but a few hundred annotated text inputs do not provide insights into the prevailing societal-level sentiments of migrant-related discourse. To the best of our knowledge, none of the prior studies performed large-scale endorsement analysis of societal-level sentiments in the context of migrant-related deliberations. This issue is intriguing in the context of the 2022 Ukrainian refugee crisis because of the Russian invasion on February 24, 2022. European countries have wholeheartedly extended their solidarity to Ukrainians during this crisis. However, one portion of the society argues that Europeans are more sympathetic to Ukrainians, but they expressed xenophobic and hostile behaviors during the earlier refugee crises (Hauck, n.d.; Khalid, n.d.). Thus, the intriguing question is – what do we endorse? In other words, our research questions are: *Do users endorse migrant-related comments with positive or negative sentiments (such as abusive comments)? Notably, does it depend on who the migrants are?*

To address the above research questions, we perform a large-scale analysis of user interactions on the YouTube platform. We have considered two datasets: 110,803 migrant-related comments from 2778 videos posted from January 2018 to December 2020, i.e., before the 2022 Ukrainian crisis, and 21,453 migrant-related comments from 342 videos posted from February 2022 to April 2022, i.e., during this crisis. We form a set of competing hypotheses. To analyze the endorsement patterns, we employ Logit and Tobit models (Greene, 2003). We find users endorse migrant-related positive comments for both datasets. We also note a negative propensity to endorse migrant-related comments with abusive words. However, negative sentiments may depend on who the migrants are, i.e., whether they are *cultural others* or *cultural us*.

5.2 Migrants: Cultural us, Cultural others, or Both?

Prior social science studies have employed the racism theory to explain the attitude toward migrants. Racism theory considers cultural aspects, like race or skin color, to explore the behavior toward migrants from different cultures. Racism not only considers migrants as *'Others'* with different skin color but also make a biased assumption that migrants are mostly "drug dealer or pimp" (Peterie & Neil, 2020). Interestingly, the anti-migrant section of society tries to justify their stance by arguing that they do not want to form negative opinions, 'but' their prior or recent experience compelled them to think differently (Peterie & Neil, 2020). In addition to ethnic differences, antipathy toward migrants is also due to cultural and religious differences like *'Islamophobia'* in Christian-majority countries. European countries perceive Islam as a cultural threat and associate Islam with no freedom of expression (in Denmark), lack of gender equality (in France), or lack of tolerance (in Switzerland) (Peterie & Neil, 2020). These cultural differences are also known as 'Otherness'- a form of cultural racism.

Cultural racism also gets captured on social media platforms. For instance, Lee & Nerghes (2018) analyzed the comments of two YouTube videos posted in September 2015 to unravel the sentiments toward the European refugee/migrant crisis. They have identified five themes. Three of them are generic: Refugee/refugee crisis, Migrant/migrant crisis, and Immigrant/immigrant crisis. The remaining two themes were context-specific: Syrian migrant and refugee; and the final theme is related to *threat perceptions*, and some of the labels are *Jihadist*, *terrorist*, *criminal*, and *rapefugee*. This last theme implicitly reflects cultural racism, as we noted above. Similarly, Pope & Griffith (2016) employed LIWC to explore the sentiments of Twitter discourse during the Paris attack (on November 13, 2015) and sexual assaults in Cologne (on December 31, 2015). They found the sentiment of Twitter deliberation was negative for both events; however, the negativity of the discourse was severe after the *terrorist activities* in Paris. Additionally, they also note higher anxiety and anger toward migrants after these incidents. Öztürk & Ayvaz (2018) explored English and Turkish tweets in the context of the Syrian refugee crisis. Interestingly, Turkish tweets were more positive toward Syrian refugees, whereas English tweets were primarily negative. This also implicitly hints at the '*cultural other'* argument. Thus, drawing insights from the extant literature, we argue that if social media users are broadly xenophobic, they will endorse migrant-related negative comments instead of positive comments, especially when migrants are *cultural others*. Contrarily, if users are xenophilic, they will endorse positive comments and not endorse migrant-related negative comments – even if migrants are *cultural others*. For instance, advanced nations, especially European countries, may consider Ukrainian migrants as *cultural us*; however, they can perceive migrants from Syria or Africa as *cultural others*. Hence, we hypothesize,

***Hypothesis 1A** (Cultural Other Assumption): Ceteris paribus, the propensity of liking migrant-related comments with negative (positive) sentiments will be positive (negative or insignificant).*

***Hypothesis 1B** (Cultural Us Assumption): Ceteris paribus, the propensity of liking migrant-related comments with negative (positive) sentiments will be negative (positive).*

One extreme case of negative comment is using swear or abusive words – also known as hate speech. Hate speech not only differs in terms of its intensity but can also be classified into categories such as aggressiveness, offensiveness, irony, and stereotype (Sanguinetti et al., 2018). Hate speech literature has identified various target groups based on race, color, gender, religion, and ethnicity (Sagredos & Nikolova, 2022; Tang et al., 2021; Waseem & Hovy, 2016). Theoretically, the genesis of a target group, which ranges from *anti-black-people* sentiment to *Islamophobia*, can be explained through the lens of cultural otherness. A plethora of studies explored hate speech toward migrants and refugees across contexts (Basile et al., 2019; Latorre & Amores, 2021; Palakodety et al., 2020; Sanguinetti et al., 2018; Vázquez & Pérez, 2019). For instance, existing literature noted hate speech toward migrants in Italy (Sanguinetti et al., 2018; Vigna et al., 2017), Spain (Calderón et al., 2020a; 2020b; Vázquez & Pérez, 2019), Germany (Karakayali, 2018), Myanmar (Palakodety et al., 2020), and Korea (Kim et al., 2020). However, most of these studies explored Twitter discourse, and only a handful of studies explored YouTube deliberations (Latorre & Amores, 2021; Palakodety et al., 2020).

As pointed out earlier, migrants differ in terms of race, color, religion, and ethnicity. A recent study, which has analyzed one of the most exhaustive hate speech datasets, pointed out that migrants are one of the most common target groups (Vidgen et al., 2021). Other target groups are 'Black people', 'Muslims', and 'Women' (Vidgen et al., 2021). Unfortunately, a migrant woman from Syria can be a target for hate speech not only for her migrant status but also for her gender, religion, and ethnicity (Hattar-Pollara, 2019). In accordance with the cultural racism argument, Vigna et al. (2017) identified various antecedents of hate speech: religious regions, racial reasons, socio-economic reasons, etc. Generally, high-profile events, such as the 2017 Rohingyas crisis or the 2018 rejection of the boat Aquarius by the Italian government and acceptance by the Spanish government, become not only trending topics but also these events may trigger hate speech on social media platforms (Palakodety et al., 2020; Vázquez & Pérez, 2019). Additionally, hate speech is also triggered by twisted narratives by right-wing politicians against migrant workers, and it agitates the vulnerable working class of host nations (Hoewe et al., 2020; Jaki & Smedt, n.d.; Ottoni et al., 2018). Thus, if most users are xenophobic, migrant-related comments with abusive or swear words would get endorsed. Otherwise, these comments would not get the endorsement. Hate speech literature has convincingly associated swear words with abusive or offensive comments (Pamungkas et al., 2020). Hence, our competing hypotheses are,

Hypothesis 2A *(Cultural Other Assumption): Ceteris paribus, the propensity of liking migrant-related comments with swear words will be positive.*

Hypothesis 2B *(Cultural Us Assumption): Ceteris paribus, the propensity of liking migrant-related comments with swear words will be negative.*

5.3 Data and Methodology

YouTube is the second most visited website after Google search (Thelwall & Foster, 2021; Foster, 2020). Unlike other social media platforms, the recency effect is not salient on the YouTube platform. Thus, prior studies considered YouTube data for probing consumption patterns (Thelwall, 2021). For instance, users mostly see trending hashtags or tweets on Twitter or Facebook. Thus, Twitter deliberations immediately after migrant-related violence can display animosity toward migrants. However, when users search for a specific topic on the YouTube platform, it recommends a list of relevant and popular videos already liked by thousands (or millions) of users – not only the recently published videos. A user can watch and post comments, and later, other users can also read and endorse these comments by liking them (Siersdorfer et al., 2010; Foster, 2020; Thelwall & Foster, 2021). Thus, the YouTube platform allows accessing and analyzing past discourse. Consequently, user endorsement analysis on the YouTube platform may not be fraught with the pitfall of temporal emotional swings in response to a specific event. However, prior migration-related studies have rarely probed migrant-related comments on the YouTube platform – except for a few event-specific studies (Aguirre & Domahidi, 2021; Lee & Nerghes, 2018; Spörlein & Schlueter, 2021). We have employed YouTube's API for data collection. We have considered an exhaustive set of keywords for crawling migration-related videos, such as *migration*, *immigration*, *refugee*, *asylum*, etc. We dropped video ids without comments. We have collected the data in two phases as follows:

Dataset 1: 110,803 migrant-related comments (i.e., comments specifically mentioning migrant-related keywords) from 2778 videos posted from January 01, 2018, to December 31, 2020. We extracted these comments in November 2021, i.e., nearly three months before the Ukrainian refugee crisis.

Dataset 2: 21,453 migrant-related comments from 342 videos posted from February 16, 2022, to April 14, 2022. However, except for 3 videos, the remaining 339 videos were posted only after February 24, 2022. We extracted these comments in April 2022.

As noted, supervised approaches require manually annotated training data and rarely allow a large-scale analysis. Consequently, a limited sample size does not allow to probe societal-level endorsement patterns reliably. Moreover, neural network models would not allow simultaneous testing of the causality of multiple input variables. Thus, we have employed regression analysis on the methodology front, specifically Logit and Tobit regression models (Greene, 2003). Following prior studies, we have considered LIWC-based analysis for probing the sentiments of YouTube comments (Pope & Griffith, 2016; Tausczik & Pennebaker, 2010).

Table 5.1 Video-level Statistics

	N	Mean	Min	1st Qu.	Median	3rd Qu.	Max	Skew.	Kurt.
Dataset 1 (January 01, 2018, to December 31, 2020)									
Views	2778	82400.8	2	2221.0	8269.0	32738.0	21966597	27.9	1023.3
Comments	2778	411.3	1	15.0	57.0	242.8	107489	31.5	1266.8
Likes	2732	1381.2	0	32.0	118.0	465.5	277935	21.4	577.9
Dataset 2 (February 16, 2022, to April 14, 2022)									
Views	342	296283.9	51	15312	43466.0	165023.0	16540758	10.2	130.7
Comments	342	1480.9	2	120.8	371.0	1181.2	33794	5.3	34.0
Likes	337	4509.4	3	191.0	628.0	2647.0	136120	6.3	43.1

5.3.1 Dependent Variables

Our DVs are user endorsement patterns. We have assumed that if users like a comment, they endorse the sentiments expressed in that comment. For dataset 1, we note that 61% of comments did not receive any endorsement (i.e., zero-like) by other users. Next, 28% of comments receive 5 or less than 5 likes. Only 5% of comments received more than 20 likes. The pattern was also similar for dataset 2. Overall, the user endorsement distribution was skewed. Hence, we have operationalized our DV as a binary dummy variable, i.e., *Like_Dummy* (1 = liked/endorsed by users, and 0 otherwise). It is worth noting that a binary dummy variable will not differentiate between 10 likes and 10,000 likes. However, from the perspective of user endorsement, they are not the same. Thus, we have also considered the log transformation of comment-level likes, i.e., *Log(Like_Counts)*, as our second measure of DV, i.e., user endorsement.

5.3.2 Explanatory Variables

Our H1A & H1B investigate the relationship between user endorsement and comment-level (CL) sentiments. LIWC analyzes the usage of opinionated words, both positive and negative words, of an input text to assign the sentiment scores. Similarly, H2A & H2B explore the relationship between user endorsement patterns for hate speech. For H2A & H2B, we have considered the 'swear words' score of LIWC. LIWC considers a list of offensive words (such as f*ck, d*mn, sh*t, etc.) to calculate the *Swear_Words* score. We have considered these scores for testing our proposed hypotheses.

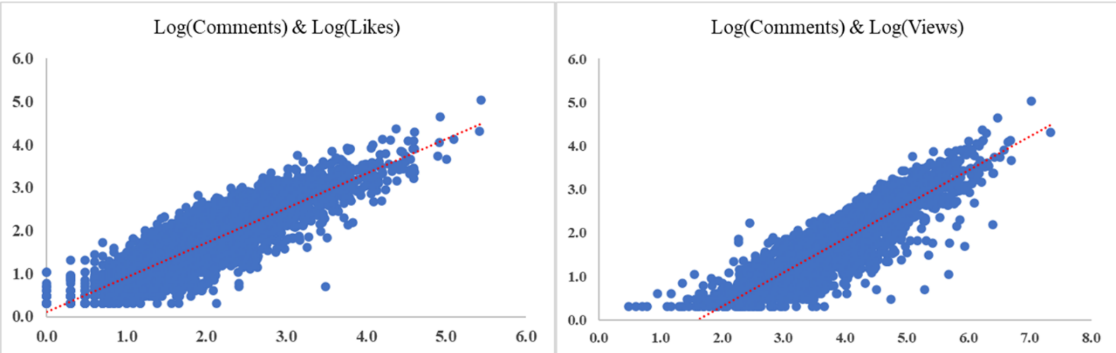


Figure 5.1 Relationships between VL Variables for Dataset 1

5.3.3 Controlling the Effects of Video-Level (VL) Features

The user endorsement also depends on VL interaction variables, such as overall video popularity (Siersdorfer et al., 2010; Foster, 2020). Hence, we need to control the VL user interactions, such as views, comments, and likes. Table 5.1 reports the distribution of these variables for both datasets. It is interesting to note that the mean values are significantly larger than the median values. The gaps between the 3rd quartiles and max values are substantially large. This indicates that user interactions for popular videos are substantially higher than the rest. Consequently, the skewness and kurtosis scores are way beyond the acceptable cutoff. Hence, we have performed a log transformation. Next, we observe a very high correlation between these indicators. In other words, a higher view of a video will also fetch more comments or likes, and vice versa. Figures 5.1 and 5.2 graphically report these relationships for Dataset 1 and Dataset 2, respectively. This high correlation will create multicollinearity issues in regression models, and we cannot incorporate all of them in the same regression model. Thus, we have considered only the log transformation of video-level views, i.e., *Log (VL-Views)*, to control the video-level popularity.

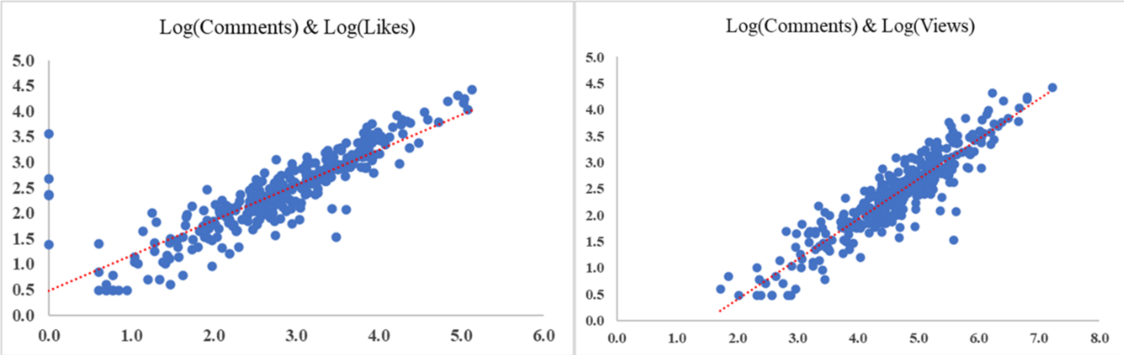


Figure 5.2 Relationships between VL Variables for Dataset 2

Media houses posted many videos in our corpus, and existing literature has pointed out that the political leanings of media houses are not similar (Bovet & Makse, 2019; Ribeiro et al., 2020). Hence, we referred to the media bias fact check (MBFC) resource for political biases (Media Bias/Fact Check News, n.d.). MBFC broadly classifies media houses into two categories: *left-leaning channels* emphasizing community over the individual. Contrarily, *right-leaning channels* emphasize individualism. In the context of immigration, left-leaning media houses 'support a moratorium on deporting' and are sympathetic to undocumented immigrants, whereas right-leaning media houses are generally against undocumented immigrants. Right-leaning channels support restrictive actions and expect more robust border control. Hence, in our regression model, we introduce two variables as follows: left and center-left channels as *Left-Leaning Chnl* variable, and right and center-right channels as *Right Leaning Chnl* variable (Ribeiro et al., 2020). For dataset 1, 39% of videos were posted by left-leaning media, whereas 14% were from right-leaning channels. However, during the 2022 crisis, left-leaning media houses posted 42% of the videos in our corpus, and only 4 % of the videos were from right-leaning media houses. As mentioned earlier, to control the confounding effects of media biases on user endorsement, we need to control the political leaning of media houses in regression models. Thus, we have created two dummy variables for political leaning (1 = left-leaning, 0 otherwise; and 1 = right-leaning, 0 otherwise) for dataset 1. We have considered all remaining video ids, which were either posted by least-biased (i.e., center) channels or not published by media houses, as the base category. For dataset 2, we have considered only one dummy variable (1 = left-leaning, 0 otherwise) because we cannot incorporate a dummy variable for 4% of our sample (i.e., 4% of videos posted by right-leaning channels – refer to Table 5.2). Thus, we have clubbed these 4% videos with our base category.

5.3.4 Controlling the Effects of Comment-Level Features

The propensity of getting an endorsement would be high for an initial comment in response to a video (Foster, 2020). So, we have controlled the gap between VL time stamps (VLTS) and CL time stamps (CLTS). To control large variances, we have taken a log of the time-stamp gaps: *Log (VLTS - CLTS)*. Next, we find that the distribution of comment lengths (i.e., words in a comment) is also skewed. For dataset 1, we note that around 30% of comments are short, i.e., 20 or less than 20 words. 37% of comments have 21 to 50 words. Only 13% of comments have more than 100 words. Interestingly, the pattern was almost similar for dataset 2: 28% of comments with 20 or less than 20 words, 38% of comments with 21 to 50 words, and only 13% of comments with more than 100 words. Hence, we have controlled the length of the comment by taking the log of word counts of the comment, i.e., *Log (CL-Word Count)*.

5.4 User Endorsement Analysis

5.4.1 Descriptive Analysis

Table 5.2 juxtaposes the descriptive statistics of both datasets. The mean value of our first dependent variable, i.e., endorsement or *Like_Dummy*, is 0.39 (for dataset 1) and 0.42 (for dataset 2). This indicates that users liked 39% of comments, and the remaining 61% did not receive endorsement for dataset 1. Similarly, users endorsed 42% of comments for dataset 2. This becomes almost similar after taking the log transformation of like counts, i.e., *Log(Like_Counts)*. As expected, for the variable *Log(VLTS - CLTS),* we note a significant difference between the two datasets because the average time gap between the videos posted and data collection was significantly higher for dataset 1 compared to dataset 2.

However, the *Log (VL-Views)* and *Log (CL-Word Count)* are slightly higher for dataset 2. Thus, compared to dataset 1, videos posted during the 2022 crisis fetched higher views; users wrote lengthy comments in response to these videos and endorsed 42% of these comments. Regarding societal sentiments, we note that the Positive_Sentiment (i.e., 4.17) score for dataset 2 is slightly higher than dataset 1 (i.e., 3.99). Contrarily, the scores of *Negative_Sentiment* (2.31), as well as Swear_Words (0.14) for dataset 2, are less than the scores of *Negative_Sentiment* (2.39) and *Swear_Words* (0.33) of dataset 1. A priori, this indicates that user involvement was slightly higher during the 2022 crisis, and comments were more (less) positive (negative and abusive).

Table 5.2 Descriptive Statistics of Datasets 1 and 2

Variables	Mean	S.D.	Min	Max
Dataset 1 (January 01, 2018, to December 31, 2020)				
Like_Dummy	0.39	0.49	0.0	1.0
Log(Like_Counts)	0.27	0.48	0.0	4.1
Log(VL-Views)	5.38	0.95	0.3	7.3
Log(VLTS - CLTS)	1.13	1.05	0.0	3.1
Left_Leaning_Chnl.	0.39	0.49	0.0	1.0
Right_Leaning_Chnl.	0.14	0.34	0.0	1.0
Log(CL-Word Count)	1.54	0.41	0.0	3.9
Positive_Sentiment	2.93	3.99	0.0	83.3
Negative_Sentiment	2.39	3.42	0.0	60.0
Swear_Words	0.33	1.51	0.0	50.0
Dataset 2 (February 16, 2022, to April 14, 2022)				
Like_Dummy	0.42	0.49	0.0	1.0
Log(Like_Counts)	0.27	0.44	0.0	3.8
Log(VL-Views)	5.48	0.74	1.7	7.2
Log(VLTS - CLTS)	0.45	0.46	0.0	1.7
Left_Leaning_Chnl.	0.42	0.49	0.0	1.0
Right_Leaning_Chnl.	0.04	0.18	0.0	1.0
Log(CL-Word Count)	1.56	0.40	0.0	3.2
Positive_Sentiment	3.11	4.17	0.0	100.0
Negative_Sentiment	2.31	3.12	0.0	50.0
Swear_Words	0.14	0.90	0.0	50.0

5.4.2 Logit and Tobit Models

Tables 5.3 and 5.4 report the results of our proposed hypotheses for dataset 1(i.e., before the 2022 crisis) and dataset 2 (i.e., during the 2022 crisis), respectively. In both these tables, models 1 to 4 report Logit models where the dependent variable is a dummy variable, i.e., *Like_Dummy*. Similarly, models 5 to 8 report Tobit models where the dependent variable is a log transformation of like counts, i.e., *Log (Like_Counts)*. Tobit models, also known as censored regression models, are appropriate when the dependent variable is either left or right-censored. In our case, comment-level-like counts are left-censored, i.e., the minimum value cannot be less than zero, and this violates the normality assumption. For instance, we observe that 58% of comments did not receive any endorsement (i.e., zero-like), 32% of comments received 5 or less than 5 likes, and only 4% of comments received more than 20 likes in our dataset 2. In this scenario, censored regression models are appropriate to estimate the proposed relationship between left-censored like counts and sentiment scores or swear words scores. For Logit and Tobit models, if an explanatory variable has a positive coefficient, the propensity to get likes will positively correlate with that variable, and vice versa (Greene, 2003).

5.4.3 Findings from 'Prior to Ukraine' Dataset

Model 1 of Table 5.3 reports the base model with control variables. We observe a negative relationship between *Log(VL-Views)* and user endorsement. Counterintuitively, this indicates that comments on highly viewed videos are not getting endorsed. So, we probed it further. First, higher views, i.e., *Log(VL-Views)*, are associated with higher comments (refer to Figure 5.1). One of the popular videos (posted in 2018) has fetched 107,489 comments. It is unlikely that a user will read these 0.1 million comments. Hence, if there are thousands of comments for a video-id, the propensity of a comment (in that video-id) getting endorsed will be very low. Instead, users will read the first few initial comments and move to another video or sign off. We also note that videos posted by pro-migrant left-leaning channels are getting endorsed. Contrarily, the propensity to get likes is negative for videos by right-leaning media. This indicates that users generally endorse pro-migrant videos – if we assume videos posted by leaf-leaning channels are pro-migrants. Next, we note a negative relationship between *Log(VLTS - CLTS)* and user endorsement. This indicates that initial comments get endorsed, but comments afterwards would get lost in hundreds of other comments. Similarly, users might not be interested to read a verbose comment, i.e., high values of *Log(CL-Word Count)* – instead, users will prefer brief comments for the paucity of time. This explains the negative relationship between *Log(CL-Word Count)* and user endorsement. Model 2 incorporates the comment level positive and negative sentiment scores, and model 3 includes the swear word scores. Finally, model 4 reports the full model with all control and explanatory variables. We observe that migrant-related comments with positive sentiments are getting endorsed, and our findings are statistically significant ($p<0.01$). However, users are also endorsing comments with negative sentiments ($p<0.01$). This partially supports both H1A and H1B. Models 3 and 4 indicate that users did not endorse comments with swear words ($p<0.01$). Overall, the findings remain consistent for Tobit models (i.e., models 5 to 8 in Table 5.3) and confirm the robustness of our findings. Overall, results from Table 5.3, i.e., data before the 2022 Ukrainian crisis, neither support the *cultural us assumption* nor *cultural others assumption*. Empirical evidence is mixed at best.

Interestingly, the coefficient of *Negative_Sentiment*, i.e., the endorsement pattern, is negative and significant ($p<0.01$) in model 6, but this coefficient becomes positive and significant ($p<0.01$) in model 8. In other words, the coefficient of *Negative_Sentiment* is not consistent across models 6 and 8. Hence, we need to probe this further. It is worth noting that the coefficients of *Swear_Words* are consistently negative and significant ($p<0.01$) for both models 7 and 8. According to LIWC, input texts with swear words mostly display negative sentiments. So, the coefficient of *Negative_Sentiment* in model 6 captures not only the endorsement pattern for comments with negative sentiments, but also the endorsement pattern of comments with swear words. However, this is not the case in model 8, because model 8 controls the endorsement pattern of comments with swear words through a separate variable, i.e., *Swear_Words*. This explains the difference between the coefficients of *Negative_Sentiment* across models 6 and 8. Hence, based on the findings of model 8, it can be concluded for Dataset 1 that users may endorse a negative comment toward migrants but not comment with swear words.

5.4.4 Findings from 'Ukraine Crisis' Dataset

Similar to Table 5.3, Table 5.4 reports the findings of our proposed hypotheses for dataset 2, i.e., during the 2022 crisis. Overall, the patterns remain the same except for a few variables. For instance, unlike table 5.3, we initially did not find a conclusive relationship between *Log(CL-Word Count)* and user endorsement. We investigate our data carefully and observe a curvilinear relationship. Hence, we introduced the quadratic term and found a statistically significant inverted U-shaped relationship. In other words, neither very brief comments (without any considerable argument) nor verbose comments are getting endorsed. Contrarily, comments with around 100 words are getting endorsed. Next, we find a weakly significant ($p<0.10$) positive coefficient for videos posted by left-leaning channels for our Logit models but no relationship for Tobit models (when we are considering the log transformation of like counts). This pattern was a bit puzzling, and we carefully went through some of the YouTube videos. We find that videos posted by left-leaning channels were primarily sympathetic to the Ukrainian people. However, some of the videos, and comments in response to these videos, also pointed out the discriminatory policies during the 2020 crisis (such as differential treatment received by Africans when they were trying to leave Ukraine) as well as before this crisis. Users did not endorse these comments. Another interesting observation was the coefficients of negative emotions – unlike Table 5.3, it was negative and statistically significant ($p<0.10$) across all the models. This indicates a negative propensity to like migrant-related comments with negative sentiments. Other results were similar to our previous findings. Hence, Table 5.4, i.e., data during the 2022 crisis, *strongly supports the cultural us assumption*.

Table 5.3 User Endorsement Analysis prior to the Ukrainian Crisis

	Logit: M1	Logit: M2	Logit: M3	Logit: M4	Tobit: M5	Tobit: M6	Tobit: M7	Tobit: M8
Log(VL-Views)	-0.389***	-0.390***	-0.385***	-0.387***	-0.157***	-0.158***	-0.154***	-0.155***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Left_Leaning_Chnl	0.047***	0.049***	0.046***	0.047***	0.025***	0.027***	0.025***	0.025***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Right_Leaning_Chnl	-0.300***	-0.297***	-0.297***	-0.296***	-0.147***	-0.145***	-0.145***	-0.144***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(VLTS - CLTS)	-0.115***	-0.117***	-0.117***	-0.118***	-0.108***	-0.109***	-0.109***	-0.109***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(CL-Word Count)	-0.146***	-0.139***	-0.156***	-0.152***	-0.125***	-0.122***	-0.132***	-0.129***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Positive_Sentiment		**0.008*** **		**0.009*** **		**0.004*** **		**0.004*** **
		(0.000)		(0.000)		(0.000)		(0.000)
Negative_Sentiment		-0.002		**0.006*** **		**-0.002*** **		**0.002** **
		(0.373)		(0.004)		(0.015)		(0.076)
Swear_Words			**-0.059*** **	**-0.063*** **			**-0.037*** **	**-0.038*** **
			(0.000)	(0.000)			(0.000)	(0.000)
Constant	2.010***	1.989***	2.025***	1.996***	0.904***	0.894***	0.911***	0.896***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Chi-square	4858.4***	4887.7***	5020.4***	5057.0***	4526.4***	4557.3***	4750.4***	4778.9***

*Note: M1 – Model 1; M2 – Model 2; and so on; M1 to M4 – Logit models with DV - Like_Dummy; M5 to M8 – Tobit models with DV – Log(Like_Counts); N= 110,803; Standard error in parenthesis; *p<0.10, **p<0.05, ***p<0.01 (one-tailed test)*

Table 5.4 User Endorsement Analysis during the Ukrainian Crisis

	Logit: M1	**Logit: M2**	**Logit: M3**	**Logit: M4**	**Tobit: M5**	**Tobit: M6**	**Tobit: M7**	**Tobit: M8**
Log(VL-Views)	-0.498***	-0.496***	-0.498***	-0.497***	-0.228***	-0.227***	-0.228***	-0.227***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Left_Leaning_Chnl	0.040*	0.042*	0.039*	0.042*	-0.005	-0.004	-0.005	-0.004
	(0.186)	(0.157)	(0.190)	(0.166)	(0.691)	(0.772)	(0.692)	(0.755)
Log(VLTS - CLTS)	-0.683***	-0.682***	-0.680***	-0.679***	-0.399***	-0.396***	-0.397***	-0.394***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(CL-Word Count)	0.925***	1.058***	0.950***	1.072***	0.579***	0.669***	0.588***	0.672***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(CL-Word Count)2	-0.161***	-0.192***	-0.170***	-0.199***	-0.120***	-0.141***	-0.124***	-0.143***
	(0.012)	(0.003)	(0.008)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Positive_ Sentiment		**0.018*** **		**0.018*** **		**0.013*** **		**0.013*** **
		(0.000)		(0.000)		(0.000)		(0.000)
Negative_ Sentiment		**-0.018*** **		**-0.013*** **		**-0.009*** **		**-0.007*** **
		(0.000)		(0.005)		(0.000)		(0.001)
Swear_Words			**-0.119*** **	**-0.109*** **			**-0.058*** **	**-0.052*** **
			(0.000)	(0.000)			(0.000)	(0.000)
Constant	1.645***	1.494***	1.645***	1.496***	0.674***	0.562***	0.676***	0.565***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Chi-square	1302.5***	1346.9***	1339.2***	1376.6***	1615.1***	1712.6***	1661.2***	1748.9***

*Note: M1 – Model 1; M2 – Model 2; and so on; M1 to M4 – Logit models with DV - Like_Dummy; M5 to M8 – Tobit models with DV – Log(Like_Counts); N= 21,453; Standard error in parenthesis; *p<0.10, **p<0.05, ***p<0.01 (one-tailed test)*

5.4.5 Limitations and Future Directions

We have carefully examined numerous comments to make sense of our data, and Table 5.5 reports a few representative comments from the corpus. We note that our LIWC-based unsupervised approach has failed to capture the semantic aspects of some comments (such as comments #2 and #10). For instance, though the tone of comment # 10 is negative, it is a pro-migrant comment, and users endorsed the same.

Table 5.5 Representative Comments from our Corpus

#	Sample YouTube Comments	Dataset*	Like(s)	Label**
1	I support this. Immigrants, Canada awaits you! We love you!	1	27	Positive
2	Send them back they're not refugees they are economic migrants here for free money not willing to work	1	1958	Positive
3	This is absolutely outrageous, ... spending our taxes on foreigners who have no right to be here... if the government persists in doing nothing about these migrants .. I'm refusing to vote at the next election...	1	0	Negative
4	Imagine a world where putting your own country's people first is racist Only the western world. No one wants to talk about middle Eastern countries outright saying no to "refugees," it's only racist ...	1	416	Negative
5	That's the point jackass, they aren't refugees, they are economic migrants. Stop committing cultural suicide you f***ing douc**bag!	1	0	Abusive
6	You f***ing idiot. No one is buying that shit anymore. "It's just immigrants even though if this continues you will be a minority in your own country in 2066. It's just immigrants man."	1	12	Abusive
7	Thank you, Poland, it's amazing how ordinary polish people accepted refugees in their own homes, Poland is providing place to stay, medical help, but Poland and other neighboring countries shall be supported by wealthy western countries.	2	203	Positive
8	These refugees should blame their president for insisting on joining NATO and forcing Putin into a corner.	2	0	Negative
9	Why the f*** would Poland take Muslim refugees....	2	0	Abusive
10	Disgraceful, there is no first-class refugee and second-class refugee when people are running for their life.	2	706	Negative

*Note: Dataset 1: before 2022 refugee crisis; Dataset 2: during 2022 refugee crisis; *Labels according to LIWC scores*

Overall, we find that users endorse positive sentiments (comments #1 and #7) and don't endorse comments with abusive words (comments #5 and #9). However, we also note that the instances of antipathy in dataset 1 (comments #2, #3, #4, #5, and #6) are significantly higher than in dataset 2 (comment #9). Some of the racial comments of Dataset 1 also got endorsed (comments #2, #4, and #6). This explains the positive coefficient of negative sentiments for models 4 and 8 in Table 5.3. We also note racial comments to non-European citizens, such as Islamophobia or discrimination toward Africans, during the Ukrainian crisis (comment #9). Some users also pointed out these discriminations and got endorsed by other users (like comment #10). To make a contextual interpretation of our findings, we also need to consider the YouTube penetration, which is mostly more than 90% in advanced nations, such as European countries or the USA or Canada, in 2022, but YouTube penetration is only 37.7% globally (Ceci, 2022). So, we may assume that a significant portion of YouTube users is primarily from advanced nations. Thus, our findings, based on LIWC-based syntactic analysis, implicitly lend support to the racism theory of *cultural us* vis-à-vis *cultural others* in the context of migrant-related discourse. Future studies need to employ a semantic approach to have a fine-grained understanding of migrant-related discourse.

5.5 Conclusion

Racism theory suggests that the host nation may consider migrants as cultural others, and it triggers antipathy and hate speech in extreme cases. The 2022 Ukrainian refugee crisis allows us to test the '*cultural us*' versus '*cultural others*' assumptions. This is one of the initial studies, except a few like Leasure et al. (2022), that explored the 2022 Ukrainian refugee crisis. We collected YouTube comments in two phases (i.e., prior to the 2022 crisis and during the 2022 crisis). We probed the relationship between sentiments of migrant-related discourse and user endorsement patterns. Our findings suggest that users mainly endorse positive comments. However, the user endorsement pattern for comments with negative sentiments across two datasets implicitly supports the *cultural other* hypothesis. We found that users did not endorse negative sentiments during the 2022 crisis, but the pattern was not identical in our earlier dataset. Probably, YouTube users from advanced nations, especially European countries, considered Ukrainian refugees as *cultural us*, but they may have considered earlier refugees as *cultural others*. It is worth noting that this study explored societal opinion about migrants, but prior studies observe that the journey from 'struggle to settlement' of refugees is the same irrespective of their home country (e.g., Chapters 5 and 7). Hence, migrants and refugees, irrespective of who they are or from where they are, need societal support and sympathy not only from the citizens of host nations but also from the netizens on social media platforms.

Contrary to the conventional belief, it is heartening to note that though there are instances of hate speech toward migrants, social media users generally do not endorse comments with abusive or swear words (though there were a few exceptions, as we noted in Table 5.5). We also note that our unsupervised approach has a few limitations in capturing the semantic meaning (Cambria et al., 2020). However, despite these limitations, this approach allowed us to perform a large-scale user interaction analysis to understand the overall societal sentiments. Our study is one of the initial attempts, if not the first, in this direction and contributes to the applied NLP domain of migration literature.

Chapter 6: Settling Down of Migrants in Host Nations

6.1 Voices of Migrants and Refugees on Twitter

6.1.1 The Struggle of Migrants to Settle Down

Prior studies from the social science domain (e.g., Atallah & Mahdi, 2017; Crawley, 2017; Dustmann et al., 2017; Mitsilegas, 2017) explored the challenges faced by migrants and refugees, such as health issues (Brown-Bowers et al., 2015; Qutob, 2018), barriers in social integration process (Alencar, 2018), and effects of refugee-related policies and law (Mitsilegas, 2017). Migrants and refugees struggle due to anxiety (Henkelmann et al., 2020), discrimination (Laban et al., 2005), hatred (Hrdina, 2016), and other stressful events before settling down in the host country. For instance, Figure 6.1 reports the total number of pending asylum applications in the host country (Source: UNHCR Refugee Population Statistics Database, 2022, Based on the latest data available on 25 June 2022). At the end of 2021, globally, 4.6 million asylum applications are pending, and 1.3 million of these are for the USA. Next, Figure 6.2 reports the Asylum or Refugee Policy Index in 2010, where 0 is the least restrictive, and 1 is the most restrictive. Figure 6.2 indicates that advanced nations have slightly restrictive policies – especially when the global average is 0.3 index.

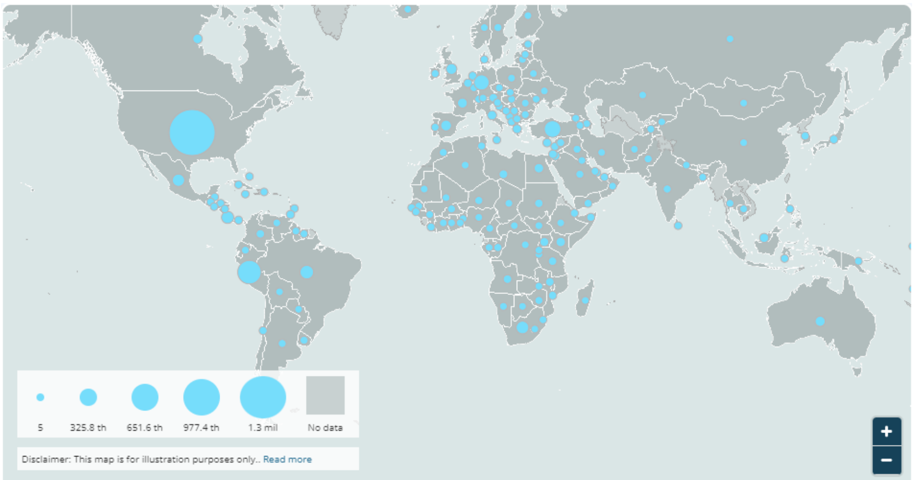


Figure 6.1 Asylum Seekers in host country, end of 2021[38]

Subsequently, Figure 6.3 reports the number of refugees who were resettled in 2021, by country of destination (Source: IMPIC, 2016; Based on the latest data available on 26 March 2019). Globally, only 57.4 thousand refugees got settled, and the maximum number of settlements were in countries like Canada (20.4 thousand), the USA (13.7 thousand),

[38] Available at https://www.migrationdataportal.org/international-data?i=asyl_host&t=2021, Accessed on July 1, 2022

Sweden (6.7 thousand), Norway (3.6 thousand), Australia (3.3 thousand) and so on. Juxtaposing the insights from Figures 6.1 and 6.3 would offer us the required reality check. For instance, in 2021, 1.3 million asylum applications were pending, and only 13.7 thousand got settled in the USA. Comparatively, the situation was significantly better in Canada – 63.1 thousand pending asylum applications and 20.4 thousand got settled. However, this pattern is intuitive because the Asylum/Refugee Policy Index score was lower for Canada (0.2 index) than the USA (0.4 index). Figures 6.1, 6.2, and 6.3 aptly capture the struggles faced by refugees and migrants before settling down in their respective host nations.

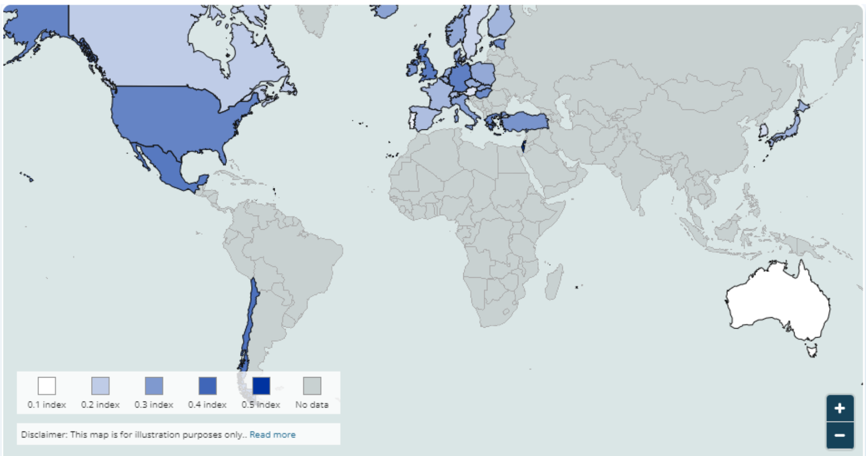


Figure 6.2 Asylum/Refugee Policy Index in 2010[39]

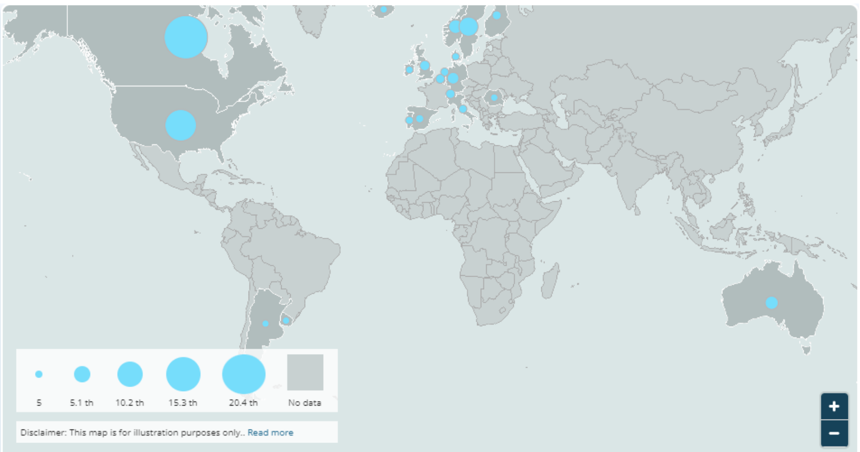


Figure 6.3 Number of refugees who were resettled in 2021[40]

As aforementioned in Chapter 2, the AI-based stream of literature mostly explored public opinions on social media platforms about refugees and migrants (Armstrong et al.,

---

[39] Available at https://www.migrationdataportal.org/international-data?i=impic_asyl&t=2010, Accessed on July 1, 2022

[40] Available at https://www.migrationdataportal.org/international-data?i=reset_destin&t=2021, Accessed on July 1, 2022

2021; Gallego et al., 2017; Kreis, 2017; Urchs et al., 2019). However, to the best of our knowledge, none of the prior studies, except a few like Palakodety et al. (2020), has exclusively examined the voices and concerns of refugees and migrants. Also, we note that on the methodology front, social science literature has considered focused group discussion (Laban et al., 2005; Schweitzer et al., 2002), questionnaire-based survey (Naja et al., 2016; Nesterko et al., 2020), and individual interviews (Henkelmann et al., 2020). These studies generally employed structured or semi-structured questionnaires to investigate the concerns of migrants and refugees. Contrarily, the Twitter platform allows users to share their initial struggles without probing. Hence, Twitter deliberations potentially allow us to explore the faint and unheard voices of voiceless migrants. Accordingly, Section 6.1.2 probes their journey from *Struggle to Settlement*.

6.1.2 Psychological Stress and Resilience of Migrants

The online behavior of social media users can be an indicator of their mental well-being. For instance, Burke & Kraut (2013) explored the online behavior of Facebook users after they experience job loss. This study reveals that talking with close friends helps them to control their stress level and provide social support; however, distant friends rarely provide this support. The covid-19 crisis is one of the most stressful global events in the history of mankind. Suh et al. (2021) explored the web query pattern, on Microsoft's Bing platform, during this pandemic to investigate the shift in human needs across societal, economic, and psychosocial dimensions. Basic human needs went up drastically during this crisis, whereas higher-level aspirational needs experienced a downfall. Additionally, temporal variations in human needs indicate instances of psychological and economic resilience (Fletcher & Sarkar, 2013; Suh et al., 2021). Similarly, prior studies noted that childbirth could lead to postpartum depression in new mothers. De Choudhury et al. (2013) probed the changes in emotion and behavior during the postpartum period and argued that postpartum depression of new mothers could be predicted by analyzing their "social engagement, emotion, ego-network, and linguistic styles" (p. 3267) on the Twitter platform around childbirth.

A similar study in the Canadian context, by Brown-Bowers et al. (2015), has found that the propensity of having postpartum depression of refugee or asylum-seeking women is five times more than Canadian-born women. Not only postpartum depression, but also migrants experience loneliness due to separation from their family members, face ethnic conflicts in the new country, face difficulties in getting appropriate accommodation, struggle to learn a new language, and face communication-related difficulties in the host country (Laban et al., 2005; Schick et al., 2016; Chapter 4). These struggles to settle down create psychological stress, and subsequently, mental disorders (Laban et al., 2005). Accordingly, prior studies have found evidence of high mental stress among refugees in the context of Australia, Canada, Denmark, Germany, Iraq, Netherlands, Switzerland, Syria, and in Middle Eastern countries (Brown-Bowers et al., 2015; Laban et al., 2005; Montgomery, 2010; Naja et al., 2016; Nesterko et al., 2020; Schick et al., 2016; Schweitzer et al., 2002). A meta-analysis of anxiety, depression, and post-traumatic stress disorder (PTSD) in refugees reveals that many refugees have self-reported anxiety, depression, and PTSD (Henkelmann et al., 2020). Schick et al. (2016) argue that exposure to traumatic experiences leads to substantial psychological impairment, and this study has also observed a strong correlation between the lack of social integration and mental health problems. Hence, Schick et al. (2016) suggest fostering the social integration process for these traumatized refugees.

Alarmingly, the prevalence of mental disorders of refugees from politically disturbed countries is substantially higher in comparison to non-refugee populations, and this pattern is similar for both child/adolescent refugees as well as adult refugees (Henkelmann et al., 2020; Naja et al., 2016). For instance, Montgomery (2010) found a strong relationship between mental health, especially in young refugees, and traumatic events, such as arrest and imprisonment of family members, loss and separation from the family, street shootings, war, and bombing in the Middle East. The struggles of migrants and refugees to settle down are broadly similar, irrespective of their origin and host countries.

Thankfully, the traumatic experiences of refugees and subsequent mental stress are not the end of the story. The resilience of migrants and refugees helps them to overcome these adverse situations (i.e., struggles) and keep them optimistic in achieving their target (i.e., settling down in the host countries). For example, Kira et al. (2014) have observed that refugees from Arabs have displayed "resilience in successfully adapting to life in the USA" (p. 183). Resilience can be defined as "the personal qualities that enables one to thrive in the face of adversity" (Connor & Davidson, 2003, p. 76). The resilience to come out of the adverse situation can be captured through qualities such as adaptability to a changed context, confidence to face a new challenge, ability to handle stress, don't give up attitude, ability to handle unpleasant feelings, and tenacity to attain the desired goal (Connor & Davidson, 2003; Fletcher & Sarkar, 2013). In brief, resilience enhances the propensity of successful adaptation and settling down in the host country after the initial phase of not-so-pleasant hostile experiences and adverse situations (Montgomery, 2010). Overall, the extant literature reveals that social media data can be insightful for probing mental well-being (from the information science domain), and the level of mental stress and resilience is high for migrants and refugees (from the social science domain). However, none of the prior studies, to the best of our knowledge, probed the *first-person tweets from migrants and refugees* to understand their voices.

6.1.3 Data

As we noted in Chapter 2, Twitter data was widely used in migration-related studies (Kreis, 2017; Özerim and Tolay, 2020; Urchs et al., 2019). Hence, we considered migrant and refugee-related tweets from May 2020 to March 2021. For the initial crawling purpose, we have considered the following keywords: '*asylums*', '*migrants*', '*refugee*', '*immigrants*' etc. Our final corpus comprises 0.15 million tweets after discarding duplicate tweet-ids and tweets with similar text content from the initial corpus. A careful introspection reveals that a significant portion of our corpus comprises tweets by non-refugee or non-migrant users, but we need to probe the refugee perspectives – their struggles, concerns, experiences, and resilience. We note that only a minuscule portion of our corpus captures these issues.

Table 6.1 Sample Tweets for Analysis

Sl.#	Representative Tweets	Themes
1.	Yes, it is so sad. *I am a refugee* with family about 7 years with horrible conditions. We are fighting days, nights. High level of depression, stress, uncertain future and mental; psychological problems cased (sic) him end his life. *#UNHCR* please do something.	Struggles of Migrants & Refugees (168 tweets)
2.	1178 days in a NZ jail. No crime, no charge, just for *being a refugee*.	
3.	*I am an immigrant* every year I pay my taxes, and nobody gives me anything.	Settlement in the host country (156 tweets)
4.	A century ago, my great-grandparents migrated for better economic opportunities. A few years later their daughter -my grandmother- left their new home country for love and family. I'm a proud descendant of migrants and today, *I am a migrant* myself.	
5.	*I'm an immigrant* rights activist, so you're wrong on every count. Every time you open your mouth, you're wrong. Did you get your right-wing feelings hurt? Poor thing. *#JoeBidenInauguration*	Generic Views of Migrants & Refugees (147 tweets)
6.	*I am a refugee* from Fox News who tuned in last month and now watch Newsmax. My bro and neighbor recommended we try and now we love your channel!	
7.	I demand an explanation from ____ why it takes so long to deport foreign criminals and illegal immigrants	Public views about Migrants & Refugees (152 tweets)
8.	Refugees deserve respect as much as the next person they are part of the solution you can be too *(sic)* listen to them join them support them on *#worldrefugeeday* and beyond	

Table 6.1 reveals that tweets from migrants and refugees occasionally mention phrases such as *"I'm a/an refugee/migrant/immigrant*", "*as a refugee/migrant I"*, and *"being a refugee/migrant"* (e.g., tweets #1 to 6). Some of the tweets from second-generation refugees and migrants also use phrases such as *"father/mother was a refugee/migrant"* or *"my parents were refugees/migrants"* (e.g., tweet # 4). Using the above phrases as search themes, we find that less than 3000 tweets are first-person tweets from migrants and refugees, i.e., a mere 2% of our corpus. This scant presence of first-person tweets elucidates the challenges and need to analyze the faint voices of migrants and refugees on the Twitter platform. Table 6.1 reports a few representative tweets. Our tweets can be broadly divided into two classes: migrant and refugee voices and '*public views'* (i.e., tweets by non-refugee or non-migrant users). We find public opinions are either supportive of refugees (e.g., tweet # 8) or abusive to them (e.g., tweet # 7).

Next, we have divided refugee voices into three classes: '*struggles'* (i.e., tweets sharing their initial challenges or mental stress due to traumatic experiences), '*settlement*' (i.e., successful adaptation in the host country or recovery through resilience), and '*generic views'* (i.e., tweets about larger political or social issues, mostly related to refugee and migrant issues, but not about their personal experiences). We note that refugees use the Twitter platform to share their initial struggles and mental stress (e.g., tweet #1), such as traumatic detention experiences (e.g., tweet # 2) after arriving at the host country. These tweets capture the fear, anxiety, and challenges associated with accessing primary education and essential health facilities. Our next theme tells how their resilience helps some of them to overcome adverse situations to settle down in the host country (e.g., tweets # 3 & 4). Interestingly, a few tweets are generic, and these tweets are deliberating issues that are not specific to their personal journey (e.g., tweets # 5 & 6). We have analyzed the linguistic content of tweets and manually annotated 623 tweets into the above four categories for the final analysis. Some of the tweets have multiple themes within it, but we considered the most salient theme/view for classifying these tweets.

6.1.4 Methodology

As we pointed out earlier, Bi-LSTM models can efficiently capture long-term dependencies of very lengthy sequences, whereas the CNN models can capture the spatial information of words in low-dimensional vectors, and convolutional filters are applied to the word embedding matrix (Kim, 2014). We have considered Bi-LSTM and CNN models with pre-trained embedding *fasttext* (i.e., wiki-news300d-1M of 16B tokens) for training. We have also considered transformer-based models, i.e., BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019), that evaluate the context of a given word from both left and right directions. We have considered pre-trained implementation from the HuggingFace library for our BERT and RoBERTa models (Wolf et al., 2019). We have used 16 and 32 BSs and 'adam' optimizer for all these models.

Table 6.2 Accuracies in Identifying Refugee and Migrant Concerns

BS	Bi-LSTM	CNN	DR	BERT	RoBERTa	LR
16	**61.61**	**67.86**	0.4	**75.00**	75.00	2e-5
16	60.71	65.18	0.5	70.54	**75.89**	3e-5
32	56.25	66.07	0.4	68.75	74.11	2e-5
32	54.89	64.29	0.5	71.43	72.32	3e-5

Table 6.2 reports the classification accuracies. Accuracies of CNN-based models are higher (in the range of 64% to 67%) than Bi-LSTM models (55% to 60%) for identical hyperparameters. The best F1-Score for CNN is 67.86%. As expected, BERT and RoBERTa models have outperformed Bi-LSTM and CNN models. Broadly, the accuracies of RoBERTa models are higher than those of BERT models. The best-performing RoBERTa model has reported an F1-Score of 75.89% (for BS 16 and LR 3e-5). Table 6.2 strongly indicates that our AI-based models are efficient in identifying and analyzing the concerns of migrants and refugees.

6.1.5 Discussions and Future Scope of Work

Concerns of refugees were probed by using focused group discussion, questionnaire-based analysis, or individual interviews (Gallego et al., 2017; Laban et al., 2005; Nesterko et al., 2020; Schweitzer et al., 2002). However, AI-based studies have rarely probed the voices and concerns of refugees and migrants on social media platforms. Hence, we have tried to address this gap. However, our study has a few limitations that might offer exciting avenues for future research. *First,* a further fine-grained classification of these themes might offer a more nuanced view. For example, we have clubbed a wide range of challenges and constraints as struggles. It is worth noting that some of them are struggling due to their inability to access essential medical services. In contrast, others are worried about the future of their kids in the asylum because they are not getting primary education. Some don't have legal documents, and some are traumatized due to hostile behaviors or lack of a proper social integration process. These struggles are not uniform, and their needs are different. So, a fine-grained classification will help organizations like UNCHR to take appropriate policy measures. It is worth noting that we have collected data for a year, and we could gather around 3000 tweets. Intuitively, it can be argued that refugees in the asylums might not have the facility or internet connectivity to tweet about their sufferings. Hence, understanding their voices and concerns would be a challenging task.

Second, our study has considered refugees and migrants synonymously. It is quite possible that xenophobic social media might be indifferent, and they can assume all refugees and migrants are illegally entering their country. As we pointed out in Chapter 1, UNCHR policies have clearly pointed out that these two categories are not the same. The mental stress of refugees coming from politically disturbed countries will be much higher than a migrant. Our corpus reveals that the struggles of refugees are associated more with basic human needs like food or shelter. On the contrary, the struggles of a migrant are primarily associated with a lack of proper social integration or discriminating behaviors from a specific section of the host country. Similarly, we also note that most support activities by charitable organizations are for distressed refugees. We also note that a significant portion of the success stories of settling down in the host country mostly comes from migrants. Hence, future work should tease out the voices of refugees and migrants. A fine-grained analysis will enrich the contextual understanding of migrants and refugees.

6.2 Social Media for Matching Recruiters and Migrants?

6.2.1 Migrants, Unemployment and Social media

Organization for Economic Co-operation and Development (OECD)[41] data indicates that "immigrant workers are affected to a greater extent by unemployment than native-born workers in European countries that have traditionally received migrants" (¶ 2). Section 6.1 also suggests that one of the crucial struggles for migrants is finding a suitable job. OECD has defined the foreign-born unemployment rate[42] "as the share of unemployed foreign-born persons aged 15-64 in the foreign-born labor force (the sum of employed and unemployed foreign-born) of that same age. Unemployed people consist of those persons who report that they are without work during the reference week, are available for work, and have taken active steps to find work during the four weeks preceding the interview" (¶ 2). Figure 6.4 graphically reports the same for major European countries. Migrants consider themselves settled when they find a suitable job. For instance, a tweet from Table 6.1 captures this aspect implicitly: *I am an immigrant every year I pay my taxes, and nobody gives me anything.*

[41] Available at https://www.oecd-ilibrary.org/social-issues-migration-health/foreign-born-unemployment/indicator/english_ba5d2ce0-en, Accessed on July 1, 2022
[42] ibid

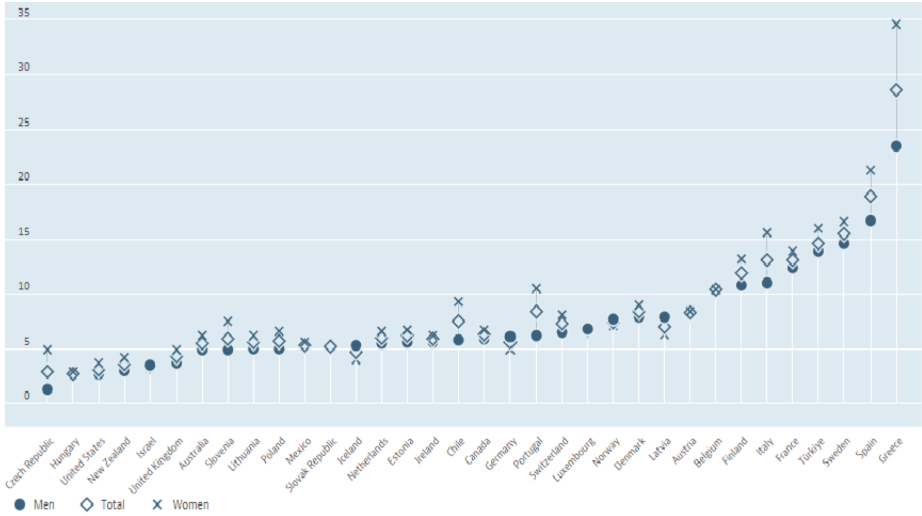


Figure 6.4 Foreign-born unemployment[43]

Recent developments in social media mining and AI have impacted the job searching. Job portals have reduced the information gap between recruiters and job seekers. Recruiters heavily use these job portals for their recruitment process. Similarly, job seekers are also exploring these job portals to search for a suitable job. Hence, AI-based research is trying to recommend the appropriate candidate to a recruiter or recommending the appropriate vacancy to a jobseeker. Prior studies have conceptualized this task as a *person-job fit problem* (Zhu et al., 2018). Existing literature has mainly considered commercial job portals as their data source. However, job-related data from these portals are proprietary data and not publicly available. Commercial job portals might not be a viable solution for economically weaker migrants. Hence, we are exploring whether publicly available social media data, such as Twitter, can be a viable substitute for commercial job portals. For our analysis, we have formulated our task as a simplistic person-job fit problem – not considering whether the job seeker is a migrant or not. Basically, we aim to connect job seekers and recruiters on the Twitter platform. However, Tables 6.4 and 6.5 clearly indicate that the best matching between the jobseekers and recruiters mostly happens across borders. Hence, extrapolating this generic framework in the context of migrants will not require much adaptation.

We have crawled job-related tweets from both recruiters (T_R) and jobseekers (T_J), but we need labeled job data for model building. Hence, we have annotated and matched tweet-pairs (t_{r1}:t_{j1}, t_{r2}:t_{j2}, ... t_{rN}:t_{jN}) from recruiters and jobseekers. We have used this annotated data for the training and evaluation of our proposed Siamese architecture.

43 OECD (2022), Foreign-born unemployment (indicator). doi: 10.1787/ba5d2ce0-en, Available at https://www.oecd-ilibrary.org/social-issues-migration-health/foreign-born-unemployment/indicator/english_ba5d2ce0-en, Accessed on July 01, 2022

Figure 6.5 graphically represents our overall research framework. We have restricted our analysis to data science, game developers, software engineers, and web developer jobs. However, our proposed framework can be extrapolated to other domains, ranging from accounting to advertising jobs or prosecutors to physiologists. On methodological fronts, we propose a semi-supervised Siamese architecture-based framework, and employed CNN, LSTM, Bi-LSTM, and Bi-LSTM with attention. The core objective is to probe – *whether publicly available Twitter data can be a viable substitute for commercial job portals in connecting migrants and recruiters.*

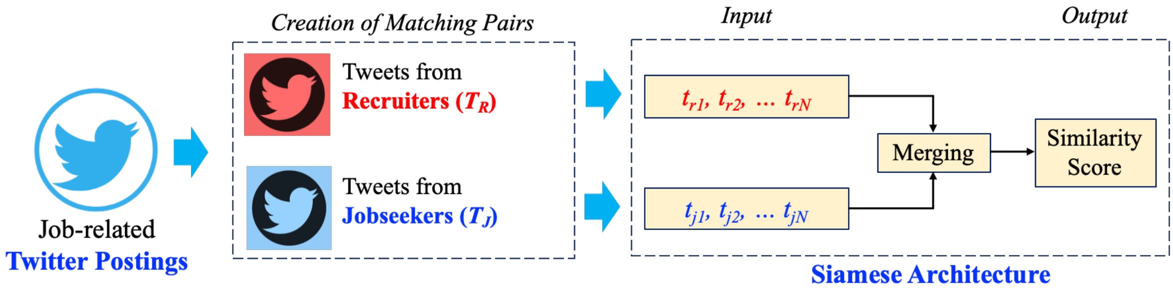


Figure 6.5 Framework for Job Recommendation

6.2.2 Prior studies on Person-Job Fit

Job-related research has become popular in the last few years. This stream of research mostly focused on job recommendation or person-job fit (Zhu et al., 2018). In addition to person-job fit, prior research in this domain has also probed job mobility (Meng et al., 2019), salary benchmarking (Meng et al., 2018), privacy issues (Kenthapadi & Tran, 2018), and personalized question recommendations for an interview (Qin et al., 2019). From the migrant's unemployment perspective, we focus on the person-job fit problem. An efficient job recommendation system needs to connect two different entities: recruiters and jobseekers (or migrants).

Prior research considered two separate sets of data as input variables: one input variable is the candidate profiles, and the other input variable is the job description from the recruiters. However, automated job recommendation is a challenging task. For example, the required skillsets for 'data science' can be very similar to the skillsets for 'data engineering', 'data analysis', or 'machine learning'. A simple rule-based approach might not be the most efficient approach to address this named entity recognition problem. For the sake of brevity, Table 6.3 reports the existing literature in tabular format.

Table 6.3 Prior Studies on Job Recommendation

Author	Data Source	Methods/Findings
Shalaby et al. (2017)	CareerBuilder	Proposed job recommendation by addressing the short-lived nature of jobs and the rapid rate at which new users and jobs enter the system
Yang et al. (2017)	CareerBuilder	Employed Statistical Relational Learning for developing their job recommendation system and ensured low inappropriate job recommendation
Dave et al. (2018)	CareerBuilder	Proposed a representation learning model that considers information from three networks (job transition network, job-skill network, and skill co-occurrence network) for job recommendation
Geyik et al. (2018)	LinkedIn	Explains the architecture of "LinkedIn Recruiter product, which enables recruiters to search for relevant candidates and obtain candidate recommendations for their job postings" (p. 1353)
Ramanath et al. (2018)	LinkedIn	Considered both deep learning models as well as representation learning approaches for talent search systems at LinkedIn
Zhu et al. (2018)	A Chinese company	Employed CNN for person-job fit. Also, it identifies which all items of the job requirements the person can satisfy
Liu et al. (2019)	CareerBuilder	Proposed a vector representation for both job postings and resumes and considered three information graphs (job-job, skill-skill, job-skill)
Meyer et al. (2019)	Indeed	Content analysis of U.S. healthcare data scientist job postings to understand the job requirements
Ozcaglar et al. (2019)	LinkedIn	Developed an entity-personalized talent search model by combining generalized linear mixed models and gradient-boosted decision tree models

Our literature review reveals that prior studies have employed a diverse range of methodology that ranges from simple content analysis (Meyer, 2019) to CNN-based approaches (Zhu et al., 2018), from statistical relational learning (Yang et al., 2017) to representation learning (Dave et al., 2018). However, prior studies mostly considered datasets that were either proprietary or owned by corporates like CareerBuilder or LinkedIn. Some of these studies are the outcome of their in-house research. Commercial job portals have two significant shortcomings. First, job-related data in these job portals are proprietary; thus, accessing the data through crawling can have legal implications. Second, most of these portals charge their users. So, big recruiters or high-end jobseekers can afford their customized service, but small organizations and not-so-rich job seekers, especially migrants, might not be able to afford it. Thus, we probe whether Twitter can address this disparity.

To the best of our knowledge, none of the prior studies has explored the Twitter platform for job recommendation. Hence, we attempt to address this research gap by exploring the feasibility of social media data for the person-job fit problem. In short, we are trying to map a tweet from a jobseeker with an appropriate and relevant tweet from a recruiter in the domain of high-end computer science jobs. The following section elaborates on why we have selected social media platforms as our data source.

6.2.3 Data: Why Twitter?

We find that social media users discuss job-related issues on the Twitter Platform. Most companies, irrespective of their size, are using the Twitter platform to promote their brand and reach customers. These companies also use the Twitter platform for sharing vacancies, new jobs, and recruitment plans. Thus, Twitter is becoming a popular communication channel not only for big organizations but also for small organizations, who cannot afford the customized service of commercial job portals to reach the labor market. Moreover, the younger generations are highly active on the Twitter platform. Table 6.4 reports some sample job-related tweets and various job attributes in those tweets. Our preliminary analysis of the linguistic contents of job-related tweets reveals that most job-related tweets clearly mention the expectations. Many of these posts mention - What is the overall scope of a particular job? Which location? Required skill sets? However, it is essential to note that all job-related tweets are not so informative. Additionally, we find some commercial job portals are very focused. For example, one job portal might be popular for only finance-related jobs in one particular location/country. As a result, others might not opt for that particular portal. However, this is not a problem for the Twitter platform. Twitter data comes from all over the world. Hence, using Twitter Search API, we have extracted job-related tweets. We have considered a set of crucial keywords for the crawling purpose, such as 'job', 'vacancy', 'hiring', and 'employment'. We have extracted 0.76 million tweets during November and December 2019.

This initial corpus has resulted in a diverse set of job-related tweets that range from architectural to accounting jobs. For the analysis purpose, we have focused on computer science jobs where the authors have the required expertise to correctly annotate and match the expectations of recruiters with the experience of the job seekers. We have identified four prominent types of computer science jobs: data science, game development, software engineering, and web development. We have manually annotated 704 unique tweets (52% of them by jobseekers and 48% of them by recruiters). Next, we matched a recruiter's tweet with an appropriate jobseeker's tweet and created 3980 recruiter-jobseeker tweet pairs (this includes 38% correctly and 62% wrongly matched pairs) for training purposes. It is worth noting that while every tweet pair is unique, every tweet within the tweet pair is not. Table 6.5 has reported a few correctly (C) and wrongly (W) matched pairs that we used for modeling purposes.

Table 6.4 Job-related Tweets from Recruiters

Tweets from Job Recruiters	Job Attributes
With SAS analysis experience, an opportunity to join a lovely agency in SW London in this 9 – 12-month Mat Cover Data Analyst role. #marketingjobs #newjob #dataanalyst #analytics #marketinganalyst #SAS #SASprogramme	Experience, Location, Tenure
Are you interested in designing #fullstack #code for web applications in the financial industry? Do you have experience using #Angular, #typescript, #SQS, #Nodejs or #Oauth? Click here #careers #WisconsinJobs #JavaDevelopment #Engineer #JavaScript #HTML5	Experience, Location, Role
iOS Developer Johannesburg, Gauteng, South Africa We are looking for an iOS developer responsible for the development and maintenance of applications aimed at a range of iOS devices including mobile phones and tablet... #jobs #recruiting #careers	Location, Role, Scope
Charles Taylor PLC are now recruiting Senior #Analyst .NET #Developers to join their business to build new innovative systems for insurance industry and support existing #IT applications. Location London. Competitive salary + package. Apply #jobs	Location, Role, Scope, Salary

6.2.4 Methodology: Siamese Architecture

Following prior Twitter-based studies, we have preprocessed our corpus. We have followed the standard steps such as tokenization, word normalization, and lowercasing of all words. We have also removed URLs, email-ids, and user handles and replaced these URLs, email-ids, and user handles with blank space. Next, we have employed Siamese architecture that compares the semantic meaning of two different but similar types of text. In other words, Siamese architecture explores the relationship between two different texts based on their semantic meaning (Marco et al., 2014). This is commonly known as the text pair comparison. Hence, in our research context, this text pair comparison approach measures the semantic similarity of a tweet pair to determine whether one tweet is closer to another or not, i.e., comparing the semantic similarity between the tweets from recruiters with the tweets from jobseekers for developing an efficient job recommendation framework. Existing literature has obtained state-of-the-art results by using CNN (He et al., 2015) and RNN (Mueller & Thyagarajan, 2016; Neculoiu et al., 2016) for the above sentence similarity task. Hence, we have followed a similar approach and considered four models for our analysis: CNN, LSTM, Bi-LSTM, and attention-based Bi-LSTM for text-pair comparison.

Table 6.5 Sample Training Data for Analysis

Tweets from Recruiter	Tweets from Jobseeker	Label
We're hiring a new web developer! If you or someone you know might be a good fit, take a look at the job posting right here	I'm a graphics designer, database administrator, web developer and designer, android developer. I'm looking to collaborate or work with anyone on any project. Please help	**C**
Cognizant is looking for teammates like you. See our latest #IT job openings, including "Data Analyst", via the link in our bio. #Lynnwood, WA	are you looking web designer and developer? I'm a professional web designer and developer with 2 years of experience working with international clients and agency's.	**W**
I am looking for a fantastic and professional web designer. Focused on online shopping, subscriptions and creativity. PLEASE let me know your favorite recommendations. #webdesigner	I am WordPress web developer I have 3years experience. looking for a working opportunity with agency or team.. #wordpress #hire #WebsiteDesign #webagency #agency	**C**
New job opening! We are looking for a full-time (permanent contract) game designer to join our core team More info on our website, feel free to check it out! #swissgames #gamedev	Hi, I saw that you are looking for expert web designer / developer, Well I can assist you as I have designed / developed 100+ website for various clients with different	**W**

Following prior studies, we have created two identical sub-networks that read the corpus and generate a fixed representation. In other words, we have considered two identical sub-networks for the tweets from recruiters and jobseekers. Each sub-network reads the tweets and produces its vector representation for the next layer. Both subnetworks share the same weight for comparing a tweet pair (one from the recruiter and the other one from the jobseeker) in the same vector space. We have considered 80% of our corpus for training and the remaining 20% of the unexposed corpus for testing using stratified sampling for our analysis. As we mentioned earlier, we trained, validated, and tested by using our manually annotated tweet-pairs.

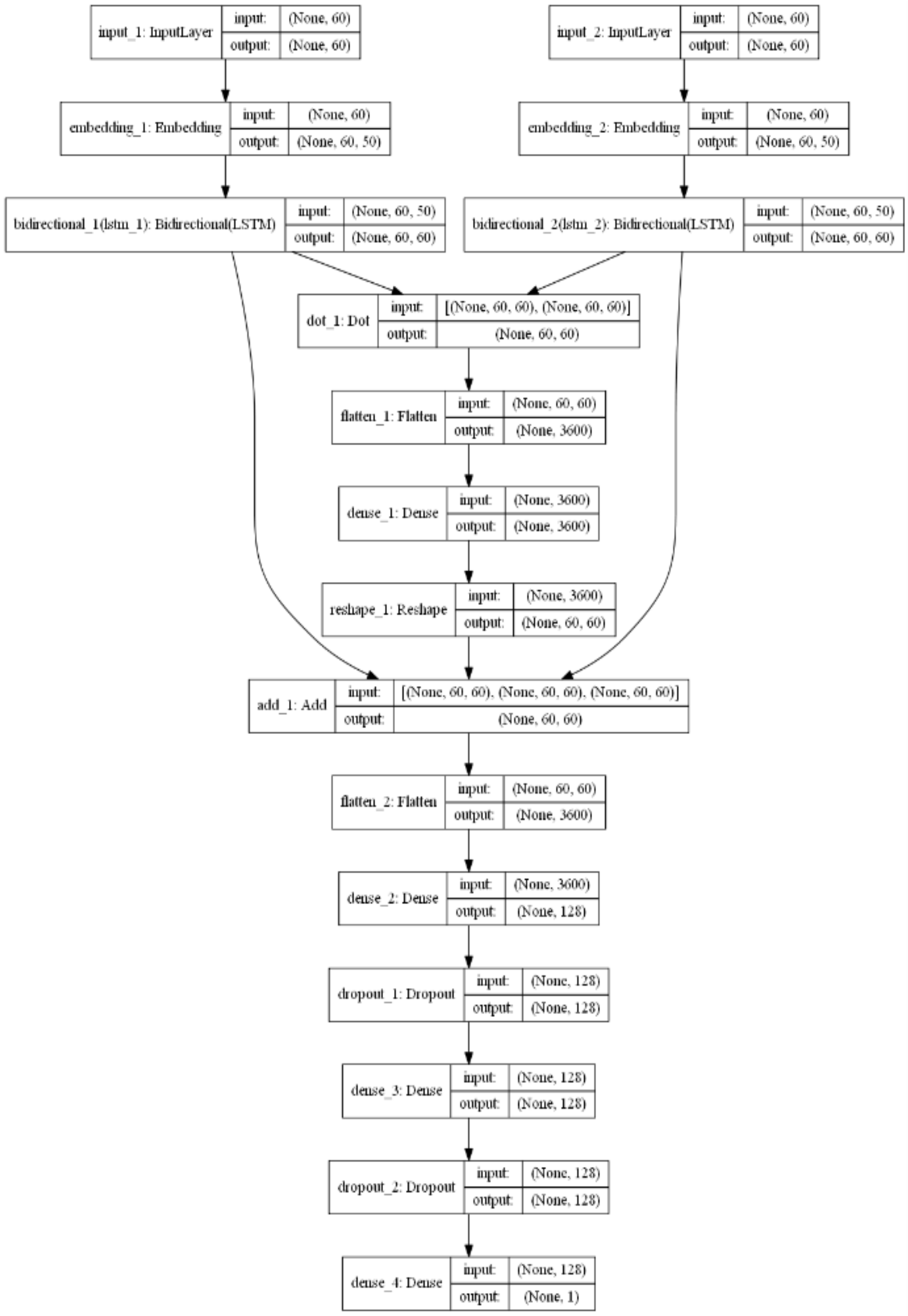


Figure 6.6 Architecture of Siamese Bi-LSTM with Attention

We have considered different pre-trained and publicly available word embeddings i.e., GloVe (Pennington et al., 2014). In the first step, we represent our tweets in a low-dimensional distributed representation, i.e., word embeddings. Specifically, we use four pre-trained word embeddings as follows: 6B50d (50-dimensional Glove embeddings based on Wikipedia 2014 & Gigaword 5 with 6 billion tokens), 6B100d (100-dimensional Glove embeddings based on Wikipedia 2014 & Gigaword 5 with 6 billion tokens), 27B50d (50-dimensional GloVe embeddings based on the Twitter corpus with 27 billion tokens), 27B100d (100-dimensional GloVe embeddings based on the Twitter corpus with 27 billion tokens) (Pennington et al., 2014). We have considered these word embeddings to ensure the robustness of our findings.

First, we use a Siamese CNN model to analyze the contextual similarity of our tweet pairs. The Siamese structure with two identical sub-networks processes the sentences (or tweets) parallelly with identical weights for each layer. Next, our fully connected layers compute the similarity score between two tweets. We have used an Ecludian distance to measure the similarity. We have used the contrastive loss as a loss function and the Adam method to optimize our model's parameters. The last layer in the Siamese architecture decides whether the tweets from the recruiter and the jobseeker (or migrants) match correctly or not.

Table 6.6 Accuracies with Glove Embedding

Model	6B50d	6B100d	27B50d	27B100d
CNN	0.8101	0.8341	0.6202	0.7837
LSTM	0.8832	0.9774	0.6181	0.8065
BI-LSTM	0.9008	0.9422	0.6759	0.9736
Bi-LSTM + Attn.	0.9497	0.9749	0.6093	**0.9661**

Second, we consider the LSTM model that incorporates a gating mechanism to ensure proper gradient propagation through the network (Hochreiter & Schmidhuber, 1997). Standard RNN models suffer from vanishing gradient problems and cannot capture the long sequence context, and this is a dominant objective for NLP tasks. Like our earlier model, we have considered the output of two LSTM-based sub-networks as an input for the next level dense feed-forward layer. This network is comprised of two dense layers with 128 hidden units each. The input strings are also post-padded to produce an equal-length sequence. We have considered a DR of 0.45 for our LSTM model. Results remain consistent for other DRs in the range of 0.2 to 0.5. Similarly, in addition to 128 hidden units, we have also considered 64 and 256 hidden units, and our results remain broadly consistent.

Third, we employ a Bi-LSTM model to extract the contextual information from both directions. This model's hyperparameters are similar to the previous LSTM model - except for the subnetwork designed by the Bi-LSTM unit for improvement. Finally, we incorporate the attention mechanism (Rocktäschel et al., 2015) in our previous Bi-LSTM model to amplify the contribution of critical keywords within a tweet. An attention mechanism assigns weight to each word, and this reflects its importance. This attention mechanism aggregates all the intermediate hidden states using their relative importance and feed-forward to the subsequent dense layer for the classification task. Hence, this approach consists of an embedding layer, a Bi-LSTM layer, an attention layer, and the final classification layer with dropout to prevent overfitting (refer to Figure 6.6 for the model details).

6.2.5 Discussions

Table 6.6 reports the classification accuracies. We find advanced sequence-based models have outperformed the CNN model, and higher-dimensional word embeddings have outperformed lower-dimensional word embeddings. Our best-performing models have reported accuracies of around 97%, and this is significantly high. Intuitively, tweets are short texts - in comparison to the long job description and elaborate resumes. Thus, users try to incorporate multiple job attributes within a tweet. Hence, the word overlapping between a correctly matched pair can be potentially high.

We find that the average word counts of tweets from recruiters and jobseekers are 30.2 and 38.0, respectively. Next, we looked into common words and word share [=common words/(word count of recruiter's tweet + word count of jobseeker's tweet)] between a pair of tweets. We find that the average common words (word share) for correctly matched pairs are 4.39 (6.3%). Similarly, the average common words (word share) for wrongly matched pairs is 3.44 (4.8%). Thus, the proportion of common words is not very high in our corpus – even for correctly matched pairs. Intuitively, our approach might not be very appropriate for job domains that require soft skills. Soft skills might not be adequately expressed through keywords on the Twitter platform. Job-related tweets from different domains and a greater number of annotated tweets can enhance the robustness of our findings. Future studies also need to consider context-specific word embeddings instead of pre-trained GloVe. These approaches might open some exciting avenues for future research. To sum up, our proposed job recommendation framework using social media data can help a migrant worker in finding a suitable job. However, real-life implementation of our proposed framework would require voluminous training data across different domains.

Chapter 7: Conclusions and Future Research Directions

7.1 The Way Ahead: Explainable AI in the Domain of Migration

Extant literature has repeatedly emphasized that pre-trained models, such as transformer-based models like BERT or RoBERTa, can have inherent biases due to the training corpus (Klare et al., 2012; Buolamwini & Gebru, 2018; Raji & Buolamwini, 2019; Abid et al., 2021; Bender et al., 2021). For instance, if the training corpus comprises racial views toward migrants or refugees, implementing these models can be debatable. Hence, it is crucial for future research to employ explainable AI and take appropriate measures like debiasing. To examine the need for explainable AI, this section makes a preliminary attempt to interpret the functioning of BERT and RoBERTa models in the context of migration.

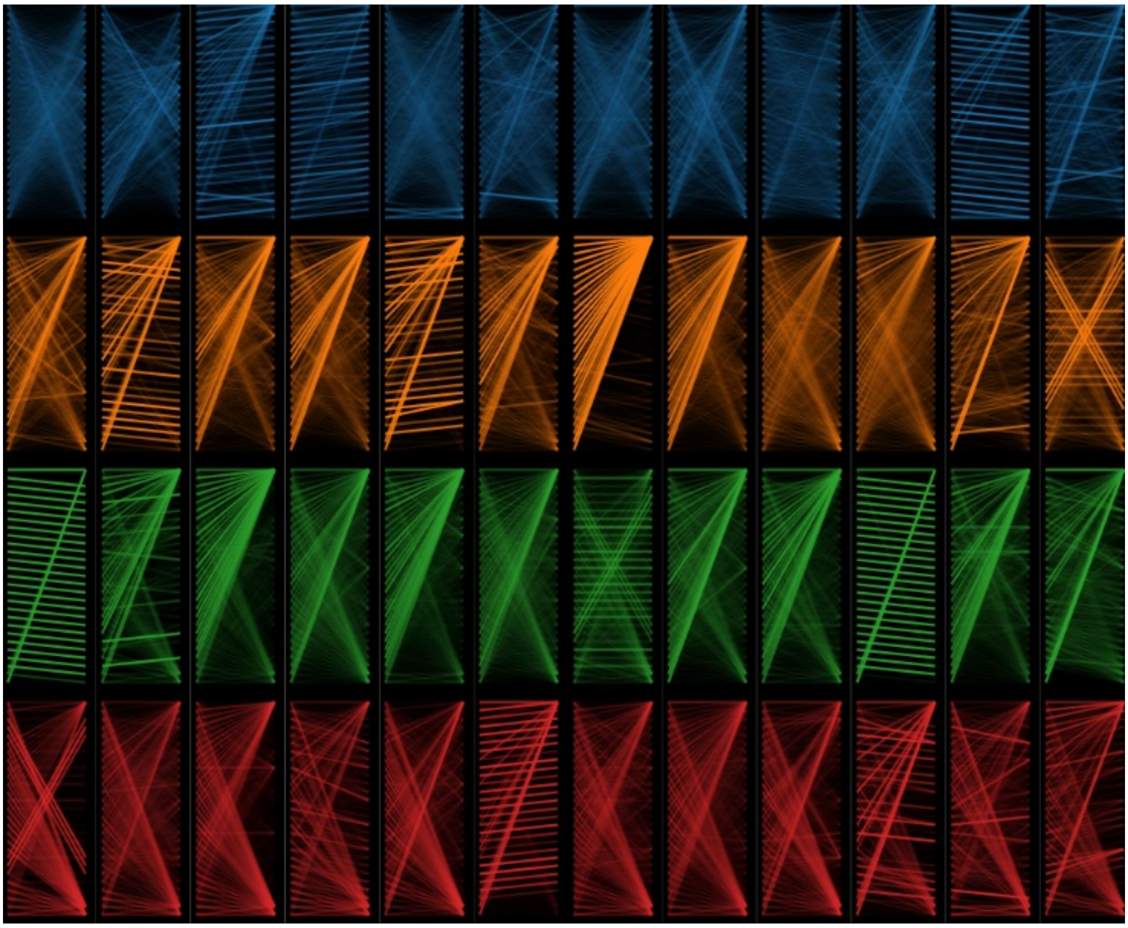


Figure 7.1 Model view of BERT for a sample tweet (excludes layers 4 to 11)

Following prior studies, such as Vig (2019) and Vig & Belinkov (2019), we aim to visually explore pre-trained models' attention weights using the BertViz library. This library is an improvement of the original *tensor2tensor-based* implementation (Jones, 2017). BertViz allows us to interpret and analyze the multi-head self-attention weights of the tokens from different BERT and RoBERTa model layers. BertViz tool can also generate the overall model view that helps us to see the "attention across all of the model's layers and heads for a particular input" tweet (Vig, 2019, p. 4). All attention heads are reported in a matrix format in this model view, where rows represent the layers and columns represent the heads. Vig (2019) pointed out that this snapshot view captures the contextual relation between tokens for all 12 layers and all heads. Figure 7.1 reports the model view for a tweet: *'If you haven't realized that nationalism and antiimmigrant violence is alarmingly in Europe then you haven't been paying attention'*. Reporting the entire 12*12 matrix will make the graphics clumsy. Hence, we have reported a truncated version of 12*4 (i.e., 0 to 11 heads and 0-3 layers). Careful observation reveals a few horizontal-stripe patterns (e.g., layer 0 & heads 2/3/10; layer 2 & heads 1/4) and a few triangular patterns (e.g., most heads in layer 1). This is in accordance with prior studies. Vig & Belinkov (2019) argue that a horizontal stripe indicates that "tokens attend to the current position," whereas a triangular pattern indicates that "they attend to the first token" (p.3). However, this overall model view does not allow to probe the exact contextual relationship between tokens.

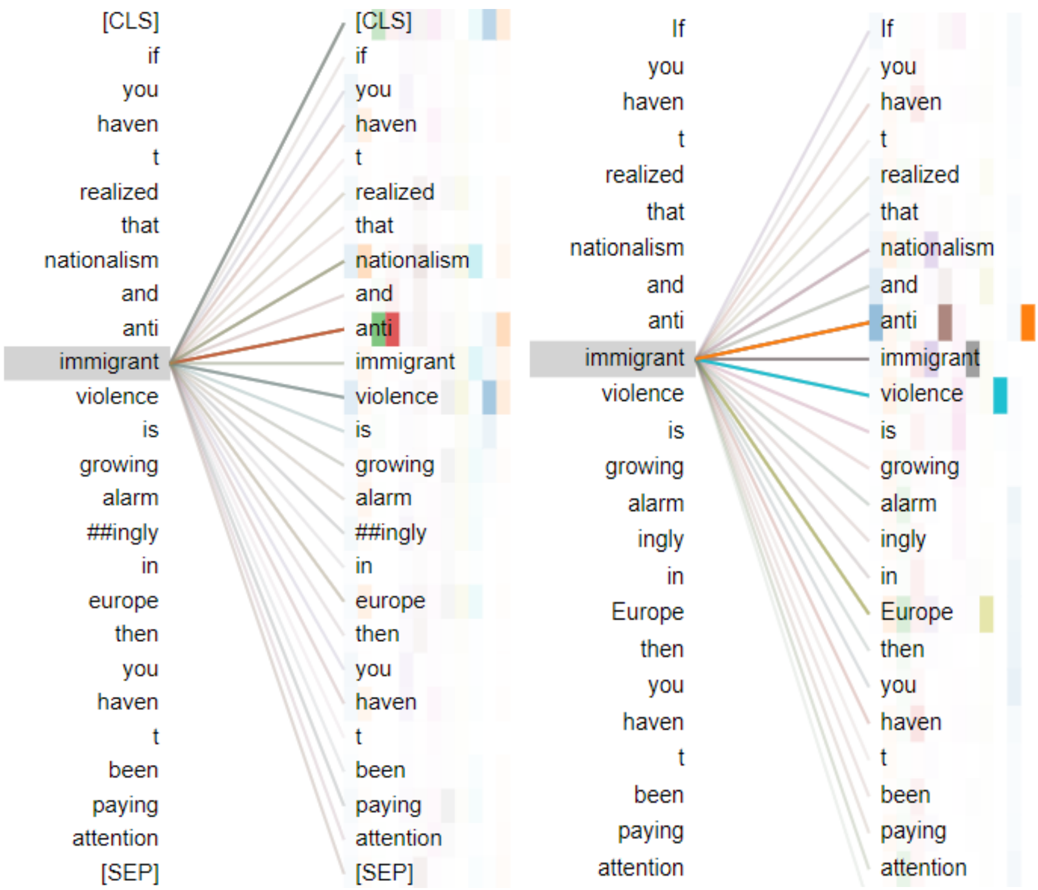


Figure 7.2 Attention to the word 'immigrant'

Hence, we need to explore the attention-head view to better understand these models. The attention-head view allows us to interpret the attention of one or more heads. In this graphical interface, self-attention is represented as lines from left to right, and the thickness of these lines represents the attention weight between tokens. However, some of the attention weights between tokens can be very low, and they will be nearly invisible in visualization. Different colors used in this visualization represent the individual heads in the model. This attention-head view allows us to explore our input sentences' precise attention heads. An in-depth analysis of these attention weights between tokens reveals how models, such as BERT and RoBERTa in our case, interpret the underlying relationship between words/tokens by allocating more attention to specific parts of the input text. Subsequently, these higher attentions, i.e., thicker lines, play a crucial role in the downstream task, which can be classification.

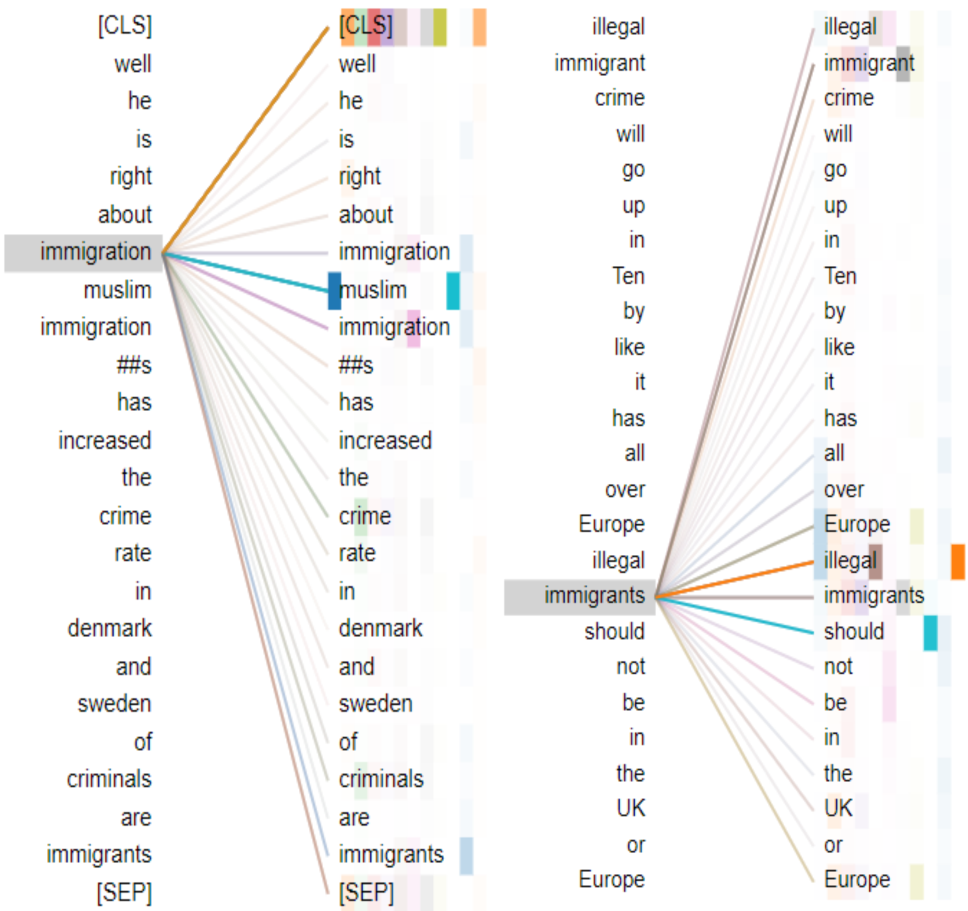


Figure 7.3 Attention to the words 'immigration' and 'immigrants'

Figure 7.2 reports the functioning of the BERT (the left figure) and RoBERTa (the right figure) models for the previous tweet of Figure 7.1. This figure elucidates why RoBERTa models can perform slightly better than BERT models. For instance, the BERT model links the token *'immigrant'* with the token *'anti'*. However, in addition to this *'anti'* token, the RoBERTa also links *'immigrant'* with *'violence'* and *'Europe'*. Thus, the RoBERTa model is also linking immigrants with the *European* context and violence. However, it is worth noting that the pre-trained RoBERTa model also links the issue of *'immigrant'* with the token *'anti'* and *'violence'*. Thus, this displays the pre-trained model's negative bias or the corpus's apprehension toward immigrants that the model has considered for training purposes. Intuitively, this also indicates that the RoBERTa model is potentially more biased than the BERT model.

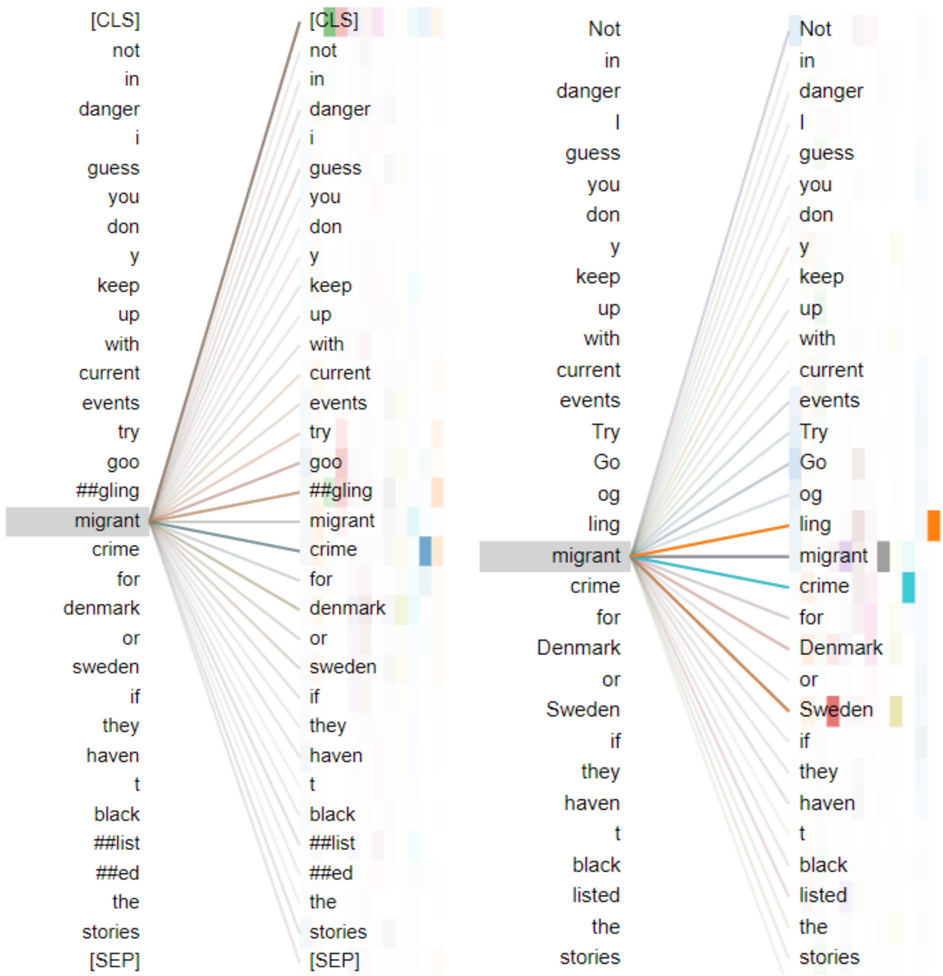


Figure 7.4 Attention to the word 'migrant'

Figure 7.3 considers two more tweets. The left figure reports the attention-head view of the tweet: *'Well, he is right about immigration. Muslim immigrations have increased the crime rate in Denmark and Sweden of criminals are immigrants'* (sic). Here, the BERT model connects the token *'immigration'* with *'Muslim'*. Similarly, the right figure reports the attention-head view of the tweet: *'Illegal immigrant crime will go up in Tenby like it has all all over Europe. Illegal immigrants should not be in the UK or Europe'*(sic). Here, the pre-trained RoBERTa model links the token *'immigrants'* with *'illegal'*. We note in our pilot study of Chapter 1 that a significant portion of migrants and refugees in European countries come from Middle Eastern countries, and many of them enter the European continent illegally. However, all immigrants are neither Muslim nor illegal. Thus, the attention-head views of these tweets reconfirm the bias of pre-trained models.

Figure 7.4 reports the attention-head view of another tweet as follows: '*Not in danger I guess you dony keep up with current events try Googling migrant crime for Denmark or Sweden if they haven't blacklisted the stories'* (sic). Both BERT and RoBERTa are linking the token *'migrant'* with *'crime'*. RoBERTa is also interpreting the context of *'Denmark'* and *'Sweden'* and linking the same with the token 'migrant'. In other words, the pre-trained RoBERTa model more accurately interprets the token 'migrant' context than the pre-trained BERT model. However, both models associate *'crime'* with *'migrant'*. We do agree that a particular section of the migrants may get involved in illegal activities. However, lack of equal opportunities or social discrimination in the host nation can be the antecedents for these illegal activities. Notably, all migrants are not criminals. In fact, skilled and educated migrants do add value to host nations. Contrarily, our pre-trained models display only a negative bias against migrants. This preliminary interpretability analysis strongly indicates that pre-trained models can be biased – indeed, the association between tokens is very similar to some of the myths and misconceptions. In short, this is alarming and AI-based future studies must consider these finer aspects.

7.2 Concluding Thoughts

Migration, as a research topic, has attracted the attention of researchers from multiple disciplines, like sociology, communication, economics, psychology, and sociology, over the years. **Chapter 2** indicates that social media-based research is a relatively newer domain for migration studies, and AI-based analysis of user-generated data is a prevalent approach in this domain. Our dissertation has made a humble effort to contribute to this literature. However, the potential of social media data is (probably) unlimited. Hence, we are fully aware that the scope of a dissertation cannot do justice to a diverse, intricate, and well-researched topic like migration.

Based on our literature review in Chapter 2, extant AI-based migration research can be classified into two domains: *opinion mining* (this also includes hate speech toward migrants and refugees) and *integration studies*. Accordingly, this dissertation has attempted to address a few unexplored issues in these two streams of research. We initiated this journey with a pilot study in **Chapter 1**. This initial study aimed to identify various salient themes of migrant-related social media discourse. We have identified 5 themes: *safety concerns*; *economic conditions*; *employment opportunities; healthcare support*; and *inequality & discrimination*. Interestingly, some themes are *pro-migrant* (like employment opportunities; healthcare support; and inequality & discrimination). However, we also observed *apprehensive views* toward migrants due to safety concerns, such as violence by migrants, in the host economies.

Taking a cue from the pilot study, **Chapter 3** deeply delves into societal opinion mining. Drawing insights from the social psychology literature, we have argued that classifying social media discourse into simplistic pro-migrant and anti-migrant categories would not be able to capture the finer nuances like perceptions and behaviors. So, we have employed the *perception-behavior theory framework* to probe social media discourse. Social psychology literature argues that 'perception' mostly leads to 'behavior' (Dijksterhuis & Bargh, 2001). Accordingly, sympathy may lead to solidarity, and antipathy may lead to animosity. However, social media data, especially Twitter data, did not allow us to test this causality. Hence, Chapter 3 has identified and examined various nuances of two prevailing perceptions (i.e., *sympathy* and *antipathy*) and two dominant behaviors (i.e., *solidarity* and *animosity*). Furthermore, we note that both antipathy and animosity tweets can be considered hate speech toward migrants. However, hate speech's perceptual and behavioral aspects are not the same. Hence, Chapter 3 has also fine-tuned the binary hate speech detection task. We also note that a pro-migrant tweet can also use 'swear words' against discrimination. We have crawled 0.8 million tweets from May 2020 to September 2020 and annotated 1193 tweets for final analysis. On the methodology front, we have employed ZSLMs, DL models with fastText embedding, and transformer-based models. We find that our proposed BERT + CNN architecture has outperformed other models for this complex perception-behavior identification task. Finally, our study had a few crucial takeaways for policymakers. An intervention mechanism can endorse positive perceptions and debunk myths or inappropriate perceptions. Consequently, we may observe a trend of lower animosity and more solidarity toward migrants in the long run.

In addition to opinion mining, extant literature has also explored the social integration of migrants and refugees. However, most of these studies investigated a specific aspect of integration, such as cultural integration ,learning a new language, or how they overcome psychological stress. The literature has rarely probed the arduous journey from their home to host nations. Hence, **Chapter 4** attempts to investigate the journey from displacement to emplacement. We refer to Arnold van Gennep's anthropological work *Les Rites de Passage* for theoretical anchoring. Gennep's study has systematically analyzed an individual's transition from one group (or society) to another. Based on Gennep's separation-transition-incorporation framework, we have identified four phases of refugee journeys: *Arrival of Refugees*, *Temporal stay at Asylums*, *Rehabilitation*, and *Integration of Refugees* into the host nation. We find multimodal tweets are more insightful than unimodal tweets. Hence, we considered 0.23 million refugee-related multimodal English tweets from April 2020 to March 2021 and annotated 1722 tweets to test our proposed framework. We have employed an early fusion of various unimodal models; however, our multimodal BERT+ LSTM + InceptionV4 model has outperformed other models. Subsequently, the 2022 Ukrainian refugee crisis allowed us to test – whether our proposed framework is generic or context-specific. To investigate this, we have considered 10,000 multimodal tweets from February 24, 2022, to March 15, 2022, and annotated 234 tweets. We find that refugees' struggles, risks, and traumas are identical – irrespective of who they are or wherever they come from. Once again, the multimodal BERT+ LSTM + InceptionV4 model has outperformed other models for this complex unseen data. In our follow-up analysis, we find that a few phases, such as *rehabilitation* and *integration of refugees*, are primarily associated with positive emotions, whereas *arrivals of refugees* and *temporal stay at asylums*, mostly display negative emotions. Similarly, we note that female references are significantly higher for the *rehabilitation* and *integration phase*, but the *arrival of refugee* phase is more associated with male references.

Our literature review in Chapter 2 reveals that AI-based migration studies have mostly considered a sample of manually annotated data. This manually annotated structured data offered a nuanced understanding of a specific issue related to migrants, but a few thousand annotated data points do not provide insights about general societal opinion toward migrants. To the best of our knowledge, none of the prior studies employed an unsupervised approach to perform a large-scale analysis to explore whether social media users are generally xenophobic or xenophilic toward migrants. To address this research gap, **Chapter 5** has explored whether users endorse xenophobic or xenophilic comments on the YouTube platform. To probe this, we referred to *racism theory* and tried to understand whether migrants are *cultural us* or *cultural others* through endorsement analysis on the YouTube platform. This question becomes intriguing in the 2022 Ukrainian refugee crisis context. Thus, the follow-up question became - *does our endorsement pattern depend on who the migrants are or from where they are?* Drawing insights from racism theory, we propose a set of competing hypotheses. To test our proposed hypotheses, we have considered two datasets: 110,803 migrant-related comments from 2778 videos posted from January 01, 2018, to December 31, 2020, i.e., before the Ukraine crisis, and 21,453 migrant-related comments from 342 videos posted from February 16, 2022, to April 14, 2022, i.e., during the Ukraine crisis. On the methodology front, we have employed regression analysis, specifically Logit and Tobit regression models, considering the nature of our data. Our main dependent variable is endorsement. We have assumed that if users like a comment, they endorse the sentiments expressed in that comment. Our unsupervised approach has considered LIWC scores for the operationalization of our explanatory variables. We have controlled for an exhaustive set of video-level (video popularity and video-level media biases) and comment-level (comment length and time gap between the video posting and comment) parameters. Empirical evidence suggests that social media users mainly endorse positive comments and generally do not endorse comments with abusive or swear words. Endorsement patterns for comments with negative sentiments across two datasets offer an exciting insight. Social media users did not endorse negative sentiments during the Ukraine crisis, but the pattern was different for the earlier dataset. Overall, the endorsement pattern during the Ukraine crisis strongly supported the cultural us hypothesis, but our findings were inconclusive for the earlier dataset.

Chapter 6 explores the psychological trauma and stress migrants face before settling down in the host nation and the potential of social media platforms to find a suitable job. Extant literature noted that refugees and migrants experience high mental stress and struggles. AI-based studies examined social media data for probing psychological stress. However, extant literature has rarely investigated the (faint and unheard) voices of migrants and refugees to analyze their psychological concerns. We have employed AI-based approaches for identifying their generic views, initial struggles, and subsequent settlement in the host country. Our best-performing transformer-based model has reported an accuracy of 75.89%. This analysis also suggests that the initial struggles of refugees and migrants mostly revolve around finding a suitable job in the host nation. Prior studies on job recommendation have mostly considered datasets from commercial job portals. However, these proprietary portals charge their clients and, thus, might not be feasible for economically vulnerable migrants. Hence, we explore whether publicly available Twitter data can be a viable alternative to commercial job portals. From 0.76 million job-related tweets, we annotated tweet pairs from recruiters and job seekers in the domain of computer science jobs and created 3980 recruiter-jobseeker tweet pairs (this includes 38% correctly and 62% wrongly matched pairs). Our Siamese architecture-based approaches report an accuracy of around 97% and demonstrate the potential of the Twitter platform for job recommendations.

As we pointed out earlier, the scope of this dissertation is not even the tip of the iceberg for a topic like migration. We have explored a few selected aspects of social media data, but a significant portion of social media data is untapped, unobserved, and unexplored. For instance, one stream of migration studies tried to predict the stock of migrants based on user-generated social media data (Zagheni et al., 2014; 2017; 2018). This dissertation did not probe in this direction. Contrarily, this dissertation revolved around opinion mining, specifically hate speech toward migrants who are cultural others (in Chapters 3 and 5), and some specific aspects of social integration of migration like how social media can help them to find a suitable job (Chapter 6) or different phases of their journey (Chapter 4), i.e., from struggle to settling down (Chapter 6). This dissertation made a small progress toward identifying the concerns shared by refugees or migrants. However, we feel that future studies should give equal emphasis to opinion mining and exploration of the faint (and often unheard) voices of voiceless migrants. Finally, Chapter 7 points out the need for explainable AI.

We conclude this dissertation by quoting a few selected excerpts (in Section 7.3) from one the most thought-provoking speeches on international migrants by Ban Ki-moon, the former Secretary-General of the United Nations from 2007 to 2016. This dissertation echoes the views and concerns expressed by Ban Ki-moon. This powerful speech was delivered in the year 2013, and the situation is broadly the same even a decade later. Hence, this justifies the relevance of this research. Hopefully, this dissertation has made an incremental contribution to employing AI for Social Good, especially in the context of international migrants – in accordance with the overarching themes of United Nations SDGs.

7.3 An Epitaph for Missing Migrants

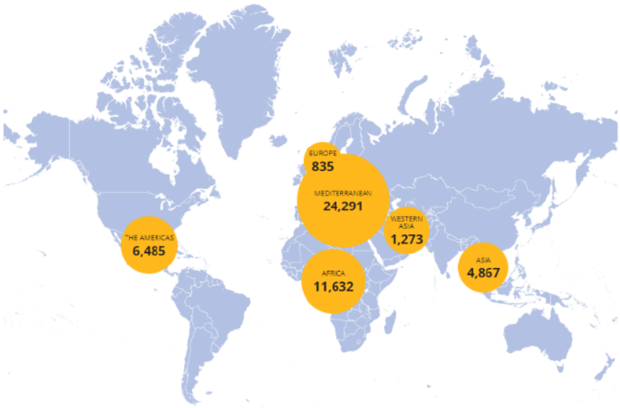


Figure 7.5 Deaths During Migration[44]

... We must do more to protect the human rights of all migrants. Too often, migrants live in fear -- of being victimized as the so-called "other"; of having little recourse to justice; or of having their wages or passports withheld by an unscrupulous employer. We cannot remain silent. We need to eliminate all forms of discrimination against migrants, including those related to working conditions and wages.

... We must end the exploitation to which migrants are vulnerable, including human trafficking. These crimes often perpetuate vicious cycles of abuse, violence and poverty, to which women and children are particularly vulnerable.

... We need to improve public perceptions of migrants. Migrants contribute greatly to host societies. As entrepreneurs, they create jobs. As scientists, they are engines of innovation. They are doctors, nurses and domestic workers and often the unheralded heart of many service industries. Yet far too often they are viewed negatively. Too many politicians seek electoral advantage by demonizing migrants.

... We are fortunate to live in an era of information. Yet reliable data on migration and its impact on development are often very hard to come by. Migration policies should be guided by facts, rather than hunches and hearsay.

Migration is an expression of the human aspiration for dignity, safety and a better future. It is part of the social fabric, part of our very make-up as a human family.

~ Ban Ki-moon[45]*, Secretary-General of the United Nations, 2007 – 2016*

[44] Missing Migrants Project tracks incidents involving migrants, including refugees and asylum-seekers, who have died or gone missing in the process of migration toward an international destination. The above image reports the deaths during migration recorded since 2014, by region of incident. The figures for 2022 were last updated on July 4, 2022. Available at https://missingmigrants.iom.int/, Accessed on July 6, 2022

[45] Selected excerpts from the Secretary-General's remarks to High-Level Dialogue on International Migration and Development on October 3, 2013. Available at https://www.un.org/sg/en/content/sg/statement/2013-10-03/secretary-generals-remarks-high-level-dialogue-international, Accessed on July 1, 2022

References

Abid, A., Farooqi, M., & Zou, J. (2021). Large language models associate Muslims with violence. *Nature Machine Intelligence*, *3*(6), 461-463.

Abou-Khalil, V., Helou, S., Flanagan, B., Pinkwart, N., & Ogata, H. (2019). Language learning tool for refugees: Identifying the language learning needs of Syrian refugees through participatory design. *Languages*, *4*(3), 71.

Adler-Nissen, R., Andersen, K. E., & Hansen, L. (2020). Images, emotions, and international politics: the death of Alan Kurdi. *Review of International Studies*, *46*(1), 75-95.

Aguirre, L., & Domahidi, E. (2021). Problematic Content in Spanish Language Comments in YouTube Videos about Venezuelan Refugees and Migrants. Journal of Quantitative Description: Digital Media, 1.

Ahmad, K., Pogorelov, K., Riegler, M., Ostroukhova, O., Halvorsen, P., Conci, N., & Dahyot, R. (2019). Automatic detection of passable roads after floods in remote sensed and social media data. *Signal Processing: Image Communication*, *74*, 110-118.

Al Qutob, M. F. (2018). Refugee healthcare: Refugees in crisis. *British Dental Journal*, *225*(3), 201-201.

Alam, F., Ofli, F., & Imran, M. (2018). Crisismmd: Multimodal twitter datasets from natural disasters. In *Twelfth international AAAI conference on web and social media.*

Alcántara-Plá, M., & Ruiz-Sánchez, A. (2018). Not for Twitter: Migration as a silenced topic in the 2015 Spanish general election. In *Exploring Silence and Absence in Discourse* (pp. 25-64). Palgrave Macmillan, Cham.

Alencar, A. (2018). Refugee integration and social media: A local and experiential perspective. *Information, Communication & Society*, *21*(11), 1588-1603.

Alloush, M., Taylor, J. E., Gupta, A., Valdes, R. I. R., & Gonzalez-Estrada, E. (2017). Economic life in refugee camps. *World Development*, *95*, 334-347.

Armstrong, C., Poorthuis, A., Zook, M., Ruths, D., & Soehl, T. (2021). Challenges when identifying migration from geo-located Twitter data. *EPJ Data Science*, *10*(1), 1.

Aswad, F. M. S., & Menezes, R. (2018). Refugee and Immigration: Twitter as a Proxy for Reality. In *FLAIRS Conference* (pp. 253-258).

Atallah, S., & Mahdi, D. (2017). Law and politics of "safe zones" and forced return to Syria: Refugee politics in Lebanon. *Beirut: The Lebanese Center for Policy Studies.*

Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2018). Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, *41*(2), 423-443.

Banerjee, A. V., & Duflo, E. (2019). *Good Economics, Bad Economics: Six Ways We Get the World Wrong and How to Set It Right*. PublicAffairs.

Barbier, G., & Liu, H. (2011). Data mining in social media. In *Social network data analytics* (pp. 327-352). Springer, Boston, MA.

Basile, V., Bosco, C., Fersini, E., Nozza, D., Patti, V., Pardo, F. M. R., ... & Sanguinetti, M. (2019,). Semeval-2019 task 5: Multilingual detection of hate speech against immigrants and women in twitter. In *Proceedings of the 13th international workshop on semantic evaluation* (pp. 54-63).

Bellino, M. J., & Dryden-Peterson, S. (2019). Inclusion and exclusion within a policy of national integration: Refugee education in Kenya's Kakuma Refugee Camp. *British Journal of Sociology of Education*, *40*(2), 222-238.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623).

BenEzer, G., & Zetter, R. (2015). Searching for directions: Conceptual and methodological challenges in researching refugee journeys. *Journal of Refugee Studies*, *28*(3), 297-318.

Bi, Y., Wang, S., & Fan, Z. (2020). A multimodal late fusion model for e-commerce product classification. *arXiv preprint arXiv:2008.06179*.

Blandfort, P., Patton, D. U., Frey, W. R., Karaman, S., Bhargava, S., Lee, F. T., ... & Chang, S. F. (2019). Multimodal social media analysis for gang violence prevention. In *Proceedings of the International AAAI conference on web and social media* (Vol. 13, pp. 114-124).

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, *2*(1), 1-8.

Bondi, E., Xu, L., Acosta-Navas, D., & Killian, J. A. (2021). Envisioning communities: a participatory approach towards AI for social good. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 425-436).

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, *10*(1), 1-14.

Bozdag, C., & Smets, K. (2017). Understanding the images of Alan Kurdi with "small data": A qualitative, comparative analysis of tweets about refugees in Turkey and Flanders (Belgium). *International Journal of Communication*, *11*, 24.

Brown-Bowers, A., McShane, K., Wilson-Mitchell, K., & Gurevich, M. (2015). Postpartum depression in refugee and asylum-seeking women in Canada: a critical health psychology perspective. *Health:*, *19*(3), 318-335.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.

Burke, M., & Kraut, R. (2013). Using Facebook after losing a job: Differential benefits of strong and weak ties. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1419-1430).

Cabot, H. (2016). "Refugee Voices" Tragedy, Ghosts, and the Anthropology of Not Knowing. *Journal of Contemporary Ethnography*, *45*(6), 645-672.

Calderón, C. A., Blanco-Herrero, D., & Apolo, M. B. V. (2020a). Rejection and hate speech in Twitter: Content analysis of Tweets about migrants and refugees in Spanish. *Revista Española de Investigaciones Sociológicas (REIS)*, *172*(172), 21-56.

Calderón, C. A., Blanco-Herrero, D., Frías-Vázquez, M., & Seoane-Pérez, F. (2021). Refugees welcome? Online hate speech and sentiments in Twitter in Spain during the reception of the boat Aquarius. *Sustainability*, *13*(5), 2728.

Calderón, C. A., de la Vega, G., & Herrero, D. B. (2020b). Topic modeling and characterization of hate speech against immigrants on Twitter around the emergence of a far-right party in Spain. S*ocial Sciences*, 9(11), 188.

Calderón, C. A., Sánchez-Holgado, P., Quintana-Moreno, C., & Amores, J. J. (2022). Hate speech and social acceptance of migrants in Europe: Analysis of tweets with geolocation. *Comunicar*, *30*(71), 21-35.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155-205.

Camacho, D., Luzón, M. V., & Cambria, E. (2021). New trends and applications in social media analytics. *Future Generation Computer Systems*, *114*, 318-321.

Cambria, E., & Hussain, A. (2015). *Sentic computing. A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*. Socio-Affective Computing. Springer

Cambria, E., Li, Y., Xing, F. Z., Poria, S., & Kwok, K. (2020). SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis. In *Proceedings of the 29th ACM international conference on information & knowledge management* (pp. 105-114).

Cambria, E., Poria, S., Gelbukh, A., & Kwok, K. (2014). A common-sense based api for concept-level sentiment analysis. *Making Sense of Microposts (# Microposts2014)*, 2.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems*, *28*(2), 15-21.

Capozzi, A., Francisci Morales, G. D., Mejova, Y., Monti, C., Panisson, A., & Paolotti, D. (2020). Facebook Ads: Politics of Migration in Italy. In *International Conference on Social Informatics* (pp. 43-57). Springer, Cham.

Castle, S., & Diarra, A. (2003). The international migration of young Malians: Tradition, necessity or rite of passage. Available at https://sarahcastle.co.uk/docs/Traffickingreport_final_Octo-ber.pdf Accessed on July 1, 2022

Ceci, L. (2022). YouTube penetration in selected countries and territories 2022. Statista. https://www.statista.com/statistics/1219589/youtube-penetration-worldwide-by-country/, Accessed on July 1, 2022

Chen, H., & Zimbra, D. (2010). AI and opinion mining. *IEEE Intelligent Systems*, *25*(3), 74-80.

Chen, T., & Dredze, M. (2018). Vaccine images on Twitter: analysis of what images are shared. *Journal of medical Internet research*, *20*(4), e8221.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, *5*(1), 146-166.

Coletto, M., Esuli, A., Lucchese, C., Muntean, C. I., Nardini, F. M., Perego, R., & Renso, C. (2016). Sentiment-enhanced multidimensional analysis of online social networks: Perception of the mediterranean refugees crisis. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 1270-1277). IEEE.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8440-8451).

Conneau, A., Schwenk, H., Cun, Y. L., & Barrault, L. (2017). Very deep convolutional networks for text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers* (pp. 1107-1116). Association for Computational Linguistics.

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and anxiety*, *18*(2), 76-82.

Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2021). A definition, benchmark and database of AI for social good initiatives. *Nature Machine Intelligence*, *3*(2), 111-115.

Crawley, H. (2017). Migration: refugee economics. *Nature*, *544*(7648), 26-27.

Crawley, H., & Skleparis, D. (2018). Refugees, migrants, neither, both: categorical fetishism and the politics of bounding in Europe's 'migration crisis'. *Journal of Ethnic and Migration Studies*, *44*(1), 48-64.

Dave, V. S., Zhang, B., Al Hasan, M., AlJadda, K., & Korayem, M. (2018). A combined representation learning approach for better job and skill recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (pp. 1997-2005).

Davidson, T., Bhattacharya, D., & Weber, I. (2019). Racial Bias in Hate Speech and Abusive Language Detection Datasets. In *Proceedings of the Third Workshop on Abusive Language Online* (pp. 25-35).

Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1, pp. 512-515).

de Bruijn, J. A., de Moel, H., Weerts, A. H., de Ruiter, M. C., Basar, E., Eilander, D., & Aerts, J. C. (2020). Improving the classification of flood tweets with contextual hydrological information in a multimodal neural network. *Computers & Geosciences*, *140*, 104485.

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3267-3276).

Dekker, R., Engbersen, G., Klaver, J., & Vonk, H. (2018). Smart refugees: How Syrian asylum migrants use social media information in migration decision-making. *Social Media+ Society*, *4*(1), 2056305118764439.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

Devlin, J.; Chang, M.W.; Lee, K.; and Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In NAACL-HLT (1).

Dijksterhuis, A. and Bargh, J.A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In *Advances in Experimental Social Psychology*. 33, 1-40.

Dilrukshi, I., & De Zoysa, K. (2014). A feature selection method for twitter news classification. *International Journal of Machine Learning and Computing*, *4*(4), 365.

Dong, J., He, F., Guo, Y., & Zhang, H. (2020). A commodity review sentiment analysis based on BERT-CNN model. In *2020 5th International Conference on Computer and Communication Systems (ICCCS)* (pp. 143-147). IEEE.

Dubois, A., Zagheni, E., Garimella, K., & Weber, I. (2018). Studying migrant assimilation through Facebook interests. In *International Conference on Social Informatics* (pp. 51-60). Springer, Cham.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., & Schönberg, U. (2017). On the economics and politics of refugee migration. *Economic policy*, *32*(91), 497-550.

Ekman, M. (2018). Anti-refugee mobilization in social media: The case of soldiers of Odin. *Social Media+ Society*, *4*(1), 2056305118764431.

Fang, M., & Cohn, T. (2016). Learning when to trust distant supervision: An application to low-resource POS tagging using cross-lingual projection. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning* (pp. 178-186).

Farkas, J., Schou, J., & Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, *20*(5), 1850-1867.

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in cognitive sciences*, *8*(1), 33-39.

Fiorio, L., Abel, G., Cai, J., Zagheni, E., Weber, I., & Vinué, G. (2017). Using Twitter data to estimate the relationship between short-term mobility and long-term migration. In *Proceedings of the 2017 ACM on web science conference* (pp. 103-110).

FitzGerald, D. S., & Arar, R. (2018). The sociology of refugee migration. *Annual Review of Sociology*, *44*, 387-406.

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European psychologist*, *18*(1), 12.

Flores, R. D. (2017). Do anti-immigrant laws shape public sentiment? A study of Arizona's SB 1070 using Twitter data. *American Journal of Sociology*, *123*(2), 333-384.

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: seven essential factors. *Science and Engineering Ethics*, *26*(3), 1771-1796.

Foster, D. (2020). *Factors influencing the popularity of YouTube videos and users' decisions to watch them*. Doctoral Dissertation, School of Mathematics and Computing, University of Wolverhampton

Francisco, V. (2015). 'The internet is magic': Technology, intimacy and transnational families. *Critical Sociology*, *41*(1), 173-190.

Gallego, M., Gualda, E., & Rebollo, C. (2017). Women and refugees in Twitter: Rhetorics on abuse, vulnerability and violence from a gender perspective. *Journal of Mediterranean Knowledge (ISSN 2499-930X)*, *2*(1).

Gallo, I., Ria, G., Landro, N., & La Grassa, R. (2020). Image and text fusion for upmc food-101 using bert and cnns. In *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)* (pp. 1-6). IEEE.

Geyik, S. C., Guo, Q., Hu, B., Ozcaglar, C., Thakkar, K., Wu, X., & Kenthapadi, K. (2018). Talent search and recommendation systems at LinkedIn: Practical challenges and lessons learned. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 1353-1354).

Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, *101*, 417-428.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012-1014.

Goodman, S., Sirriyeh, A., & McMahon, S. (2017). The evolving (re) categorisations of refugees throughout the "refugee/migrant crisis". *Journal of Community & Applied Social Psychology*, *27*(2), 105-114.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Gualda, E., & Rebollo, C. (2016). The refugee crisis on Twitter: A diversity of discourses at a European crossroads. *Journal of Spatial and Organizational Dynamics*, *4*(3), 199-212.

Gui, T., Zhu, L., Zhang, Q., Peng, M., Zhou, X., Ding, K., & Chen, Z. (2019). Cooperative multimodal approach to depression detection in twitter. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 110-117).

Guidry, J. P., Austin, L. L., Carlyle, K. E., Freberg, K., Cacciatore, M., Meganck, S., ... & Messner, M. (2018). Welcome or not: Comparing# refugee posts on Instagram and Pinterest. *American Behavioral Scientist*, *62*(4), 512-531.

Hadgu, A. T., Naini, K. D., & Niederée, C. (2016). Welcome or not-welcome: Reactions to refugee situation on social media. *arXiv preprint arXiv:1610.02358*.

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review*, *61*(4), 5-14.

Hattar-Pollara, M. (2019). Barriers to education of Syrian refugee girls in Jordan: Gender-based threats and challenges. *Journal of nursing scholarship*, *51*(3), 241-251.

Hauck, G. (2022, March 3). Black and brown refugees are once again being turned away in Europe amid Ukraine migrant crisis. USA Today. Available at https://eu.usatoday.com/story/news/world/2022/03/03/ukraine-refugee-crisis-racism-european-union/9329667002/ , Accessed on July 1, 2022

He, C., Chen, S., Huang, S., Zhang, J., & Song, X. (2019). Using convolutional neural network with BERT for intent determination. In *2019 International Conference on Asian Language Processing (IALP)* (pp. 65-70). IEEE.

He, H., Gimpel, K., & Lin, J. (2015). Multi-perspective sentence similarity modeling with convolutional neural networks. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 1576-1586).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hedderich, M. A., Adelani, D., Zhu, D., Alabi, J., Markus, U., & Klakow, D. (2020). Transfer Learning and Distant Supervision for Multilingual Transformer Models: A Study on African Languages. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 2580-2591).

Henkelmann, J. R., de Best, S., Deckers, C., Jensen, K., Shahab, M., Elzinga, B., & Molendijk, M. (2020). Anxiety, depression and post-traumatic stress disorder in refugees resettling in high-income countries: systematic review and meta-analysis. *BJPsych open*, *6*(4).

Himelboim, I., Smith, M. A., Rainie, L., Shneiderman, B., & Espina, C. (2017). Classifying Twitter topic-networks using social network analysis. *Social media+ society*, *3*(1), 2056305117691545.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Hoewe, J., Peacock, C., Kim, B., & Barnidge, M. (2020). The relationship between Fox News use and Americans' policy preferences regarding refugees and immigrants. International Journal of Communication, 14, 21.

Hrdina, M. (2016). Identity, activism and hatred: Hate speech against migrants on Facebook in the Czech Republic in 2015. *Naše společnost*, *14*(1), 38-47.

i Orts, Ò. G. (2019). Multilingual detection of hate speech against immigrants and women in Twitter at SemEval-2019 task 5: Frequency analysis interpolation for hate in speech detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 460-463).

Ibrahim, Y. (2018). The unsacred and the spectacularized: Alan Kurdi and the migrant body. *Social Media+ Society*, *4*(4), 2056305118803884.

International Organization for Migration, & United Nations. (2000). *World migration report*. Geneva, International Organization for Migration.

Jaki, S., & De Smedt, T. (2019). Right-wing German hate speech on Twitter: Analysis and automatic detection. arXiv preprint arXiv:1910.07518.

Jayatunga, M. K., Xie, W., Ruder, L., Schulze, U., & Meier, C. (2022). AI in small-molecule drug discovery: A coming wave. *Nat. Rev. Drug Discov*, *21*, 175-176.

Jones, L. (2017). Tensor2tensor transformer visualization. Available at https://github.com/tensorflow/tensor2tensor/tree/master/tensor2tensor/visualization

Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). Fasttext. zip: Compressing text classification models. *arXiv preprint arXiv:1612.03651*.

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A Convolutional Neural Network for Modelling Sentences. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 655-665).

Karakayali, S. (2018). The Flüchtlingskrise in Germany: Crisis of the refugees, by the refugees, for the refugees. *Sociology*, *52*(3), 606-611.

Karsu, O., Kara, B. Y., & Selvi, B. (2019). The refugee camp management: a general framework and a unifying decision-making model. *Journal of Humanitarian Logistics and Supply Chain Management*, *9*(2), 131-150.

Kaufmann, K. (2018). Navigating a new life: Syrian refugees and their smartphones in Vienna. *Information, Communication & Society*, *21*(6), 882-898.

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. *Gastrointestinal endoscopy*, *92*(4), 807-812.

Kenthapadi, K., & Tran, T. T. (2018). Pripearl: A framework for privacy-preserving analytics and reporting at linkedin. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (pp. 2183-2191).

Khalid, S. (2022, March 10). Q&A: Understanding Europe's response to Ukrainian refugee crisis, Al Jazeera. Available at https://www.aljazeera.com/news/2022/3/10/qa-why-europe-welcomed-ukrainian-refugees-but-not-syrians , Accessed on July 1, 2022

Khan, F., & Amatya, B. (2017). Refugee health and rehabilitation: challenges and response. *Journal of rehabilitation medicine*, *49*(5), 378-384.

Khatua, A., & Khatua, A. (2016). Leave or remain? Deciphering Brexit deliberations on Twitter. In *2016 IEEE 16th international conference on data mining workshops (ICDMW)* (pp. 428-433). IEEE.

Khatua, A., & Nejdl, W. (2020). Matching recruiters and jobseekers on twitter. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 266-269). IEEE.

Khatua, A., & Nejdl, W. (2021a). Analyzing European Migrant-related Twitter Deliberations. In *Companion Proceedings of the Web Conference 2021* (pp. 166-170).

Khatua, A., & Nejdl, W. (2021b). Struggle to Settle down! Examining the Voices of Migrants and Refugees on Twitter Platform. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 95-98).

Khatua, A., & Nejdl, W. (2022a). Endorsement Analysis of Migrant-related Deliberations on YouTube: Prior to and During 2022 Ukrainian crisis. In *Open Challenges in Online Social Networks* (pp. 31-38).

Khatua, A., & Nejdl, W. (2022b). Rites de Passage: Elucidating Displacement to Emplacement of Refugees on Twitter. In *Proceedings of the 33rd ACM Conference on Hypertext and Social Media* (pp. 214-219).

Khatua, A., & Nejdl, W. (2022c). Unraveling Social Perceptions & Behaviors towards Migrants on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 16, pp. 512-523).

Khatua, A., Khatua, A., & Cambria, E. (2019). A tale of two epidemics: Contextual Word2Vec for classifying twitter streams during outbreaks. *Information Processing & Management*, 56(1), 247-257.

Khatua, A., Khatua, A., & Cambria, E. (2020). Predicting political sentiments of voters from Twitter in multi-party contexts. *Applied Soft Computing*, 97, 106743.

Khawaja, N. G., Ibrahim, O., & Schweitzer, R. D. (2017). Mental wellbeing of students from refugee and migrant backgrounds: The mediating role of resilience. *School Mental Health*, *9*(3), 284-293.

Kiela, D., Firooz, H., Mohan, A., Goswami, V., Singh, A., Ringshia, P., & Testuggine, D. (2020). The hateful memes challenge: Detecting hate speech in multimodal memes. *Advances in Neural Information Processing Systems*, *33*, 2611-2624.

Kim, D. O. D., Curran, N. M., & Kim, H. T. C. (2020a). Digital Feminism and Affective Splintering: South Korean Twitter Discourse on 500 Yemeni Refugees. *International Journal of Communication*, *14*, 19.

Kim, J., Sîrbu, A., Giannotti, F., & Gabrielli, L. (2020b). Digital footprints of international migration on twitter. In *International Symposium on Intelligent Data Analysis* (pp. 274-286). Springer, Cham.

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),* Doha, Qatar (pp. 1746–1751). Association for Computational Linguistics.

Kira, I. A., Amer, M. M., & Wrobel, N. H. (2014). Arab refugees: Trauma, resilience, and recovery. In *Biopsychosocial perspectives on arab americans* (pp. 175-195). Springer, Boston, MA.

Klare, B. F., Burge, M. J., Klontz, J. C., Bruegge, R. W. V., & Jain, A. K. (2012). Face recognition performance: Role of demographic information. *IEEE Transactions on Information Forensics and Security*, *7*(6), 1789-1801.

Kreis, R. (2017). # refugeesnotwelcome: Anti-refugee discourse on Twitter. *Discourse & Communication*, *11*(5), 498-514.

Laban, C. J., Gernaat, H. B., Komproe, I. H., Van der Tweel, I., & De Jong, J. T. (2005). Postmigration living problems and common psychiatric disorders in Iraqi asylum seekers in the Netherlands. *The Journal of nervous and mental disease*, *193*(12), 825-832.

Latorre, J. P., & Amores, J. J. (2021). Topic modelling of racist and xenophobic YouTube comments. Analyzing hate speech against migrants and refugees spread through YouTube in Spanish. In *Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality* (TEEM'21) (pp. 456-460).

Leasure, D. R., Kashyap, R., Rampazzo, F., Elbers, B., Dooley, C., Weber, I., ... & Mills, M. C. (2022). Ukraine Crisis: Monitoring population displacement through social media activity.

Lee, J. S., & Nerghes, A. (2017). Labels and sentiment in social media: On the role of perceived agency in online discussions of the refugee crisis. In *Proceedings of the 8th International Conference on Social Media & Society* (pp. 1-10).

Lee, J. S., & Nerghes, A. (2018). Refugee or migrant crisis? Labels, perceived agency, and sentiment polarity in online discussions. *Social Media+ Society*, 4(3), 2056305118785638.

Leurs, K. (2014). The politics of transnational affective capital: Digital connectivity among young Somalis stranded in Ethiopia. *Crossings: Journal of Migration & Culture*, *5*(1), 87-104.

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 7871-7880).

Lingel, J., Naaman, M., & Boyd, D. M. (2014). City, self, network: transnational migrants and online identity work. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 1502-1510).

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-167.

Liu, M., Wang, J., Abdelfatah, K., & Korayem, M. (2019). Tripartite vector representations for better job recommendation. *arXiv preprint arXiv:1907.12379*.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A Robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*.

Maclachlan, M., & McAuliffe, E. (1993). Critical incidents for psychology students in a refugee camp: implications for counselling. *Counselling Psychology Quarterly*, *6*(1), 3-11.

Madukwe, K., Gao, X., & Xue, B. (2020). In data we trust: A critical analysis of hate speech detection datasets. In *Proceedings of the Fourth Workshop on Online Abuse and Harms* (pp. 150-161).

Marco, M., Luisa, B., Raffaella, B., Stefano, M., & Roberto, Z. (2014). SemEval-2014 Task 1: Evaluation of compositional distributional semantic models on full sentences through semantic relatedness and textual entailment. In *Proc. SemEval* (pp. 1-8).

Mazzoli, M., Diechtiareff, B., Tugores, A., Wives, W., Adler, N., Colet, P., & Ramasco, J. J. (2020). Migrant mobility flows characterized with digital data. *PloS one*, *15*(3), e0230264.

McAuliffe, M., Khadria, B., & Bauloz, C. (2019). World migration report 2020. IOM.

Ménard, P. A., & Mougeot, A. (2019). Turning silver into gold: error-focused corpus reannotation with active learning. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing* (RANLP 2019) (pp. 758-767).

Meng, Q., Zhu, H., Xiao, K., & Xiong, H. (2018). Intelligent salary benchmarking for talent recruitment: A holistic matrix factorization approach. In *2018 IEEE International Conference on Data Mining (ICDM)* (pp. 337-346). IEEE.

Meng, Q., Zhu, H., Xiao, K., Zhang, L., & Xiong, H. (2019). A hierarchical career-path-aware neural network for job mobility prediction. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 14-24).

Meyer, M. A. (2019). Healthcare data scientist qualifications, skills, and job focus: a content analysis of job postings. *Journal of the American Medical Informatics Association*, *26*(5), 383-391.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning--based text classification: a comprehensive review. *ACM Computing Surveys (CSUR)*, *54*(3), 1-40.

Mitsilegas, V. (2017). Humanizing solidarity in European refugee law: The promise of mutual recognition. *Maastricht Journal of European and Comparative Law*, *24*(5), 721-739.

Monsutti, A. (2007). Migration as a rite of passage: Young Afghans building masculinity and adulthood in Iran. *Iranian Studies*, *40*(2), 167-185.

Montgomery, E. (2010). Trauma and resilience in young refugees: A 9-year follow-up study. *Development and psychopathology*, *22*(2), 477-489.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *Proceedings of the international AAAI conference on web and social media* (Vol. 7, No. 1, pp. 400-408).

Mortensen, M. (2017). Constructing, confirming, and contesting icons: The Alan Kurdi imagery appropriated by# humanitywashedashore, Ai Weiwei, and Charlie Hebdo. *Media, Culture & Society*, *39*(8), 1142-1161.

Mozafari, M., Farahbakhsh, R., & Crespi, N. (2020). Hate speech detection and racial bias mitigation in social media based on BERT model. *PloS one*, *15*(8), e0237861.

Mueller, J., & Thyagarajan, A. (2016). Siamese recurrent architectures for learning sentence similarity. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 30, No. 1).

n.d. Media Bias/Fact Check News. Media Bias/Fact Check. Available at https://mediabiasfactcheck.com/ , Accessed on July 1, 2022

Naja, W. J., Aoun, M. P., El Khoury, E. L., Abdallah, F. J. B., & Haddad, R. S. (2016). Prevalence of depression in Syrian refugees and the influence of religiosity. *Comprehensive psychiatry*, *68*, 78-85.

Neculoiu, P., Versteegh, M., & Rotaru, M. (2016). Learning text similarity with siamese recurrent networks. In *Proceedings of the 1st Workshop on Representation Learning for NLP* (pp. 148-157).

Nerghes, A., & Lee, J. S. (2018). The refugee/migrant crisis dichotomy on Twitter: A network and sentiment perspective. In *Proceedings of the 10th ACM conference on web science* (pp. 271-280).

Nerghes, A., & Lee, J. S. (2019). Narratives of the refugee crisis: A comparative study of mainstream-media and Twitter. *Media and Communication*, *7*(2), 275-288.

Nesterko, Y., Jäckle, D., Friedrich, M., Holzapfel, L., & Glaesmer, H. (2020). Prevalence of post-traumatic stress disorder, depression and somatisation in recently arrived refugees in Germany: an epidemiological study. *Epidemiology and psychiatric sciences*, *29*.

Nie, Y., Williams, A., Dinan, E., Bansal, M., Weston, J., & Kiela, D. (2020). Adversarial NLI: A New Benchmark for Natural Language Understanding. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 4885-4901).

Ogan, C., Pennington, R., Venger, O., & Metz, D. (2018). Who drove the discourse? News coverage and policy framing of immigrants and refugees in the 2016 US presidential election. *Communications*, *43*(3), 357-378.

O'Leary, D. E. (2013). Artificial intelligence and big data. *IEEE intelligent systems*, *28*(2), 96-99.

Olesen, T. (2018). Memetic protest and the dramatic diffusion of Alan Kurdi. *Media, Culture & Society*, *40*(5), 656-672.

O'Mara, B. (2013). Social media, digital video and health promotion in a culturally and linguistically diverse Australia. *Health Promotion International*, *28*(3), 466-476.

Ottoni, R., Cunha, E., Magno, G., Bernardina, P., Meira Jr, W., & Almeida, V. (2018). Analyzing right-wing youtube channels: Hate, violence and discrimination. In Proceedings of the 10th ACM conference on web science (pp. 323-332).

Ozcaglar, C., Geyik, S., Schmitz, B., Sharma, P., Shelkovnykov, A., Ma, Y., & Buchanan, E. (2019). Entity personalized talent search models with tree interaction features. In *The World Wide Web Conference* (pp. 3116-3122).

Özerim, M. G., & Tolay, J. (2021). Discussing the populist features of anti-refugee discourses on social media: an anti-Syrian hashtag in Turkish Twitter. *Journal of Refugee Studies*, *34*(1), 204-218.

Öztürk, N., & Ayvaz, S. (2018). Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telematics and Informatics*, *35*(1), 136-147.

Palakodety, S., KhudaBukhsh, A. R., & Carbonell, J. G. (2020). Voice for the voiceless: Active sampling to detect comments supporting the rohingyas. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 01, pp. 454-462).

Pamungkas, E. W., Basile, V., & Patti, V. (2020). Do you really want to hurt me? predicting abusive swearing in social media. In The *12th Language Resources and Evaluation Conference* (pp. 6237-6246). European Language Resources Association.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).

Peterie, M., & Neil, D. (2020). Xenophobia towards asylum seekers: A survey of social theories. *Journal of Sociology*, 56(1), 23-35.

Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., & Patti, V. (2021). Resources and benchmark corpora for hate speech detection: a systematic review. *Language Resources and Evaluation*, *55*(2), 477-523.

Pope, D., & Griffith, J. (2016). An Analysis of Online Twitter Sentiment Surrounding the European Refugee Crisis. In *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (pp. 299-306).

Pushp, P. K., & Srivastava, M. M. (2017). Train once, test anywhere: Zero-shot learning for text classification. *arXiv preprint arXiv:1712.05972*.

Qin, C., Zhu, H., Zhu, C., Xu, T., Zhuang, F., Ma, C., ... & Xiong, H. (2019). Duerquiz: A personalized question recommender system for intelligent job interview. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2165-2173).

Quan, K. A. C., Nguyen, V. T., Nguyen, T. C., Nguyen, T. V., & Tran, M. T. (2020). Flood level prediction via human pose estimation from social media images. In *Proceedings of the 2020 international conference on multimedia retrieval* (pp. 479-485).

Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 429-435).

Ramanath, R., Inan, H., Polatkan, G., Hu, B., Guo, Q., Ozcaglar, C., ... & Geyik, S. C. (2018). Towards deep and representation learning for talent search at LinkedIn. In *Proceedings of the 27th ACM international conference on information and knowledge management* (pp. 2253-2261).

Rampazzo, F., Bijak, J., Vitali, A., Weber, I., & Zagheni, E. (2021). A framework for estimating migrant stocks using digital traces and survey data: an application in the United Kingdom. *Demography*, *58*(6), 2193-2218.

Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2017). Snorkel: Rapid training data creation with weak supervision. In *Proceedings of the VLDB Endowment. International Conference on Very Large Data Bases* (Vol. 11, No. 3, p. 269). NIH Public Access.

Reel, S., Wong, P., Wu, B., Kouadri, S., & Liu, H. (2018). Identifying tweets from Syria refugees using a Random Forest classifier. In *2018 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 1277-1280). IEEE.

Rettberg, J. W., & Gajjala, R. (2016). Terrorists or cowards: negative portrayals of male Syrian refugees in social media. *Feminist Media Studies*, *16*(1), 178-181.

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira Jr, W. (2020). Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 131-141).

Rijhwani, S., Zhou, S., Neubig, G., & Carbonell, J. G. (2020). Soft Gazetteers for Low-Resource Named Entity Recognition. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 8118-8123).

Rivera-Pagán, L. N. (2013). Xenophilia or xenophobia: Toward a theology of migration. In *Contemporary Issues of Migration and Theology* (pp. 31-51). Palgrave Macmillan, New York.

Riyadi, N. A. E., & Widhiasti, M. R. (2020). Racism and stereotyping of refugees: The use of the hashtag# aufschrei by Twitter users in Germany. In *International University Symposium on Humanities and Arts (INUSHARTS 2019)* (pp. 151-157). Atlantis Press.

Rocktäschel, T., Grefenstette, E., Hermann, K. M., Kočiský, T., & Blunsom, P. (2015). Reasoning about entailment with neural attention. *arXiv preprint arXiv:1509.06664.*

Ross, B., Rist, M., Carbonell, G., Cabrera, B., Kurowsky, N., & Wojatzki, M. (2017). Measuring the reliability of hate speech annotations: The case of the european refugee crisis. *arXiv preprint arXiv:1701.08118.*

Rychalska, B., & Dabrowski, J. (2020). Synerise at SIGIR Rakuten Data Challenge 2020: Efficient Manifold Density Estimator for Cross-Modal Retrieval. *SIGIR eCom Challenge 2020.*

Sagredos, C., & Nikolova, E. (2022). 'Slut I hate you': A critical discourse analysis of gendered conflict on YouTube. *Journal of Language Aggression and Conflict*, *10*(1), 169-196.

Saha, K., Sugar, B., Torous, J., Abrahao, B., Kıcıman, E., & De Choudhury, M. (2019). A social media study on the effects of psychiatric medication use. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 13, pp. 440-451).

Sajir, Z., & Aouragh, M. (2019). Solidarity, social media, and the" refugee crisis": Engagement beyond affect. *International Journal of Communication*, *13*, 28.

Sanguinetti, M., Poletto, F., Bosco, C., Patti, V., & Stranisci, M. (2018). An italian twitter corpus of hate speech against immigrants. In *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018).*

Schick, M., Zumwald, A., Knöpfli, B., Nickerson, A., Bryant, R. A., Schnyder, U., ... & Morina, N. (2016). Challenging future, challenging past: the relationship of social integration and psychological impairment in traumatized refugees. *European journal of psychotraumatology*, *7*(1), 28057.

Schweitzer, R., Buckley, L., & Rossi, D. (2002). The psychological treatment of refugees and asylum seekers: what does the literature tell us?. *Mots Pluriels*, *21*.

Shalaby, W., AlAila, B., Korayem, M., Pournajaf, L., AlJadda, K., Quinn, S., & Zadrozny, W. (2017). Help me find a job: A graph-based approach for job recommendation at scale. In *2017 IEEE international conference on big data (big data)* (pp. 1544-1553). IEEE.

Sharma, M., & Sharma, S. (2020). The rising number of COVID-19 cases reflecting growing search trend and concern of people: a Google Trend analysis of eight major countries. *Journal of Medical Systems*, *44*(7), 1-3.

Siapera, E., Boudourides, M., Lenis, S., & Suiter, J. (2018). Refugees and network publics on Twitter: Networked framing, affect, and capture. *Social Media+ Society*, *4*(1), 2056305118764437.

Siersdorfer, S., Chelaru, S., Nejdl, W., & San Pedro, J. (2010). How useful are your comments? Analyzing and predicting YouTube comments and comment ratings. In *Proceedings of the 19th international conference on World wide web* (pp. 891-900).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Spörlein, C., & Schlueter, E. (2021). Ethnic insults in YouTube comments: Social contagion and selection effects during the german “Refugee Crisis”. *European Sociological Review*, *37*(3), 411-428.

Stark, O. (2004). On the economics of refugee flows. *Review of Development Economics*, *8*(2), 325-329.

Stewart, I., Flores, R. D., Riffe, T., Weber, I., & Zagheni, E. (2019). Rock, Rap, or Reggaeton?: Assessing Mexican immigrants' cultural assimilation using Facebook data. In *The World Wide Web Conference* (pp. 3258-3264).

Suh, J., Horvitz, E., White, R. W., & Althoff, T. (2021). Population-scale study of human needs during the covid-19 pandemic: Analysis and implications. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (pp. 4-12).

Sun, S., Luo, C., & Chen, J. (2017). A review of natural language processing techniques for opinion mining systems. *Information fusion*, *36*, 10-25.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

Tang, J., Chang, Y., & Liu, H. (2014). Mining social media with social theories: a survey. *ACM Sigkdd Explorations Newsletter*, *15*(2), 20-29.

Tang, L., Fujimoto, K., Amith, M. T., Cunningham, R., Costantini, R. A., York, F., ... & Tao, C. (2021). “Down the rabbit hole” of vaccine misinformation on YouTube: Network exposure study. *Journal of Medical Internet Research*, 23(1), e23262.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24-54.

Thelwall, M. (2021). Lifestyle information from YouTube influencers: some consumption patterns. *Journal of Documentation*, 77(6), 1209-1222.

Thelwall, M., & Foster, D. (2021). Male or female gender-polarized YouTube videos are less viewed. *Journal of the Association for Information Science and Technology*, 72(12), 1545-1557.

Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., ... & Clopath, C. (2020). AI for social good: unlocking the opportunity for positive impact. *Nature Communications*, *11*(1), 1-6.

Turing, A. M. (2009). Computing machinery and intelligence. In Parsing the turing test (pp. 23-65). Springer, Dordrecht.

Turner, S. (2016). What is a refugee camp? Explorations of the limits and effects of the camp. *Journal of Refugee Studies*, *29*(2), 139-148.

Udwan, G., Leurs, K., & Alencar, A. (2020). Digital resilience tactics of Syrian refugees in the Netherlands: Social media for social support, health, and identity. *Social Media+ Society*, *6*(2), 2056305120915587.

UNHCR Data Portal. Available at https://data2.unhcr.org/en/situ-ations/ukraine . Accessed on July 1, 2022.

Urbina, F., Lentzos, F., Invernizzi, C., & Ekins, S. (2022). Dual use of artificial-intelligence-powered drug discovery. *Nature Machine Intelligence*, *4*(3), 189-191.

Urchs, S., Wendlinger, L., Mitrović, J., & Granitzer, M. (2019). MMoveT15: A Twitter Dataset for Extracting and Analysing Migration-Movement Data of the European Migration Crisis 2015. In *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 146-149). IEEE.

Van Gennep, A. (1960). Les rites de passage (Paris, 1909). *The Rites of Passage: Anthropology and Ethnography*.

Vázquez, M. F., & Pérez, F. S. (2019). Hate speech in spain against aquarius refugees 2018 in twitter. In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 906-910).

Vidgen, B., Thrush, T., Waseem, Z., & Kiela, D. (2021). Learning from the Worst: Dynamically Generated Datasets to Improve Online Hate Detection. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1667-1682).

Vig, J. (2019). A multiscale visualization of attention in the transformer model. *arXiv preprint arXiv:1906.05714*.

Vig, J., & Belinkov, Y. (2019). Analyzing the structure of attention in a transformer language model. *arXiv preprint arXiv:1906.04284*.

Vigna, F. D., Cimino, A., Dell'Orletta, F., Petrocchi, M., & Tesconi, M. (2017). Hate me, hate me not: Hate speech detection on facebook. In *Proceedings of the First Italian Conference on Cybersecurity (ITASEC17)* (pp. 86-95).

Vis, F., & Goriunova, O. (Eds.) (2015). *The Iconic Image on Social Media: A Rapid Research Response to the Death of Aylan Kurdi*. Visual Social Media Lab.

Volkan, V. D. (2018). Refugees as the other: Large-group identity, terrorism and border psychology. *Group analysis*, *51*(3), 343-358.

Waldinger, R. (2018). Immigration and the election of Donald Trump: Why the sociology of migration left us unprepared... and why we should not have been surprised. *Ethnic and Racial Studies*, *41*(8), 1411-1426.

Wang, X., Kumar, D., Thome, N., Cord, M., & Precioso, F. (2015). Recipe recognition with large multimodal food dataset. In *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (pp. 1-6). IEEE.

Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? predictive features for hate speech detection on twitter. In *Proceedings of the NAACL student research workshop* (pp. 88-93).

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2019). Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.

Yang, S., Korayem, M., AlJadda, K., Grainger, T., & Natarajan, S. (2017). Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach. *Knowledge-Based Systems*, *136*, 37-45.

Yin, W., Hay, J., & Roth, D. (2019). Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3914-3923).

Yoshida, M., Arase, Y., Tsunoda, T., & Yamamoto, M. (2015, June). Wikipedia page view reflects web search trend. In *Proceedings of the ACM web science conference* (pp. 1-2).

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing. *IEEE Computational Intelligence Magazine*, 13(3), 55-75.

Zagheni, E., Garimella, V. R. K., Weber, I., & State, B. (2014). Inferring international and internal migration patterns from twitter data. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 439-444).

Zagheni, E., Polimis, K., Alexander, M., Weber, I., & Billari, F. C. (2018). Combining social media data and traditional surveys to nowcast migration stocks. In *Annual Meeting of the Population Association of America*.

Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review*, 721-734.

Zeng, D., Chen, H., Lusch, R., & Li, S. H. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems*, *25*(6), 13-16.

Zhu, C., Zhu, H., Xiong, H., Ma, C., Xie, F., Ding, P., & Li, P. (2018). Person-job fit: Adapting the right talent for the right job with joint representation learning. *ACM Transactions on Management Information Systems (TMIS)*, *9*(3), 1-17.

Zunes, S. (2017). Europe's refugee crisis, terrorism, and islamophobia. *Peace review*, *29*(1), 1-6.