



Article

Using Social Media to Monitor Conflict-Related Migration: A Review of Implications for A.I. Forecasting

Hamid Akin Unver

Department of International Relations, Özyeğin University, Istanbul 34337, Turkey

Abstract: Following the large-scale 2015–2016 migration crisis that shook Europe, deploying big data and social media harvesting methods became gradually popular in mass forced migration monitoring. These methods have focused on producing ‘real-time’ inferences and predictions on individual and social behavioral, preferential, and cognitive patterns of human mobility. Although the volume of such data has improved rapidly due to social media and remote sensing technologies, they have also produced biased, flawed, or otherwise invasive results that made migrants’ lives more difficult in transit. This review article explores the recent debate on the use of social media data to train machine learning classifiers and modify thresholds to help algorithmic systems monitor and predict violence and forced migration. Ultimately, it identifies and dissects five prevalent explanations in the literature on limitations for the use of such data for A.I. forecasting, namely ‘policy-engineering mismatch’, ‘accessibility/comprehensibility’, ‘legal/legislative legitimacy’, ‘poor data cleaning’, and ‘difficulty of troubleshooting’. From this review, the article suggests anonymization, distributed responsibility, and ‘right to reasonable inferences’ debates as potential solutions and next research steps to remedy these problems.

Keywords: conflict; forced migration; artificial intelligence; event data; big data ethics

Citation: Unver, Hamid Akin. 2022.

Using Social Media to Monitor Conflict-Related Migration: A Review of Implications for A.I. Forecasting. *Social Sciences* 11: 395. <https://doi.org/10.3390/socsci11090395>

Academic Editor: Carlos Arcila Calderón, Tuba Bircan and Nigel Parton

Received: 19 May 2022

Accepted: 22 August 2022

Published: 1 September 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction: Opportunities and Pitfalls of Extracting Information from Violent and Migration-Prone Regions

Mass human displacement has often been a by-product of organized violence. Some of the largest forced migration events in world history have been triggered by anthropogenic disasters, and international or intrastate conflicts still remain among the most immediate causes of refugee crises (Lozano-Gracia et al. 2010; Schon 2019; Selby and Hoffmann 2012). Since civil wars are especially fought near populated areas to establish control over population centers, they are particularly prone to generating mass displacement (Steele 2009; Lichtenheld 2020). Indeed, frequent exchanges of territory as a result of conflict—such as in Ukraine, Syria, South Sudan, Myanmar, DR Congo, Somalia, Central African Republic, Afghanistan, and Iraq—generate an overwhelming majority of the world refugee population. Both in human history and today, armed conflicts remain one of the most significant predictors of forced displacement and, often, the size of forced migration.

Violent events lead to the mass uprooting of noncombatants for two reasons. First, the immediate life-threatening potential of conflicts leads to the departure of entire villages, towns, and sometimes cities, fleeing into safer regions (Basu and Pearlman 2017). Often, large-scale population displacement gets triggered because civilians expect harsh treatment or targeting by the conquerors and leave to escape from such fate (Conte and Migali 2019). Second, the aftershocks of a conflict create significant infrastructure and sustenance problems that lead to the mass departure of civilians to gain access to essential resources and services (Humphrey 2013; Rajabali et al. 2009). Quite frequently, secondary effects such as the destruction of housing, sanitation and electrical infrastructure, absence

of law and order, and disruptions in food supplies generate escalating hardships in a population center, creating sustained, low-intensity migration (Sowers and Weinthal 2021; Crush 2013). This is not to suggest that armed violence automatically generates forced migration (sometimes civilians either choose to stay in the clash zone for various reasons, or have mobility problems that prevent such prospects), but one could safely assume that organized violence tends to create poverty, grievance, and threat stimuli that usually contributes to locals' assessments of staying or leaving (Tellez 2022; Epstein 2010).

Violent organized conflict is not solely an independent variable in the study of displacement. Often, climate and seasonal adversities (like bad harvest), in addition to natural disasters, may trigger forced migration, and are exacerbated by pre-existing or newly emerging forms of conflict (Ash and Obradovich 2020). In such cases, although violence does not necessarily initiate migration events, it nonetheless affects the tempo, duration, and direction of migration flows (Burke et al. 2015). Such disasters and climate-related effects also contribute to increased violence due to dwindling access to resources, generating additional competition over water, food, and medical supplies. Dormant grievances re-emerge due to a wide array of displacement triggers, and ethnic/religious groups may target each other during migration, or they may be targeted by outside hostile armed actors while on the move (Salehyan 2007).

Since armed conflict is either (or both) an independent and intervening variable in migration research, relief agencies and governments have been exploring ways to quantify and log event data to produce more informed analyses about both forms of interlinked crises. From the need to establish more detailed and robust mechanisms to monitor and forecast forced migration and to prevent violence, arose macro-level event datasets (Lubeck et al. 2003; Nam 2006; Eck 2012). These datasets initially logged violence in binary forms, and in a dyadic fashion, but in tandem, a number of highly granular and multi-directional datasets have begun appearing in scientific space (some examples are Chojnacki et al. 2012; Weidmann 2013; Demarest and Langer 2018). In the last decade, these datasets have been used not just for explaining conflict and migration, but also to forecast them (Blair and Sambanis 2020; Carammia et al. 2022).

The need to forecast organized violence and use it in turn to predict forced migration has long been a priority for state actors and international organizations. Prior to the proliferation of the scientific and private event data revolution, information about conflicts and migration was thus largely supplied by formal military observers. The main problem with these military reports was that they were prone to censorship and reporting bias, as states tended to misrepresent field events in line with their national interests (Banerjee and MacKay 2020). From the mid-20th century onwards, the onset of high-circulation newspapers, radio, and television changed the nature of war reporting and extracting information about conflict zones by the rise of war reporters and aid workers as an important additional source of conflict event data. Without a formal military chain of command constraints, war reporters and aid agencies have begun generating a more diverse array of field information for both military and non-military purposes, which was a welcome addition, since they suffered from fewer national interest censorship constraints. However, ultimately, they began suffering from other forms of observation bias. Reporters had editorial constraints (Knightley 2002), whereas aid workers were increasingly constrained by their superiors when field events did not conform to either editorial or agency/donor interests (Bunce et al. 2019). Gradually, more academic/scientific projects took pace by the 1970s, such as David Singer's 'Correlates of War Project', which brought local, national, and international news reports together to generate a richer account of field events through traditional media reports (Eck 2012).

However, even when used as part of scientific datasets, traditional media data has its bias limitations: Primarily, both local and international news outlets have editorial interests and political prerogatives that often murk their reporting efforts (Ravi 2005). In most cases, media groups choose to or not to report events based on their pre-existing political ties or editorial biases. Newspaper outlets have varying degrees of autonomy and

editorial exposure to governmental pressures, rendering the independence of media-based event reports highly skewed (Lee and Maslog 2005). Second, media data tends to focus on ‘big events’ that either affect a large number of people, or a wide swathe of territory in order to maximize its readership potential (Baum and Zhukov 2015). This leaves out ‘smaller events’ that are still field events but are omitted due to readership or sales considerations. Third, media groups tend to over-focus on events that concern their host country and leave out potentially important events that do not concern them directly (Weidmann 2016). These ‘irrelevant events’ can often be discarded—again, for home country readership considerations—but may otherwise contain important details to infer the inner workings of conflicts. Finally, as stipulated in the Herman–Chomsky Propaganda Model (Chomsky and Herman 2002; Mullen and Klaehn 2010), there are five filters through which events have to go through in order to appear in the news. These are (a) media ownership interests, (b) advertiser interests, (c) censorship or nudges from the government, (d) threats against established power (corporate or state), and (e) limiting the mobilizing effect of dissent. These filters form significant barriers against the utilization of mainstream, ‘established’ media news streams and as illustrated in the bombing of Serbia in 1999, they result in the significant overrepresentation of violence conducted by US adversaries.

Largely due to these limitations of ‘classical’ crisis media event data collection protocols, the advent of social media was heralded as an important novelty. The most critical novelty social media brought was its ‘disintermediated’ nature. Disintermediation is the reduction or removal of intermediaries in a system (in this case, news and communication) and within the context of digital communication, it defines the reduced importance of traditional news intermediaries—editors, reporters, censors—that serve as the middlemen between the news sources and news consumers (Sampedro et al. 2022). Social media enables a violent event or a migration movement on the ground, to be documented by participants themselves, or passerby civilians, through videos, images, and texts in real time, without any intermediaries, and very little (and obscure) content moderation, directly into the news feeds of digital users across the world (Eriksson 2018). Social media platforms have both overlapping and niche capabilities. Twitter excels in short-text/short-media format, whereas Facebook can be used for gallery-style extended documentation. Instagram offers a more visual platform experience, whereas TikTok’s comparative advantage is short-term ‘mood’ videos (Jaidka 2022). While Twitter data has so far been used extensively for crisis research due to its granular API system, researchers are increasingly experimenting with other social media platforms for crisis event data.

The downsides of social media data have precisely been by-products of the same disintermediation that also serves as its strength. Due to the absence of verification and fact-checking filters such as investigative reporters or desk editors, social media has been rife with disinformation and redundancy (Gohdes 2018; Zeitzoff 2018). Additionally, it has been falling prey to another form of availability bias, whereby social media field data can only be produced where smartphones, internet access or cell phone towers are present. This is an important factor because in war and disaster zones where all three can often be missing, social media as crisis event data can be very difficult to produce. A number of recent studies demonstrate this clearly as cell phone coverage and infrastructure has a strong impact on the volume and reliability of social media data coming from difficult-to-access regions (Pierskalla and Hollenbach 2013).

The remainder of this paper will focus on the following issues:

- Advantages and disadvantages of different forms of event data that feed currently deployed A.I. monitoring/forecasting models.
- A discussion of why social media is becoming more popular as field data that is being used to train forecasting models.
- Ethical considerations in using social media data to train A.I. conflict/migration forecasting classifiers.

2. The 'Achilles Heel' of Forecasting: Data Reliability

Forecasting critical and resource-intensive events like crises, migration, violence, and wars have long been in demand among government, military, and international organization circles. Building prediction and early warning systems allows governments and agencies to prepare for relief, aid, and law enforcement planning, and build resilience against major shocks. Given the limited resources of emergency response and relief institutions, forecasting can be an important cost optimization process, allowing such agencies to be ready during time-sensitive episodes that require substantial resources.

Forecasting is also viewed with waves of suspicion in social sciences. Most direct criticism is that the very basis of forecasting: Collating prior instances of cases to infer the timing and type of future events, has been viewed as unrealistic, or unable to properly capture social uncertainty (Dowding 2021). While more transformative and large-scale of such events—'black swans' (Ahmad et al. 2021)—have indeed been difficult to forecast, proponents of this approach viewed forecasting as nonetheless useful in estimating the future trajectory of existing events. This implies that, for any forecast, at least two types of input are required: *historical data*, which can be in the form of event, population, measurement or survey; and a model, which will form the foundation of the estimator that will determine the way historical data will be processed to produce the output, namely, *prediction* (Efendi et al. 2018).

Event data has been an important input in violence and migration forecasting. Given the fact that snow melts during springtime, insurgency and counter-insurgency operations record an uptick in mountainous areas during these periods. Monsoon seasons witness heavy rainfall, which restricts the movement of armor in muddy regions, resulting in reduced offensives by state militaries. Mass migration away from the conflict areas is expected when the monsoon season ends, the soil dries, and heavy armor can once again move into combat near population centers. Heavy artillery shelling often indiscriminately targets civilian population centers, and therefore the movement of mobile artillery units closer to such population centers tends to generate mass exodus. In winter, heavy snowfall and frost render combat operations in rugged areas more difficult, so combatants prepare for spring offensives in order to pursue their tactical objectives around these areas. As forced migration often happens around these seasonal variances, battle repertoires, and violence patterns, these independent variables provide researchers with some degree of confidence that past instances of events can be harnessed to predict unrealized outcomes. It is important to emphasize that conflicts and migration are spatio-temporally multi-dependent events and standard linear forecasts often fail to address the nuances of such variance (Christiansen et al. 2021).

Given its high-level granularity and data volume advantages compared to more traditional forms of crisis information, social media data is being increasingly leveraged to monitor, as well as forecast, conflicts and migration events. It is also being used as part of the broader suite of additional data inputs to train artificial intelligence (A.I.) classifiers that are being deployed to monitor and forecast migration (Alexander et al. 2022; Willekens 2018; Ning et al. 2019; Salah 2022; Bircan and Korkmaz 2021). Given bias limitations of traditional military observer, war reporter, classical media, and aid worker-based field data, conflict and migration forecasting big data and A.I. approaches are becoming increasingly interested in harvesting social media as an additional layer of field reports either in producing new datasets, or to complement existing ones. Social media data can increase the robustness of prediction models and forecasting confidence as it provides researchers with a greater volume of field data from crisis events, and if cleaned for redundancy and disinformation sufficiently, they can strengthen the validity of claims inferred through forecasts (Kaufhold et al. 2020; Shan et al. 2019).

In the migration and conflict forecasting context, A.I. could be defined as "a growing resource of interactive, autonomous, self-learning agency, which enables computational artifacts to perform tasks that otherwise would require human intelligence to be executed successfully" (Taddeo and Floridi 2018). Among these self-learning tasks are the

methodical assessment of algorithms and data inputs to improve their knowledge of performance with experience (Flach 2012). A.I.-based analytics and forecasting methods have been increasingly used by international organizations and governments to control and manage human mobility. These methods include identity recognition, automated border monitoring, and analysis of asylum applications (Beduschi 2021). Recently, these methods also began to harvest conflict event data (including social media data) in order to train A.I. models to predict migration (Molnar 2019a). Algorithmic migration monitoring is being increasingly relied on in the United States, United Kingdom, Canada, and most European Union members states, and forms the basis of next-generation migration technology investment in major international organizations.

A.I.-based analysis and forecasting systems require significant volumes of data—especially ‘big data’ in the form of not just large volume information, but also high granularity, high velocity, and high complexity information. While there is not always a direct relationship, data size is often an important variable in more sophisticated and accurate A.I.-based systems. To that end, training machine learning classifiers is a data-hungry endeavor, which requires not just a greater number of events, but multiple reports of the same event for greater robustness check and data validity. In addition, greater data input can render decision and output chains more efficient in further iterations by triangulation, as results are more robust results incrementally train further iterations with greater precision (Tarasyev et al. 2018; Quinn et al. 2018).

3. Ethics of Social Media as Forecasting Data

Forecasting systems that are built on machine learning principles can grow increasingly more accurate in their predictions and can ultimately make inferences about large-scale human flows during conflicts or natural disasters. This could aid in the optimization of relief aid or refugee camp preparations, as well as proper staffing of migration processing centers (Molnar 2019b).

However, forecasting migration broadly, in and of itself, may not always generate relief-oriented results. Mentioning these general pitfalls are important prior to connecting this argument to social media data. Countries can close borders or engage in direct preventative action to render refugee pathways more dangerous (Pécoud and de Guchteneire 2006; Vives 2017). Rebel groups may use forecast results published online to crack down on civilians and cut off their escape routes (Bruce 2001; Larson and Lewis 2018). An improved ability to predict migration may also cause local populations to flee earlier and in greater numbers, given the fact that such predictions are often shared through social media and can be seen by the locals through smart phones (Dekker et al. 2018). This, in turn, plays into the hands of smugglers, who may provide lesser-known pathways across countries for migrants escaping state precautions, endangering refugees (Sanchez 2017). These prospects grow more problematic as insufficiently optimized systems get deployed in decision and analytics roles, and end up misidentifying, miscalculating, and misjudging refugees and their actions. Furthermore, cyber-attacks and data protection problems may lead to migration-related datasets to be stolen, leaked, and used by malicious actors. It is important to underline that the European Data Protection Supervisor (EDPS) in its consultation with the European Asylum Support Office (EASO) (D(2019) 1961 C 2018-1083) had concluded that social media monitoring to predict migration patterns is not in line with the EU regulations. Yet, in its disclosure, the EASO has revealed that it has been actively harvesting social media data in order to support the operations of the Department of Operations, Country of Origin Information (COI), Information and Analysis Unit (IAU), and the Department of Asylum Support (DAS), among others.

Bringing the argument closer to the scope of this article, the use of social media as migration forecasting data on the other hand, brings in an even larger set of ethical considerations. First, the scale of migration data available through social media and mobile phone signals yields a false sense of sophistication in analyses, that often leads decision-makers to treat these data sources as representative, or formatted sufficiently to train

machine learning classifiers (Stewart and Wilson 2016). A number of important studies have time and again demonstrated that, even with its immense volume, social media or cell phone data from forced migrations cannot be treated as representative, or are not robust enough in their current form to properly train classifiers that lead to real world decisions (Silver and Andrey 2019). This lack of representativeness is a major contestation area between researchers that warn against using most media-based data systems on their own to guide decisions, and decision-makers that aim to leverage the surface sophistication of A.I. models built on flawed data to market the ‘validity’ of their decisions (Lachlan et al. 2016). Since predictive and analytical models are based on inferential statistics, they operate on degrees of probability, and thus the ‘acceptable’ threshold for decisions becomes a political benchmark, rather than technical (Schroeder et al. 2013). Determining which threshold is ‘sufficient’ to make life-altering inferences from data usually becomes a murky process that decision-makers find too complicated to think about, and gets delegated to engineers who have neither the political or legal legitimacy, nor appropriate training to make such decisions (Rahwan 2018; Cunneen et al. 2019). Given the complexity of the process, important details such as the variables and measurements used in benchmarking becomes obscured from public debate and creates a problematic ethical gap.

Although algorithmic decision-making structures can often be useful with sufficient oversight and control mechanisms, it is difficult to assert that big data migration forecasting protocols have sufficient safeguards in place, rendering the very core of the process detached from, and inaccessible to the very people it deals with (Zednik 2021). Ultimately, A.I. predictive and forecasting analytics operate on inferences based on a number of correlations, but to what extent these correlations represent causal mechanisms is not always straightforward (Milano et al. 2021). Given the data size and speed of social media-based information, establishing flexible and adjustable causal mechanisms becomes very difficult, which often forces engineers to offer choice alternatives for decision-makers based solely on correlations (Buhmann and Fieseler 2021). In cases where decisionmakers do not have the sufficient background or advisor support to dig deeper into the questioning of such causal mechanisms, as well as their direction and weight, big data migration forecasting and decision procedures become opaque and poorly calibrated.

This is in line with Molnar (2021), where the advent of COVID-19 pandemic has generated greater reliance on biosurveillance (virus-targeting robots, phone tracking, A.I.-enabled thermal cameras), which connects to Tendayi Achiume’s (2021) claim that surveillance technologies are reinforcing spatial racism. By using pandemic-related motives, border control agencies are relying increasingly on automated forecasting and prediction models to keep refugees and migrants within confined spaces—camps, border wall areas, or segregated processing centers within cities. These are what Jason de Leon (2015) calls as ‘land of open graves’, where authorities confine refugee movement and settlement into dangerous areas where self-sustenance is often difficult. Although more traditional border protection tactics seek similar outcomes, newer technologies render automated push-back decisions less accountable by exporting the authority of decisions to ambivalent models built on murky training data. As underlined in Molnar (2021), the existing track record of automated technologies on race and gender, open ups the path for the deployment of similar methods on migration surveillance.

Second, algorithms (as opposed to human supervised inference methods) are increasingly being used in a self-learning fashion during crisis forecasting and analysis. Most conflict and migration-related social media data are being used to train machine learning classifiers, which then ‘learn’ from this limited data and extract future instances of data collection protocols (Wachter et al. 2017). Automated data collection and analysis parameters are thus doubly detached from human agency, control, and legal responsibility: not only does it perform in an automated fashion, but it grows increasingly less influenced by the original human control over it that formed its primary legal and ethical basis (Mittelstadt et al. 2016). While this problem remains intact even for politically and legally well-controlled data protection mechanisms like the GDPR, for migration and refugee

forecasting, it is further problematically hidden from vulnerable populations, who cannot see or challenge the inference mechanism that leads to a particular decision that concerns their lives. Since the specific processes by which the collected data are used to generate likelihoods of an outcome are murky, civilians are exposed not only to decisions made by inconclusive evidence, but from a legal standpoint, cannot reliably challenge such decisions because parameters that create such decisions are abstracted from human agency by degrees (iterations) of self-training processes that generate machine learning classifiers (Zarsky 2013).

In discussing the ethics and legality of migration analytics and forecasting protocols, two key concepts form the core of the debate. These are *transparency* of the data collection, modeling, and decision chain; and the *explainability/comprehensibility* of the technicalities that lie within this link. Often, these protocols and classifiers, as well as the data streams that are integrated into them, are acquired second hand from second or third-party suppliers as part of 'integrated solutions' that contain preset classifiers and decision thresholds (Renda 2019). When using such second-hand solutions, state institutions or international organizations risk being challenged on the accessibility and comprehensibility grounds, since quite often these institutions themselves do not have the engineering capacity to understand the detailed analytics protocols themselves (Mittelstadt et al. 2016). Since institutions have low technical capacity to alter these thresholds, they quite often operate with 'one size fits all' parameters, that frequently produce inaccurate forecasting and analytics. While there is an ongoing debate on whether this accessibility gap is intentional or not, there is nonetheless such a legal and ethical gap that remains in place with such solutions.

Third, the richest aspect of social media data—that it is diverse, high-volume, and representing a broad range of views and perspectives from field events—has the danger of running counter to the very fundamental basis of A.I: it regularly modifies and alters its detection, data collection, processing, and measurement approaches in a self-learning fashion (Vallor and Bekey 2017). This renders legal and judicial oversight and legitimacy a moving target, as it brings in a critical question: Which parameter, threshold, and confidence interval range will be the basis of legal, ethical and political responsibility? If, for example, a parliament or a court approves the deployment of a particular algorithmic structure to be deployed during emergencies, migration crises and in monitoring violent conflict at $t = 0$, and if the algorithms adjust its fundamental entity recognition, statistical inference, and confidence interval optimization parameters as $t + 1$, $t + 2$... how will this approval translate over newer iterations of the algorithm? Will the algorithm be approved 'as is', or can it be discussed and deliberated within a specific parameter oscillation range, expecting its future self-learning alterations?

Given the wealth of social media data and its dramatic peaks and plateaus during key events, using such data as machine learning training input will likely cause significant episodic changes to how it collects and models emergency information. The fundamental question thus becomes: are parliaments, governments, and courts equipped to deal with the question of whether their $t = 0$ approval of an algorithm will be valid or not at $t + 1$ and beyond? If the algorithm decides to collect data beyond its initial legal and political confines to optimize its robustness, where will the legal and electoral responsibility lie? In political discourse, these algorithms are generally constructed as 'semi-supervised' suggesting that there will always be an engineer to optimize/oversee these parameter changes, recent scholarship posits that this is not always the case after an algorithm is legally and politically approved (Shneiderman 2016; Castets-Renard 2019; Elkin-Koren 2020). Lawyers, bureaucrats, and politicians, in turn, usually do not have the knowledge or advisory assistance to make such assessments beyond the medium-term as well, suggesting that self-learning algorithms that are vetted prior to their approval may steer further away from their initial parameters, and become too large to retrospectively troubleshoot, as time progresses and data size and variability increases.

Fourth, the use of social media data as one of the inputs of classifier training risks a high degree of data cleaning and wrangling problems. Often, data streams that harvest social media platforms and feed it into training and analytics dashboards have less-than-optimal data cleaning practices, which lead to the inclusion of potentially redundant, misleading, and sometimes incoherent information into the training chain (Chu et al. 2016; Jain et al. 2020; Pavlyshenko 2019). When a classifier is trained based on—for example—a social media data pool that includes heavy bot campaign, high-volume disinformation, or misleading information, it is going to generate significantly mismatched inferences compared to the ground reality. To illustrate, an attacker during a civil war can spread social media disinformation about the presence of ‘potential jihadi terrorists’ among the fleeing civilian populace, which may be picked up by automated migration crisis data extraction protocols and generate a flawed inference about the refugees, generating disproportionately preventative or harsh treatment by state security services, refugee camp administrations, or even by the locals of the towns and villages along the migration path. Since a proper cleaning and formatting of social media data takes time, current A.I. protocols have a high likelihood of generating noisy and flawed decisions during crises with time constraints.

This increases the risks of discrimination in big data migration research. While Romei and Ruggieri (2014) suggest that a controlled distortion of training data, introduction of anti-discrimination inputs into the classifiers, post-processing of classifiers for a second round check, and modifying fairness-related parameters in further iterations, this protocol becomes tricky with social media data that renders these safeguards extremely time-consuming. Given the fact that refugee crises are extensively securitized in traditional and social media data, narratives of migration tend to be highly securitized as well (Colombo 2018). Using social media directly as a text input also injects this discursive discrimination into analytical and forecasting chains, creating automatic biases about mass human mobility. Although technically social media text data can be post-processed to mitigate or at least soften these underlying discursive biases, it is both very time consuming, and also highly context-specific given the peculiarities of foreign languages, slang, and satire/sarcasm dynamics (Batinca and Treleaven 2015). While these parameters are well set in English, much more work is required to conduct this cleaning work in other languages—especially in languages in which data cleaning engineers are not proficient. This issue not only creates a generalized unfairness since an algorithm actively being trained on social media data will create immediate biases against refugees, but it also produces a second set of bias and unfairness due to the algorithm’s text corpus variances across different languages.

Fifth, and connected to the third, the sheer size of the A.I. forecasting and analytics data infrastructures, renders reverse engineering of mistakes very difficult and time consuming (Sejnowski 2020). Once such decisions lead to flawed outcomes—especially that result in harm against refugees—returning to training data and trying to understand which component caused the flawed outcome becomes a daunting task with social media data. Since such troubleshooting and fixing work requires extensive staffing, time, and financial resources, deciding to engage in such enterprise itself becomes a political and bureaucratic decision (Agrawal et al. 2018). Even if such decision is taken, trying to find the needle in the proverbial data haystack necessitates redirecting resources from daily monitoring, analytics, and decision-producing protocols. Such diminishing returns renders the entirety of the troubleshooting process deterring and potentially unsustainable, causing analytics and prediction teams to go ahead with insufficiently trained classifiers that continue to produce flawed results—and worse: learn from these flawed priors (Coglianese and Lehr 2016). When such cases persist, it becomes nearly impossible to find which data stream caused the flawed interpretation, which engineering department was responsible for it, and most importantly: who is to be held legally accountable for such flawed interpretations. This creates not just an ethical problem whereby the source of harm against migrants/refugees cannot be identified reliably, but it also poses a significant

legal problem where responsibility cannot be traced, and binding revisions cannot be applied to proper individuals, teams, and inference mechanisms. Not only that this mechanism puts people at risk into harm, but it also creates a digitally authoritarian construct which invisibly shuts the door to complaints, information requests and legal reparation options.

There are other problems with big data migration and violence forecasting that is beyond the scope of this paper, which focuses on using social media as classifier training data. These problems include (but are not limited to) problems with the analysis and modeling robustness aspect of algorithms, i.e., the fact that the statistical parameters that generate decisions or help the algorithm learn secondary tasks may not always be calibrated well enough, or accurate. Since there is a know-how and formational disconnect between engineers who calibrate these models and the decision-makers who act on the findings of such models, inference model accuracy becomes less of a mathematical or technical endeavor and more of a political and ethical one in which confidence intervals, variable weights, directions/strength of correlations, and how within-sample evidence is used to generate out-of-sample forecasts become biased, and potentially discriminatory.

4. Avenues for Ethical Social Media Data Use in Migration and Violence Forecasting

The above account may suggest that the number of issues that currently obscure the use of social media data as a reliable emergency information for training A.I.-based systems and assisting in their inference and forecast protocols are too insurmountable. This does not imply that social media data cannot be used for such purposes, but a number of broader ethical and legal decisions have to be made before more technical/programming optimization work can be undertaken.

One of the first pathways to render social media data 'more useable' in algorithmic format would be to find ways to anonymize and render such data untraceable back to a vulnerable person, or a group. While some anonymization approaches are available for accounts whose social media posts are part of training classifiers, such anonymization becomes harder in video and image data, where the content is crucial for inference, and can often contain information about actual identity of individuals and groups that are both vulnerable and in danger (Beigi et al. 2018; Townsend and Wallace 2017; Zhang et al. 2018). A new anonymization protocol is necessary to create a disconnect between actual identities and the inference produced as a result of it, as well as further studies that explore how lack of anonymity creates unintended discriminatory inferences by algorithms.

This logic, however, needs to be dissected when considering individual versus collective harm. Harm and discrimination towards one, or a small number of migrants fits into a different anonymization debate compared to harm and discrimination towards a larger flow of migration. In the former context, anonymization can provide a viable solution (as it is already deployed as such by the EASO and UNHCR), although it starts to become less relevant to the migrant protection debate within the context of large-scale flows. This point becomes a larger political and ideological issue (anti-immigration), as opposed to a technical and human rights-related debate because discrimination of one or a group of migrants within the basis of ethnicity, religion and biometrics sits in a different corner of the anonymization debate compared to extracting data from a mass exodus. In the latter context, identity-related variables form a secondary consideration in light of the more important first consideration: where the migration flow has originated, where it is headed to, and which pathways it transits through. The difference between the two also contains the difference between flawed state responses (a) by denying the right to claim asylum by a wholesale prevention or migration, and (b) by targeting one, or a group of people in a discriminatory fashion using data collected through social media.

The second path is to broaden the judicial, bureaucratic, and parliamentary discussions on the responsibility of A.I. systems. While this decision will likely differ across countries and political systems, countries will most likely opt for a shared responsibility infrastructure for A.I.-based decision errors, distributing legal and political error margins

between engineers, system managers and A.I. protocols. In a flawed decision that creates a significant level of human suffering, the legal attribution chain will likely follow through (1) the institution, (2) the director, (3) the sub-manager, (4) the chief engineer, (5) the junior engineer that runs, operates, and maintains the code structure of the algorithm. Which line of the hierarchy is to blame will likely be a political decision that different countries will take differently based on power dynamics between various related institutions, and within the institution that has undertaken the flawed decision (Shah 2018). In recent years, a new debate emerged on the ‘moral agency’ of algorithms—i.e., whether non-human actors of an algorithm such as sequence, selection, and iteration can be held responsible in a court of law (Véliz 2021; Cunneen et al. 2019). This debate risks what Wachter et al. define as ‘de-responsibilization of human actors’, which in simple terms is to ‘hide behind the computer’ and exporting the responsibility of flawed decisions to unprosecutable agents (Lösch et al. 2017; Kirkpatrick 2016). This trend runs closer to Hannah Arendt’s observation of how inhuman and structurally hostile actions can dispassionately be undertaken by bureaucratic networks and actors that are following rules and protocols that create automated iterations of suffering without agency on their part.

The third path follows the ‘right to reasonable inferences’ debate and its introduction into the debate on the use of social media data for big data migration and violence forecasting purposes (Wachter and Mittelstadt 2019; Veronese et al. 2019). In GDPR context, this right posits that individuals whose data is harvested for inference and prediction tasks that concern their lives (‘high-risk inferences’), have the right to ask for justification by the appropriate data manager to disclose whether said inference has been produced through justifiable and explainable protocols. When used as part of a decision or action that concerns migrants and refugees, the algorithm managers and engineers have to produce a public explanation that outlines the rationale for using the dataset to produce said output, whether the form of inference and modeling is truly appropriate for said decision, and whether the output is sufficiently robust and statistically meaningful within the legal and legislative parameters set by that country’s laws and regulations.

To conclude, high-velocity social media data streams contain significant potential for emergency monitoring and forecasting; however, many contemporary examples of such attempts remain ethically and legally problematic. This problem is not limited to the use of social media data, but concerns several other components of big data migration and conflict forecasting practices, such as flawed inferences, opaque decisions, and policy-engineering mismatch between thresholds that generate profiling, monitoring, and scenario-building. This paper has argued that the use of social media data can exacerbate existing problems with explainability, interpretability, and transparency of A.I. forecasting and decision systems that deal with violence and forced migration. This paper has argued that the use of social media data can exacerbate existing problems with explainability, interpretability, and transparency of A.I. forecasting and decision systems that deal with violence and forced migration. Large data injection through the inclusion of social media data into self-learning A.I. monitoring systems generate a false sense of sophistication that forces decision-makers to disregard the validity, representativeness, and robustness of such systems. Since they are self-learning systems, learning classifiers that produce decision options become detached from the people they is monitoring, become unreachable from the migrants’ standpoint, and can potentially endanger vulnerable populations under duress.

Third, since social media data contains rapidly changing stances, discourses, and narratives, A.I. systems that use such data as training input risks steering outside the confines of legal and parliamentary oversight. While courts and legislators may ethically and legally approve a self-learning algorithm initially, weeks and months after these algorithms start learning on their own, they have a great likelihood of straying into an oversight ‘gray zone’. Further problems arise from the fact that most A.I. migration and conflict monitoring systems have suboptimal data cleaning practices. This renders models vulnerable to bot campaigns and disinformation and risks bias and discrimination against migrant

populations. Finally, although social media data brings immense data diversity to help teach machine learning classifiers, its sheer scale renders retrospective troubleshooting very difficult, especially when these systems make a costly flawed decision and reverse-engineering becomes essential to find out which data streams caused the problem. Although it is still early, once these issues are discussed and potentially resolved, social media data can be used ethically and cleanly in future A.I. migration and conflict monitoring tasks.

An important step in these directions were taken by the European Data Protection Supervisor (EDPS) in its consultation with the European Asylum Support Office (EASO) (D(2019) 1961 C 2018-1083), where social media monitoring for migration prediction was deemed illegal within the parameters of the EU law. However, it is important to keep in mind that the EDPS conclusion is only limited to EASO's operations, and asserts that tracking migration within the context of migrant smuggling and human trafficking monitoring are beyond the scope of EASO, and therefore, leaves a gray area for other EU agencies that may be monitoring social media for migration on the pretext that such monitoring is done to 'prevent crime'.

Still, however, there is significant discussion space as to what extent EDPS judgement will have a binding effect on other EU nations' daily border protection practices, or form a model to emulate for the border protection agencies of the rest of the world. As the competition to train more accurate models—either in conflict or migration prediction—social media data will remain a controversial, yet an increasingly popular choice, as the debate over how best to use it ethically and to meet strategic objectives will likely continue in the following years.

Funding: This research was partially funded by the Scientific and Technological Research Council of Turkey, ARDEB 1001 Program, Grant Number 120K986; and The Science Academy Society of Turkey: 2021-BAGEP Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest

References

- (Achieme 2021) Achieme, E. Tendayi. 2021. Digital Racial Borders. *American Journal of International Law* 115: 333–38.
- (Agrawal et al. 2018) Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business Review Press.
- (Ahmad et al. 2021) Ahmad, Wasim, Ali M. Kutan, and Smarth Gupta. 2021. Black Swan Events and COVID-19 Outbreak: Sector Level Evidence from the US, UK, and European Stock Markets. *International Review of Economics & Finance* 75: 546–57. <https://doi.org/10.1016/j.iref.2021.04.007>.
- (Alexander et al. 2022) Alexander, Monica, Kivan Polimis, and Emilio Zagheni. 2022. Combining Social Media and Survey Data to Nowcast Migrant Stocks in the United States. *Population Research and Policy Review* 41: 1–28. <https://doi.org/10.1007/s11113-020-09599-3>.
- (Ash and Obradovich 2020) Ash, Konstantin, and Nick Obradovich. 2020. Climatic Stress, Internal Migration, and Syrian Civil War Onset. *Journal of Conflict Resolution* 64: 3–31. <https://doi.org/10.1177/0022002719864140>.
- (Banerjee and MacKay 2020) Banerjee, Kiran, and Joseph MacKay. 2020. Communities of Practice, Impression Management, and Great Power Status: Military Observers in the Russo-Japanese War. *European Journal of International Security* 5: 274–93. <https://doi.org/10.1017/eis.2020.11>.
- (Batrinca and Treleaven 2015) Batrinca, Bogdan, and Philip C. Treleaven. 2015. Social Media Analytics: A Survey of Techniques, Tools and Platforms. *AI & SOCIETY* 30: 89–116. <https://doi.org/10.1007/s00146-014-0549-4>.
- (Baum and Zhukov 2015) Baum, Matthew A., and Yuri M. Zhukov. 2015. Filtering Revolution: Reporting Bias in International Newspaper Coverage of the Libyan Civil War. *Journal of Peace Research* 52: 384–400. <https://doi.org/10.1177/0022343314554791>.
- (Basu and Pearlman 2017) Basu, Sunkaya, and Sarah Pearlman. 2017. Violence and migration: evidence from Mexico's drug war. *IZA Journal of Development and Migration* 7: 1–29.
- (Beduschi 2021) Beduschi, Ana. 2021. International Migration Management in the Age of Artificial Intelligence. *Migration Studies* 9: 576–96. <https://doi.org/10.1093/migration/mnaa003>.

- (Beigi et al. 2018) Beigi, Ghazaleh, Kai Shu, Yanchao Zhang, and Huan Liu. 2018. Securing Social Media User Data: An Adversarial Approach. In *Proceedings of the 29th on Hypertext and Social Media*. New York: Association for Computing Machinery, pp. 165–73. <https://doi.org/10.1145/3209542.3209552>.
- (Bircan and Korkmaz 2021) Bircan, Tuba, and Emre Eren Korkmaz. 2021. Big Data for Whose Sake? Governing Migration through Artificial Intelligence. *Humanities and Social Sciences Communications* 8: 1–5. <https://doi.org/10.1057/s41599-021-00910-x>.
- (Blair and Sambanis 2020) Blair, Robert A., and Nicholas Sambanis. 2020. Forecasting Civil Wars: Theory and Structure in an Age of ‘Big Data’ and Machine Learning. *Journal of Conflict Resolution* 64: 1885–915. <https://doi.org/10.1177/0022002720918923>.
- (Bruce 2001) Bruce, Beverlee. 2001. Toward Mediating the Impact of Forced Migration and Displacement Among Children Affected by Armed Conflict. *Journal of International Affairs* 55: 35–57.
- (Buhmann and Fieseler 2021) Buhmann, Alexander, and Christian Fieseler. 2021. Towards a Deliberative Framework for Responsible Innovation in Artificial Intelligence. *Technology in Society* 64: 101475. <https://doi.org/10.1016/j.techsoc.2020.101475>.
- (Bunce et al. 2019) Bunce, Mel, Martin Scott, and Kate Wright. 2019. Humanitarian Journalism. In *Oxford Research Encyclopedia of Communication*. Oxford: Oxford Research Encyclopedia of Communication. <https://doi.org/10.1093/acrefore/9780190228613.013.821>.
- (Burke et al. 2015) Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. Climate and Conflict. *Annual Review of Economics* 7: 577–617. <https://doi.org/10.1146/annurev-economics-080614-115430>.
- (Carammia et al. 2022) Carammia, Marcello, Stefano Maria Iacus, and Teddy Wilkin. 2022. Forecasting Asylum-Related Migration Flows with Machine Learning and Data at Scale. *Scientific Reports* 12: 1457. <https://doi.org/10.1038/s41598-022-05241-8>.
- (Castets-Renard 2019) Castets-Renard, Celine. 2019. Accountability of Algorithms in the GDPR and beyond: A European Legal Framework on Automated Decision-Making. *Fordham Intellectual Property, Media & Entertainment Law Journal* 30: 91.
- (Chojnacki et al. 2012) Chojnacki, Sven, Christian Ickler, Michael Spies, and John Wiesel. 2012. Event data on armed conflict and security: New perspectives, old challenges, and some solutions. *International Interactions* 38: 382–401.
- (Chomsky and Herman 2002). Chomsky, Noam, and Edward Herman. 2002. A propaganda model. In *Manufacturing Consent: The Political Economy of the Mass Media*, 2nd ed. New York: Pantheon Books.
- (Christiansen et al. 2021) Christiansen, Rune, Matthias Baumann, Tobias Kuemmerle, Miguel D. Mahecha, and Jonas Peters. 2021. Toward Causal Inference for Spatio-Temporal Data: Conflict and Forest Loss in Colombia. *Journal of the American Statistical Association* 117: 591–601. <https://doi.org/10.1080/01621459.2021.2013241>.
- (Chu et al. 2016) Chu, Xu, Ihab F. Ilyas, Sanjay Krishnan, and Jiannan Wang. 2016. Data Cleaning: Overview and Emerging Challenges. In *Proceedings of the 2016 International Conference on Management of Data*. New York: Association for Computing Machinery, pp. 2201–6. <https://doi.org/10.1145/2882903.2912574>.
- (Coglianese and Lehr 2016) Coglianese, Cary, and David Lehr. 2016. Regulating by Robot: Administrative Decision Making in the Machine-Learning Era. *Georgetown Law Journal* 105: 1147.
- (Colombo 2018) Colombo, Monica. 2018. The Representation of the ‘European Refugee Crisis’ in Italy: Domopolitics, Securitization, and Humanitarian Communication in Political and Media Discourses. *Journal of Immigrant & Refugee Studies* 16: 161–78. <https://doi.org/10.1080/15562948.2017.1317896>.
- (Conte and Migali 2019) Conte, Alessandra, and Silvia Migali. 2019. The role of conflict and organized violence in international forced migration. *Demographic Research* 41: 393–424.
- (Crush 2013) Crush, Jonathan. 2013. Linking food security, migration and development. *International Migration* 51: 61–75.
- (Cunneen et al. 2019) Cunneen, Martin, Martin Mullins, Finbarr Murphy, and Seán Gaines. 2019. Artificial Driving Intelligence and Moral Agency: Examining the Decision Ontology of Unavoidable Road Traffic Accidents through the Prism of the Trolley Dilemma. *Applied Artificial Intelligence* 33: 267–93. <https://doi.org/10.1080/08839514.2018.1560124>.
- (de Leon 2015) de Leon, Jason. 2015. *The Land of Open Graves: Living and Dying on the Migrant Trail*. Oakland: University of California Press.
- (Dekker et al. 2018) Dekker, Rianne, Godfried Engbersen, Jeanine Klaver, and Hanna Vonk. 2018. Smart Refugees: How Syrian Asylum Migrants Use Social Media Information in Migration Decision-Making. *Social Media + Society* 4: 2056305118764439. <https://doi.org/10.1177/2056305118764439>.
- (Demarest and Langer 2018). Demarest, Leila, and Arnim Langer. 2018. The study of violence and social unrest in Africa: a comparative analysis of three conflict event datasets. *African Affairs* 117: 310–25.
- (Dowding 2021) Dowding, Keith. 2021. Why Forecast? The Value of Forecasting to Political Science. *PS: Political Science & Politics* 54: 104–6. <https://doi.org/10.1017/S104909652000133X>.
- (Eck 2012) Eck, Kristine. 2012. In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets. *Cooperation and Conflict* 47: 124–41. <https://doi.org/10.1177/0010836711434463>.
- (Efendi et al. 2018) Efendi, Riswan, Nureize Arbaiy, and Mustafa Mat Deris. 2018. A New Procedure in Stock Market Forecasting Based on Fuzzy Random Auto-Regression Time Series Model. *Information Sciences* 441: 113–32. <https://doi.org/10.1016/j.ins.2018.02.016>.
- (Elkin-Koren 2020) Elkin-Koren, Niva. 2020. Contesting Algorithms: Restoring the Public Interest in Content Filtering by Artificial Intelligence. *Big Data & Society* 7: 2053951720932296. <https://doi.org/10.1177/2053951720932296>.
- (Epstein 2010) Epstein, Andrew. 2010. Education refugees and the spatial politics of childhood vulnerability. *Childhood in Africa* 2: 16–25.

- (Eriksson 2018) Eriksson, Mats. 2018. Lessons for Crisis Communication on Social Media: A Systematic Review of What Research Tells the Practice. *International Journal of Strategic Communication* 12: 526–51. <https://doi.org/10.1080/1553118X.2018.1510405>.
- (Flach 2012) Flach, Peter. 2012. *Machine Learning*, 1st ed. Cambridge: Cambridge University Press.
- (Gohdes 2018) Gohdes, Anita R. 2018. Studying the Internet and Violent Conflict. *Conflict Management and Peace Science* 35: 89–106. <https://doi.org/10.1177/0738894217733878>.
- (Humphrey 2013) Humphrey, Michael. 2013. Migration, security and insecurity. *Journal of Intercultural Studies* 34: 178–95.
- (Jaidka 2022) Jaidka, Kokil. 2022. Cross-Platform- and Subgroup-Differences in the Well-Being Effects of Twitter, Instagram, and Facebook in the United States. *Scientific Reports* 12: 3271. <https://doi.org/10.1038/s41598-022-07219-y>.
- (Jain et al. 2020) Jain, Abhinav, Hima Patel, Lokesh Nagalapatti, Nitin Gupta, Sameep Mehta, Shanmukha Guttula, Shashank Mujumdar, Shazia Afzal, Ruhi Sharma Mittal, and Vitobha Munigala. 2020. Overview and Importance of Data Quality for Machine Learning Tasks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, pp. 3561–62. <https://doi.org/10.1145/3394486.3406477>.
- (Kaufhold et al. 2020) Kaufhold, Marc-André, Nicola Rupp, Christian Reuter, and Matthias Habdank. 2020. Mitigating Information Overload in Social Media during Conflicts and Crises: Design and Evaluation of a Cross-Platform Alerting System. *Behaviour & Information Technology* 39: 319–42. <https://doi.org/10.1080/0144929X.2019.1620334>.
- (Kirkpatrick 2016) Kirkpatrick, Keith. 2016. Battling Algorithmic Bias: How Do We Ensure Algorithms Treat Us Fairly? *Communications of the ACM* 59: 16–17. <https://doi.org/10.1145/2983270>.
- (Knightley 2002) Knightley, Phillip. 2002. Journalism, Conflict and War: An Introduction. *Journalism Studies* 3: 167–71. <https://doi.org/10.1080/14616700220129946>.
- (Lachlan et al. 2016) Lachlan, Kenneth A., Patric R. Spence, Xialing Lin, Kristy Najarian, and Maria Del Greco. 2016. Social Media and Crisis Management: CERC, Search Strategies, and Twitter Content. *Computers in Human Behavior* 54: 647–52. <https://doi.org/10.1016/j.chb.2015.05.027>.
- (Larson and Lewis 2018) Larson, Jennifer M., and Janet I. Lewis. 2018. Rumors, Kinship Networks, and Rebel Group Formation. *International Organization* 72: 871–903. <https://doi.org/10.1017/S0020818318000243>.
- (Lee and Maslog 2005) Lee, Seow Ting, and Crispin C. Maslog. 2005. War or Peace Journalism? Asian Newspaper Coverage of Conflicts. *Journal of Communication* 55: 311–29. <https://doi.org/10.1111/j.1460-2466.2005.tb02674.x>.
- (Lichtenheld 2020) Lichtenheld, Adam G. 2020. Explaining Population Displacement Strategies in Civil Wars: A Cross-National Analysis. *International Organization* 74: 253–94. <https://doi.org/10.1017/S0020818320000089>.
- (Lösch et al. 2017) Lösch, Andreas, Reinhard Heil, and Christoph Schneider. 2017. Responsibilization through Visions. *Journal of Responsible Innovation* 4: 138–56. <https://doi.org/10.1080/23299460.2017.1360717>.
- (Lozano-Gracia et al. 2010) Lozano-Gracia, Nancy, Gianfranco Piras, Ana Maria Ibáñez, and Geoffrey J. D. Hewings. 2010. The Journey to Safety: Conflict-Driven Migration Flows in Colombia. *International Regional Science Review* 33: 157–80. <https://doi.org/10.1177/0160017609336998>.
- (Lubeck et al. 2003) Lubeck, M., D. Geppert, and K. Nienartowicz. 2003. An Overview of a Large-Scale Data Migration. Paper presented at 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies, 2003 (MSST 2003), San Diego, CA, USA, April 7–10, pp. 49–55. <https://doi.org/10.1109/MASS.2003.1194835>.
- (Milano et al. 2021) Milano, Silvia, Brent Mittelstadt, Sandra Wachter, and Christopher Russell. 2021. Epistemic Fragmentation Poses a Threat to the Governance of Online Targeting. *Nature Machine Intelligence* 3: 466–72. <https://doi.org/10.1038/s42256-021-00358-3>.
- (Mittelstadt et al. 2016) Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The Ethics of Algorithms: Mapping the Debate. *Big Data & Society* 3: 2053951716679679. <https://doi.org/10.1177/2053951716679679>.
- (Molnar 2019a) Molnar, Petra. 2019a. New Technologies in Migration: Human Rights Impacts. *Forced Migration Review* 61: 7–9.
- (Molnar 2019b) Molnar, Petra. 2019b. Technology on the Margins: AI and Global Migration Management from a Human Rights Perspective. *Cambridge International Law Journal* 8: 305–30. <https://doi.org/10.4337/cilj.2019.02.07>.
- (Molnar 2021) Molnar, Petra. 2021. Chapter 10: Robots and refugees: the human rights impacts of artificial intelligence and automated decision-making in migration. In *Research Handbook on International Migration and Digital Technology*. Cheltenham: Edward Elgar Publishing.
- (Mullen and Klaehn 2010) Mullen, Andrew, and Jeffery Klaehn. 2010. The Herman–Chomsky propaganda model: A critical approach to analysing mass media behaviour. *Sociology Compass* 4: 215–29.
- (Nam 2006) Nam, Taehyun. 2006. What you use matters: Coding protest data. *PS: Political Science & Politics* 39: 281–87.
- (Ning et al. 2019) Ning, Yue, Liang Zhao, Feng Chen, Chang-Tien Lu, and Huzefa Rangwala. 2019. Spatio-Temporal Event Forecasting and Precursor Identification. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, pp. 3237–38. <https://doi.org/10.1145/3292500.3332291>.
- (Pavlyshenko 2019) Pavlyshenko, Bohdan M. 2019. Machine-Learning Models for Sales Time Series Forecasting. *Data* 4: 15. <https://doi.org/10.3390/data4010015>.
- (Pécoud and de Guchteneire 2006) Pécoud, Antoine, and Paul de Guchteneire. 2006. International Migration, Border Controls and Human Rights: Assessing the Relevance of a Right to Mobility. *Journal of Borderlands Studies* 21: 69–86. <https://doi.org/10.1080/08865655.2006.9695652>.

- (Pierskalla and Hollenbach 2013) Pierskalla, Jan H., and Florian M. Hollenbach. 2013. Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa. *American Political Science Review* 107: 207–24. <https://doi.org/10.1017/S0003055413000075>.
- (Quinn et al. 2018) Quinn, John A., Marguerite M. Nyhan, Celia Navarro, Davide Coluccia, Lars Bromley, and Miguel Luengo-Oroz. 2018. Humanitarian Applications of Machine Learning with Remote-Sensing Data: Review and Case Study in Refugee Settlement Mapping. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376: 20170363. <https://doi.org/10.1098/rsta.2017.0363>.
- (Rahwan 2018) Rahwan, Iyad. 2018. Society-in-the-loop: Programming the algorithmic social contract. *Ethics And Information Technology* 20: 5–14.
- (Rajabali et al. 2009) Rajabali, Alefiyah, Omer Moin, Amna S. Ansari, Mohammad R. Khanani, and Syed H. Ali. 2009. Communicable disease among displaced Afghans: refuge without shelter. *Nature Reviews Microbiology* 7: 609–14. <https://doi.org/10.1038/nrmicro2176>
- (Ravi 2005) Ravi, Narasimhan. 2005. Looking beyond Flawed Journalism: How National Interests, Patriotism, and Cultural Values Shaped the Coverage of the Iraq War. *Harvard International Journal of Press/Politics* 10: 45–62. <https://doi.org/10.1177/1081180X05275765>.
- (Renda 2019) Renda, Andrea. 2019. *Artificial Intelligence. Ethics, Governance and Policy Challenges*. CEPS Centre for European Policy Studies. Available online: <https://www.ceeol.com/search/book-detail?id=829907> (accessed on 19 May 2022).
- (Romei and Ruggieri 2014) Romei, Andrea, and Salvatore Ruggieri. 2014. A Multidisciplinary Survey on Discrimination Analysis. *The Knowledge Engineering Review* 29: 582–638. <https://doi.org/10.1017/S0269888913000039>.
- (Salah 2022) Salah, Albert Ali. 2022. Can Big Data Deliver Its Promises in Migration Research? *International Migration* 60: 252–55. <https://doi.org/10.1111/imig.12984>.
- (Salehyan 2007) Salehyan, Idean. 2007. Refugees and the Study of Civil War. *Civil Wars* 9: 127–41. <https://doi.org/10.1080/13698240701207286>.
- (Sampedro et al. 2022) Sampedro, Víctor, F. Javier López-Ferrández, and Patricia Hidalgo. 2022. Digital Disintermediation, Technical and National Sovereignty: The Internet Shutdown of Catalonia's 'Independence Referendum'. *European Journal of Communication* 37: 127–44. <https://doi.org/10.1177/02673231211012143>.
- (Sanchez 2017) Sanchez, Gabriella. 2017. Critical Perspectives on Clandestine Migration Facilitation: An Overview of Migrant Smuggling Research. *Journal on Migration and Human Security* 5: 9–27. <https://doi.org/10.1177/233150241700500102>.
- (Schon 2019) Schon, Justin. 2019. Motivation and Opportunity for Conflict-Induced Migration: An Analysis of Syrian Migration Timing. *Journal of Peace Research* 56: 12–27. <https://doi.org/10.1177/0022343318806044>.
- (Schroeder et al. 2013) Schroeder, Ashley, Lori Pennington-Gray, Holly Donohoe, and Spiro Kioussis. 2013. Using Social Media in Times of Crisis. *Journal of Travel & Tourism Marketing* 30: 126–43. <https://doi.org/10.1080/10548408.2013.751271>.
- (Sejnowski 2020) Sejnowski, Terrence J. 2020. The Unreasonable Effectiveness of Deep Learning in Artificial Intelligence. *Proceedings of the National Academy of Sciences* 117: 30033–38. <https://doi.org/10.1073/pnas.1907373117>.
- (Selby and Hoffmann 2012) Selby, Jan, and Clemens Hoffmann. 2012. Water Scarcity, Conflict, and Migration: A Comparative Analysis and Reappraisal. *Environment and Planning C: Government and Policy* 30: 997–1014. <https://doi.org/10.1068/c11335j>.
- (Shah 2018) Shah, Hetan. 2018. Algorithmic Accountability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376: 20170362. <https://doi.org/10.1098/rsta.2017.0362>.
- (Shan et al. 2019) Shan, Siqing, Feng Zhao, Yigang Wei, and Mengni Liu. 2019. Disaster Management 2.0: A Real-Time Disaster Damage Assessment Model Based on Mobile Social Media Data—A Case Study of Weibo (Chinese Twitter). *Safety Science* 115: 393–413. <https://doi.org/10.1016/j.ssci.2019.02.029>.
- (Shneiderman 2016) Shneiderman, Ben. 2016. The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight. *Proceedings of the National Academy of Sciences USA* 113: 13538–40. <https://doi.org/10.1073/pnas.1618211113>.
- (Silver and Andrey 2019) Silver, Amber, and Jean Andrey. 2019. Public Attention to Extreme Weather as Reflected by Social Media Activity. *Journal of Contingencies and Crisis Management* 27: 346–58. <https://doi.org/10.1111/1468-5973.12265>.
- (Steele 2009) Steele, Abbey. 2009. Seeking Safety: Avoiding Displacement and Choosing Destinations in Civil Wars. *Journal of Peace Research* 46: 419–29. <https://doi.org/10.1177/0022343309102660>.
- (Stewart and Wilson 2016) Stewart, Margaret C., and B. Gail Wilson. 2016. The Dynamic Role of Social Media during Hurricane #Sandy: An Introduction of the STREMI Model to Weather the Storm of the Crisis Lifecycle. *Computers in Human Behavior* 54: 639–46. <https://doi.org/10.1016/j.chb.2015.07.009>.
- (Sowers and Weinthal 2021) Sowers, Jeannie, and Erika Weinthal. 2021. Humanitarian challenges and the targeting of civilian infrastructure in the Yemen war. *International Affairs* 97: 157–77.
- (Taddeo and Floridi 2018) Taddeo, Mariarosaria, and Luciano Floridi. 2018. How AI Can Be a Force for Good. *Science* 361: 751–52. <https://doi.org/10.1126/science.aat5991>.
- (Tarasyev et al. 2018) Tarasyev, Alexandr A., Gavriil A. Agarkov, and Seyed Iman Hosseini. 2018. Machine Learning in Labor Migration Prediction. *AIP Conference Proceedings* 1978: 440004. <https://doi.org/10.1063/1.5044033>.
- (Tellez 2022) Tellez, Juan Fernando. 2022. Land, Opportunism, and Displacement in Civil Wars: Evidence from Colombia. *American Political Science Review* 116: 403–18. <https://doi.org/10.1017/S0003055421001003>.

- (Townsend and Wallace 2017) Townsend, Leanne, and Claire Wallace. 2017. The Ethics of Using Social Media Data in Research: A New Framework. In *The Ethics of Online Research*. Edited by Kandy Woodfield. Advances in Research Ethics and Integrity. Emerald Publishing Limited, vol. 2, pp. 189–207. <https://doi.org/10.1108/S2398-601820180000002008>.
- (Vallor and Bekey 2017) Vallor, Shannon, and George Bekey. 2017. Artificial Intelligence and the Ethics of Self-Learning Robots. In *Robot Ethics 2.0*. Edited by Patrick Lin. Oxford: Oxford University Press. <https://scholarcommons.scu.edu/phi/64>.
- (Véliz 2021) Véliz, Carissa. 2021. Moral Zombies: Why Algorithms Are Not Moral Agents. *AI & SOCIETY* 36: 487–97. <https://doi.org/10.1007/s00146-021-01189-x>.
- (Veronese et al. 2019) Veronese, Alexandre, Alessandra Silveira, and Amanda Nunes Lopes Espiñeira Lemos. 2019. Artificial Intelligence, Digital Single Market and the Proposal of a Right to Fair and Reasonable Inferences: A Legal Issue between Ethics and Techniques. *UNIO—EU Law Journal* 5: 75–91. <https://doi.org/10.21814/unio.5.2.2294>.
- (Vives 2017) Vives, Luna. 2017. The European Union–West African Sea Border: Anti-Immigration Strategies and Territoriality. *European Urban and Regional Studies* 24: 209–24. <https://doi.org/10.1177/0969776416631790>.
- (Wachter and Mittelstadt 2019) Wachter, Sandra, and Brent Mittelstadt. 2019. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review* 2019: 494.
- (Wachter et al. 2017) Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology (Harvard JOLT)* 31: 841.
- (Weidmann 2013) Weidmann, Nils B. 2013. The Higher the Better? The Limits of Analytical Resolution in Conflict Event Datasets. *Cooperation and Conflict* 48: 567–76. <https://doi.org/10.1177/0010836713507670>.
- (Weidmann 2016) Weidmann, Nils B. 2016. A Closer Look at Reporting Bias in Conflict Event Data. *American Journal of Political Science* 60: 206–18. <https://doi.org/10.1111/ajps.12196>.
- (Willekens 2018) Willekens, Frans. 2018. Towards Causal Forecasting of International Migration. *Vienna Yearbook of Population Research* 16: 199–218.
- (Zarsky 2013) Zarsky, Tal. 2013. *Transparent Predictions*. SSRN Scholarly Paper 2324240. Rochester: Social Science Research Network. Available online: <https://papers.ssrn.com/abstract=2324240> (accessed on 19 May 2022).
- (Zednik 2021) Zednik, Carlos. 2021. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology* 34: 265–88. <https://doi.org/10.1007/s13347-019-00382-7>.
- (Zeitoff 2018) Zeitoff, Thomas. 2018. Does Social Media Influence Conflict? Evidence from the 2012 Gaza Conflict. *Journal of Conflict Resolution* 62: 29–63. <https://doi.org/10.1177/0022002716650925>.
- (Zhang et al. 2018) Zhang, Jinxue, Jingchao Sun, Rui Zhang, Yanchao Zhang, and Xia Hu. 2018. Privacy-Preserving Social Media Data Outsourcing. Paper presented at IEEE INFOCOM 2018—IEEE Conference on Computer Communications, Honolulu, HI, USA, April 16–19, pp. 1106–14. <https://doi.org/10.1109/INFOCOM.2018.8486242>.