

# Northumbria Research Link

Citation: Mahoney, Jamie, Le Louvier, Kahina, Lawson, Shaun, Bertel, Diotima and Ambrosetti, Elena (2022) Ethical considerations in social media analytics in the context of migration: lessons learned from a Horizon 2020 project. Research Ethics Review. p. 174701612210875. ISSN 1747-0161 (In Press)

Published by: SAGE

URL: <https://doi.org/10.1177/17470161221087542>  
<<https://doi.org/10.1177/17470161221087542>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/48730/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

# Ethical considerations in social media analytics in the context of migration: lessons learned from a Horizon 2020 project

Research Ethics

1–15

© The Author(s) 2022

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/17470161221087542

[journals.sagepub.com/home/rea](https://journals.sagepub.com/home/rea)

**Jamie Mahoney**   
**Kahina Le Louvier**  
**Shaun Lawson**  
Northumbria University, UK

**Diotima Bertel**  
SYNYO, Austria

**Elena Ambrosetti**  
Sapienza University of Rome, Italy

## Abstract

The ubiquitous use of social platforms across the globe makes them attractive options for investigating social phenomena including migration. However, the use of social media data raises several crucial ethical issues around the areas of informed consent, anonymity and profiling of individuals, which are particularly sensitive when looking at a population such as migrants, which is often considered as ‘vulnerable’. In this paper, we discuss how the opportunities and challenges related to social media research in the context of migration impact on the development of large-scale scientific projects. Building on the EU-funded research project PERCEPTIONS, we explore the concrete challenges experienced in

## Corresponding author:

Jamie Mahoney, Department of Computer and Information Sciences, Northumbria University, Sutherland Building, Newcastle upon Tyne NE1 8ST, UK.

Email: [jamie.mahoney@northumbria.ac.uk](mailto:jamie.mahoney@northumbria.ac.uk)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

such projects regarding profiling, informed consent, bias, data sharing and ethical approval procedures, as well as the strategies used to mitigate them. We draw from lessons learned in this project to discuss implications and recommendations to researchers, funders and university ethics review panels. This paper contributes to the growing discussion on the ethical challenges associated with big social data research projects on migration by highlighting concrete aspects stakeholders should be looking for and questioning when involved in such large-scale scientific projects where collaboration, data sharing and transformation and practicalities are of importance.

## Keywords

Social media, ethics, migration

## Introduction

In 2019, the European Border and Coast Guard Agency (FRONTEX) published a call for tenders soliciting the provision of services for monitoring and forecasting irregular migratory movement at the borders of Europe via social media.<sup>1</sup> This 400,000€ contract invited the development of tools allowing the collection and analysis of the social media data published by a variety of migration-related actors, including migrants, smugglers, traffickers, but also civil society and diaspora communities in Europe. Following the launch of this call, the NGO Privacy International published an enquiry over the legality of such services, and whether appropriate checks and approvals had been sought.<sup>2</sup> Soon after this enquiry, the tender was cancelled, raising questions about the ethical implications of research in this arena. This event illustrates the complexity of the use of social media data in migration research, a highly political and sensitive area where the opportunities provided by these popular platforms often contradict the multiple ethical challenges they carry, and where the interests of different stakeholders can come into conflict.

If this specific call for tenders was cancelled, the complex ethical considerations that accompany big social media data analytics in the context of migration persist and need careful consideration (Townsend and Wallace, 2017). Yet, if ethical guidelines regarding social media analytics on the one hand (Association of Internet Researchers, 2019), and migration research on the other (European Commission, 2020), have emerged, these two aspects tend to be covered separately and are not jointly embedded within the conception and development of international research project.

To explore these ethical challenges and put forward possible mitigation strategies, we draw together the lessons learned from PERCEPTIONS,<sup>3</sup> a 3-year project funded as part of the EU Horizon 2020's 'improving border and external security' research and development strand. This project aims to investigate the impact of perceptions and narratives about Europe on migration behaviours and associated

risks to migrants, as well as host countries. In addition to surveys, interviews and focus groups with migrants and practitioners working with migrants, the project uses social media platforms as relevant sources of data to investigate narratives and how they spread. As the project developed, various tensions arose between its initial aims, the intended use of social media data, and the ethical requirements of EU projects funded under the Horizon 2020 programme. Each research partner being based in a different country, these ethical challenges were further complicated by the heterogeneity between the different national regulations and organisational ethical processes that each had to comply with. Navigating these multiple ethical requirements and designing adequate mitigation strategies was a particularly complex process that caused several delays.

This paper provides a reflective overview of these challenges with the aim of helping research stakeholders to identify and anticipate some of the legal and ethical implications of using social media analytics in migration research. First, we outline the main ethical concerns associated with big data and social media research in a migration context. Then, we discuss the ethical challenges that we encountered as part of PERCEPTIONS, and the mitigation strategies implemented. Finally, we draw implications and recommendations to different stakeholders, including researchers, funding bodies and university ethics review panels. By highlighting key points of vigilance when it comes to social media analytics and migration, this paper provides concrete recommendations for establishing processes that ensure that ethical issues are addressed and mitigated in a way that does not limit the development of research projects. It also contributes to drawing together considerations around social media and migration, two areas of ethical concerns that are particularly sensitive and increasingly overlapping (Sandberg and Rossi, 2022).

## **Ethics in migration and social media research**

As international migration is a multifaceted and constantly changing phenomenon, rich and up to date data is needed to be able to inform relevant policies and actions. However, international migration is particularly difficult to document, leading to various knowledge gaps and datasets that are inconsistent across countries (Migration Data Portal, 2020). To address those gaps, researchers and agencies have increasingly resorted to the use of big data drawn from mobile phones and Internet sources (Sirbu et al., 2021). In addition to these sources, social media platforms have increasingly been considered as attractive options for studying migration, as they provide access to large amounts of real-time, structured data, and can be used to investigate various aspects of migration. Such aspects include measuring, monitoring and forecasting migration movements (Martín et al., 2020), analysing public attitudes towards immigrants and immigration (Siapera et al., 2018) and

estimating levels of integration (Dubois et al., 2018). Thus, social media platforms are particularly valuable data sources to investigate the social determinants and the social impact of migration in both sending and receiving countries. However, the use of social media data raises a number of crucial ethical issues around the areas of informed consent, anonymity and profiling of individuals (Townsend and Wallace, 2017), which are particularly sensitive when looking at a population such as migrants, which is often considered as ‘vulnerable’.

Individuals who immigrate to a new country but do not hold the nationality from that country are particularly at risk of social exclusion, as access to basic aspects of life such as work, housing, social welfare, education or bank accounts can be restricted or even prevented (Yuval-Davis et al., 2018). The public disclosure of one’s immigration status can also lead to stigmatisation and hate speech (ENAR, 2018). Furthermore, individuals who have entered or remained in a country in an irregular way can be at risk of detention and deportation (Tazzioli, 2020). Thus, the definition of vulnerable persons as ‘those who lack the ability to make personal life choices, to make personal decisions, to maintain independence, and to self-determine’ (Moore and Miller, 1999: 1034) applies to individuals with various migration experiences. This includes subjects of human trafficking, individuals living in a country undocumented or with a short-term visa, and others whose movement across borders puts them at risk of persecution or precariousness (van Liempt and Bilger, 2012).

Categorising such groups as ‘vulnerable’ has immediate repercussions, as research with such populations needs to go beyond ‘procedural ethics’, that is, the formal rules that guide research design and certain aspects of fieldwork (Tomkinson, 2015). Informed consent is a core principle of this type of ethical process and has been the subject of various strategies and recommendations that highlight the importance of ensuring that vulnerable participants understand the implications of being involved in a research project (European Commission, 2020). If such principles and associated guidance are frequently discussed in relation to qualitative and traditional research methods (European Commission, 2020), these are not readily applicable to social media analytics and the specific ethical challenges they raise (Sandberg et al., 2022). Indeed, when analysing big data from social media platforms, it is not always possible or feasible to seek the data subjects’ explicit consent to the use of their data, meaning that individuals may not be aware of who uses their data and how it is used (Fiesler and Proferes, 2018).

When social media data related to migration is processed, the lack of informed consent can exacerbate the severity of potential risks, such as data breach, data de-anonymisation and profiling, as the disclosure of an individual’s name, location, network of family and friends, as well as other sensitive information including immigration status, can result in stigmatisation and hate crime. For instance, a picture posted on social media may be used by anti-migration groups to feed racist

campaigns (Dearden, 2015), and the profiling of migrant social media accounts to predict migration flows and close migration routes can lead people to go through even more dangerous border crossings (Dimitriadi, 2021). Thus, such use of social media data increases the power imbalance between the researchers and institutions that process data for their specific purposes, and those vulnerable individuals whose personal data is extracted (Bloemraad and Menjivar, 2022). Addressing these risks means balancing the rights of individuals with the benefits of the collective. This means that it is the responsibility of funders and researchers to take all the precautions necessary to identify and minimise any risk of harm. Yet, despite recent policies and guidelines around social media use (UK Research and Innovation, 2021), clear institutional processes and actionable recommendations to address these issues are still lacking (Taylor and Pagliari, 2018), especially in the context of migration research and of externally funded projects that involve multiple stakeholders, as we experienced as part of the PERCEPTIONS project. In this paper, we reflect on these issues to highlight concrete recommendations for the development of ethical research programmes.

## **Ethical challenges encountered in PERCEPTIONS**

In PERCEPTIONS, social media analytics are used to identify public narratives about Europe and migration to Europe on a global scale, and to analyse their content and distribution patterns. Social media data is retrieved from Twitter, a platform that offers broad opportunities to investigate public narratives, including those reflecting and influencing the migration imaginaries of various social groups. Tweets are retrieved based on a set of approximately 60 migration-related keywords, translated into a total of six languages, leading to approximately 450,000 tweets and retweets collected each day for a period of 3 months. Twitter data is used to perform three types of analyses: topic modelling, social network analysis and bot analysis. Topic modelling allows us to gain an understanding of the topics of conversation in a large, multilingual dataset. By analysing the network structure of the collected tweets, using community detection algorithms, we characterise the general flow of information, identify the key groups of actors, and analyse the topics discussed in the various groups and countries identified. Finally, implementing algorithms for bot detection and text analysis tools allows us to analyse the type of narratives vehiculated by bots.

The use of social media was a requirement set in the European Commission's call for projects, which reflects a recent trend for Horizon 2020 project to promote this approach. Yet, obtaining ethical approval from the European Commission was a challenge, as a variety of ethical concerns emerged in relation to profiling, informed consent, data sharing processes and ethical approval and data management procedures. Progressing these ethical considerations was an opportunity to

draw lessons that we discuss in this paper in order to provide stakeholders in similar research programmes with key points of vigilance and actionable recommendations.

### *Profiling*

Based on the wording of the European Commission's funding call, aspects of PERCEPTIONS initially involved investigating how migrants perceive Europe, and analyses of social media data was considered to be one potential way of achieving this. Such an approach would require profiling, a technique that consists in using automated means to categorise individuals according to their personal characteristics (EU GDPR, Article 4(2)).

Given the sensitive context of the project and the project funding call, the profiling of individuals as 'migrants', and particularly the impact that this could have, was a major concern. While it was not the intention to directly affect participants as a result of any potential profiling of their social media activity, members of the PERCEPTIONS consortium and the European Commission's expert review panel were acutely aware of the potential misuse and risk of harm of any reported research findings. The profiling of individuals as 'migrants' could potentially expose them to harm, including hate speech, detention, removal and for people fleeing persecution, potential pressures from homeland authorities on family members who remained there (Bloemraad and Menjivar, 2022). Moreover, PERCEPTIONS is funded under Horizon 2020's security strand, and the consortium includes security agencies. Yet, the profiling of migrants on social media used by European law enforcement agencies to detect and prevent migration arrivals could have resulted in adverse effects, pushing individuals into new and more dangerous migration routes (Dimitriadi, 2021). These unwanted consequences would have contradicted the principle of non-maleficence that the PERCEPTIONS project adheres to.

Thus, although some form of profiling of individual social media accounts may have facilitated a better understanding of the data, potentially leading to more insightful results and to a more direct response to the initial funding call, it was decided to avoid attempting to identify migrants entirely, and that any other form of profiling and reporting of results would be done at an aggregated level. For example, rather than understanding and reporting on the behaviour of specific accounts, this would be done by focussing on the behaviours of groups of accounts. Analyses within the project would naturally lead to groups of users being identified, and the tendency would be to try to understand who each of these users are, through profiling their names, locations, and other biographical information. However, it was decided, within the PERCEPTIONS project, to set an initial threshold of 10,000 account followers, with only accounts above the threshold

accessed and processed as necessary. In doing so, we sought to strike a balance between protecting the rights and anonymity of individuals, while still understanding the role of major, higher-profile Twitter accounts in the dissemination of narratives related to migration and migrants; in many cases, accounts with high follower counts are likely to belong to organisations, or ‘celebrities’.

### *Informed consent*

As the aim of the project is to identify public narratives about Europe and migration on a global scale, large amounts of tweets are collected using automated means. This means data is collected without directly interacting with these Twitter users, making gaining the informed consent of each of them problematic. Under EU GDPR (Article 14 Paragraph 5b), obtaining informed consent from data subjects is not required in a context such as social media analytics, where:

‘The provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes’.

However, it remains a point of discussion as to the extent that social media users are even aware that some of their data is made available to researchers. Studies have shown that once individuals are made aware that their social media activity may be included in research studies, they often state that such actions should not be allowed, at least without informing the individuals or requesting permission to include their data in the study (Fiesler and Proferes, 2018).

Thus, when working with social media data sources, it has to be assessed whether those persons actually intended to make their information public (e.g. in the light of the privacy settings or limited audience to which the data were made available). A key argument is the set of terms and conditions to which each user of a social media platform has to agree. Within these terms, there are often clauses addressing how the data may be accessed by third parties (including researchers). However, accessibility of the data is not enough; the data must have been made public to the extent that the data subjects do not have any reasonable expectation of privacy. In particular, when researching populations that are already marginalised, such as migrants, within a politically sensitive topic, conducting research without asking for consent could lead to an increased vulnerability of these populations.

In light of these points, several decisions were made in the PERCEPTIONS project regarding the collection, storage and analysis of social media data. First, it was decided that only explicitly public social media data would be collected and analysed. This, along with the availability and access restrictions of the public Application Programming Interface (API), resulted in the decisions to focus on Twitter as a social media data source, although this came with limitations with



regards to the demographics of the users whose tweets are retrieved (Blank, 2017). While this decision did not remove the consideration of informed consent entirely, it did further satisfy ethical guidelines from organisations such as British Psychological Society who state that observations (i.e. the collection of data) should be limited to public areas of platforms, where individuals have no perception or expectation of privacy (British Psychological Society, 2021).

To further minimise any potential risks to individuals, decisions were made to store collected data only for as long as was necessary, with suitable anonymisation or pseudonymisation techniques performed at the earliest opportunity. This aims to minimise the amount of time in which individuals would remain identifiable. Each of these steps help to minimise the risks that individual social media accounts may be exposed to, in a situation where gaining the informed consent of the account holders is not possible.

### *Data sharing and exchange between partners*

The PERCEPTIONS consortium is made up of partners based in multiple countries; some of these partners are based within the EU, while others are based outside of the EU. Further, the consortium consists of academics, practitioners, as well as law enforcement agencies. Having such a multi-disciplinary and multinational team allows us to bring together different perspectives, in terms of expertise, cultural and linguistic background, which helps mitigating some of the biases of the research, especially when it comes to the choice of keywords used to retrieve data and to the design of the analyses. However, it also introduces complexities regarding data storage and sharing between partners, both in terms of legal requirements (GDPR), as well as maintaining the anonymity and rights of participants, that may be compromised if law enforcement agency partners were able to access all of the raw data collected during the project.

To address these concerns, the PERCEPTIONS project introduced various measures. First, from a GDPR standpoint, the adequacy of the national legislations regarding data protection had to be determined. This step was an important one, as it determined what data could be transferred between which countries (and in which direction). As a result of this, it was decided that only EU project partners, and partners within the UK, which the European Commission recognised as providing an adequate level of data protection equivalent to EU GDPR, would be responsible for the social media data processing. This process also led to further discussions around encryption and storage methods.

Second, it was decided to restrict access to raw data, and the accompanying results, to only those directly involved in the analysis of the social media data. In doing so, partners such as law enforcement agencies had no access to the data being collected, nor the interim results of the analyses. First, this further protects the anonymity of any individuals included in the data collection and prevents

situations where, for example, it may have been possible to infer the migration status of individuals based on their social media data. Second, this also protects some project partners from being obliged to take action based on the data that they would have otherwise had access to.

As part of these processes, and to document the agreements made between partners regarding data collection, sharing and analysis within the social media research elements of the project, a joint controller document was created, and signed by the relevant partners within the project. This also outlines the responsibilities and liabilities of each partner in the event of a data breach or similar situation. This was done in conjunction with other ethical and data management processes, which are discussed in more detail in the following section.

### *Ethics and data management processes*

Projects involving multiple partners, such as PERCEPTIONS, will likely include multiple ethical approval processes, with each organisation having their own requirements. Some partners, such as universities, may have their own departmental and organisation-wide ethical review boards, with other organisations having no equivalent process. This mismatch between partners can lead to complications, depending on the requirements placed on the project partners by the funding body. The PERCEPTIONS project experienced this, with the university-based partners (for the most part) having existing ethical approval processes that were utilised. Other partners did not need to gain ethical approval from their own institutions, as the data and processes being conducted as part of the PERCEPTIONS project did not meet the threshold for requiring ethical approval. In cases where there were also no competent national ethics committees, these partners had no possibility to gain ethical approval. Further to this, some partners had to seek feedback and approval from multiple stakeholders within their own institutions, including ethics boards, data protection officers and legal departments.

As the PERCEPTIONS project required confirmation of ethical approval from each project partner, it was necessary to implement a system whereby individual partners could agree – formally – to abide by the ethical principles and constraints included in various project documentation, if they had no such ethical approval processes within their organisation. This ‘Declaration of Ethical Compliance’ was then approved and signed by an appropriate individual in each organisation, in lieu of a more ‘traditional’ ethical approval.

To maintain oversight of ethical and data management processes within the project consortium, an ethics board and data management task force were created. These consisted of various members of the project consortium, with additional external experts also forming part of the ethics board. These had oversight across various ethical processes and documentation for use within the project.

One such document overseen by the ethics board and data management task force was the Data Protection Impact Assessment (DPIA) conducted as part of the social media research within the PERCEPTIONS project. This was a necessary step, as determined by GDPR, as the collection of social media data was likely to include personal data, and its storage and processing could pose risks to the individual account holders.

The potential disruption for each of the points detailed above could have been minimised if the project consortium had a clearer understanding of the exact requirements and processes of both the funding body, and each partner's institutional processes. This could have, for example, led to the creation of the 'Declaration of Ethical Compliance' at the very beginning of the project, which could have been approved and signed by appropriate signatories at each partner institution. This shows the need for further structural guidance and processes that ensure ethical challenges are anticipated and addressed in the early stages of research programmes.

## **Considerations from multiple perspectives**

In the previous section, we outlined various ethical challenges that were encountered during the PERCEPTIONS project. In this section, we synthesise this discussion, highlighting the main points of consideration from the perspective of various interested parties: researchers, funding bodies and ethics review boards.

### **Researchers**

The previous section has highlighted numerous ethical challenges that were encountered during the PERCEPTIONS project. While many of these processes may not be avoidable, being aware of them, and planning for them from the beginning of the project could lead to a swifter and streamlined process.

First, researchers involved in the project should consider the standing of each partner's country regarding their data protection regulations, particularly their 'adequacy' in relation to EU GDPR. An understanding of what data can be collected and exchanged between which countries will help in planning the data collection and analyses aspects of the project and make undertaking processes such as the DPIA much easier.

Second, in situations where personal information may be collected, such as collecting social media content, researchers should consider elements such as informed consent, and whether it is necessary or indeed possible to collect this consent. While content that is clearly public may not require this consent, aspects of some social media platforms, or the introduction of new features or new platforms may make the identification of 'public' and 'private' spaces more complex.

This must be considered as early as possible in a project, as selecting the social media platform that provides the highest levels of data protection should always prevail, but may have consequences for the type of data collected and the research questions that can be addressed.

Third, where personal information may be collected, planning for data minimisation and anonymisation (where possible) from the outset will also be beneficial. Undertaking processes like the DPIA will often require such techniques to be implemented, to reduce the level of risk associated with data collection, exchange and processing. Further, following a ‘privacy-by-design’ and ‘data-protection-by-design’ approach from the project outset will not only benefit researchers by making processes such as the DPIA easier to navigate, but will also benefit project participants by providing additional safeguards to protect their privacy. Project partners may also, where necessary, consider asking their organisations’ legal department (or equivalent) to check the various Terms & Conditions statements from the relevant social media platforms. Designing studies that include time and resources to go through these processes from the outset will make things easier, along with the inclusion of individuals with relevant experience and expertise in such matters.

Finally, researchers should be mindful of the different ethical approval processes at each partner institution. As detailed in the previous section, funding bodies and other partner institutions can sometimes assume that each partner gaining ethical clearance from their own institution will be straightforward, when this can be a much more complex and time-consuming process. Such considerations should be made when designing project timelines, particularly as data collection and analysis will depend on gaining this ethical clearance. If planned for too early in the project timeline, such activities are likely to experience delays.

### *Funding bodies*

We have previously outlined some of the processes that were undertaken within the PERCEPTIONS project in terms of gaining ethical clearance from partner institutions, meeting the requirements of the funding body, and planning data collection and analysis processes that satisfied both the funding body and legal requirements such as EU GDPR.

To aid the timely progress of these projects, funding bodies should ensure that they make their requirements clear from the beginning of the project, allowing the project partners to meet these requirements in a timely manner. As some partners will have lengthy internal processes to achieve, for example, ethical clearance, or for data management plans to be approved by their organisation, introducing new requirements or delaying feedback until a particular pre-determined date can introduce unnecessary delays into the project timeline. Funding bodies should also

seek to develop a greater understanding of the differences between the various project partners, particularly in terms of their ethical approval processes and presence (or absence) of various legal representatives within the organisation. In doing so, this will allow funding bodies to aid the project researchers in successfully meeting the various requirements, and in a timely manner.

Further, funding bodies should ensure that the focus of research in each research topic does not contradict the ethics requirements and processes of their programme (such as Horizon 2020). Potential mismatches between these two can introduce unnecessary friction into the ethical review processes, as researchers attempt to balance meeting the research requirements of the topic and the ethical requirements and constraints of the funding programme. Related to this, the ethical and legal oversight processes from funding bodies should also be developed as the research areas they oversee develop. In areas of research that evolve quickly, such as social media research, those responsible for ensuring ethical and legal compliance should seek to keep their understanding of the key elements and principles of the research area up to date.

### *University ethics review panels*

Based on experiences within the PERCEPTIONS project, not all university ethics review processes recognise social media-based research as requiring ethical approval and oversight, due to the lack of direct interaction with potentially at-risk individuals. Given the potential implications of such research, especially if research findings are misused, such ethical approval processes should be further developed. Those responsible for such processes should consider moving their focus of research protection from ‘human subjects research’, whereby the review is driven by issues of informed consent, to ‘human harming research’, whereby a risk analysis is conducted, highlighting any potential harms stemming from the research, even when there are no direct interactions with potentially vulnerable participants (Carpenter and Dittirch, 2011). Through this change in focus, ‘researchers who might otherwise (even if incorrectly) feel no human is directly involved in the research study would be compelled to address the ethical implications of any harm to broader populations outside the immediate research project’ (Zimmer, 2018: 6).

## **Conclusion**

Big data provide a comprehensive, relatively cost-effective and timely complementary sources of data for the analysis of migration-related phenomena. These new sources of data, including social media platforms, can enrich migration research and knowledge with novel insights, helping to improve humanitarian

responses to migration and migration management. However, ‘technology is not inherently democratic, and its human rights impacts are particularly important to consider in humanitarian and forced migration contexts’ (Molnar, 2019: 7). Asymmetries in power between researchers and study participants, as well as between governments and people on the move, are particularly strong and can have severe consequences in social media analytics contexts where the principle of informed consent cannot be met (Bloemraad and Menjivar, 2022).

In discussing the ethical challenges that were encountered during the PERCEPTIONS project, and the mitigation measures undertaken, we have highlighted several issues that other research projects may also encounter, as well as areas of concerns that different stakeholders within research projects may wish to consider. We hope such reflections can help research projects meet their goals, while undertaking social media analyses in such a way that ethical and legal obligations are met, participants have their privacy respected, and processes such as ethical approval and Data Protection Impact Assessments can be undertaken in a timely and efficient manner. Although the power asymmetries inherent to social media analytics in the context of migration research may not disappear, such processes can ensure that all mitigation issues are in place to avoid causing harm to populations who are, often, already vulnerable.

## Funding

All articles in Research Ethics are published as open access. There are no submission charges and no Article Processing Charges as these are fully funded by institutions through Knowledge Unlatched, resulting in no direct charge to authors. For more information about Knowledge Unlatched please see here: <http://www.knowledgeunlatched.org>.

This research is conducted as part of the PERCEPTIONS H2020 project which has received funding from the European Union’s H2020 research & innovation programme under Grant Agreement No. 833870.

## ORCID iD

Jamie Mahoney  <https://orcid.org/0000-0003-4632-7212>

## Notes

1. <https://etendering.ted.europa.eu/cft/cft-display.html?cftId=5471>
2. <https://privacyinternational.org/advocacy/3289/privacywins-eu-border-guards-cancel-plans-spy-social-media-now>
3. <https://project.perceptions.eu/>

## References

Association of Internet Researchers (2019) Internet research: Ethical guidelines 3.0 association of internet researchers. Available at: <https://aoir.org/reports/ethics3.pdf> (accessed 7 July 2021).

- Blank G (2017) The digital divide among Twitter users and its implications for social research. *Social Science Computer Review* 35(6): 679–697.
- Bloemraad I and Menjivar C (2022) Precarious times, professional tensions: The ethics of migration research and the drive for scientific accountability. *International Migration Review* 56(1): 4–32.
- British Psychological Society (2021) Ethics guidelines for internet-mediated research. Available at: <https://www.bps.org.uk/news-and-policy/ethics-guidelines-internet-mediated-research> (accessed 5 July 2021).
- Carpenter KJ and Dittirch D (2011) Bridging the distance: Removing the technology buffer and seeking consistent ethical analysis in computer security research. In: *1st international digital ethics symposium*, Chicago, IL, 28 October 2011.
- Dearden L (2015) The fake refugee images that are being used to distort public opinion on asylum seekers. *The Independent*, 16 September.
- Dimitriadi A (2021) Countering smuggling of migrants through social media monitoring: Looking for a needle in a digital haystack. In: Sanchez G and Sanchez G (eds) *Beyond Networks, Militias and Tribes: Rethinking EU Counter-Smuggling Policy and Response*. EuroMesCo Policy Study No. 19. Barcelona: IEMed, 34–51.
- Dubois A, Zagheni E, Garimella K, et al. (2018) *Social Informatics*. St. Petersburg: Springer, 51–60.
- ENAR (2018) Racist crime and institutional racism in Europe. ENAR Shadow Report 2014–2018. Available at: [https://www.enar-eu.org/IMG/pdf/shadowreport2018\\_final.pdf](https://www.enar-eu.org/IMG/pdf/shadowreport2018_final.pdf) (accessed 6 July 2021).
- European Commission (2020) Guidance note—Research on refugees, asylum seekers & migrants. Available at: [https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/guide\\_research-refugees-migrants\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/guide_research-refugees-migrants_en.pdf) (accessed 9 July 2021).
- Fiesler C and Proferes N (2018) “Participant” perceptions of Twitter research ethics. *Social Media + Society* 4(1): 2056305118763366.
- Martín Y, Cutter SL, Li Z, et al. (2020) Using geotagged tweets to track population movements to and from Puerto Rico after Hurricane Maria. *Population and Environment* 42(1): 4–27.
- Migration Data Portal (2020) Migration data portal: International migration flows. Available at: <https://www.migrationdataportal.org/themes/international-migration-flows> (accessed 6 July 2021).
- Molnar P (2019) New technologies in migration: Human rights impacts. *Forced Migration Review* 61: 7–9.
- Moore LW and Miller M (1999) Initiating research with doubly vulnerable populations. *Journal of Advanced Nursing* 30(5): 1034–1040.
- Sandberg M, Mollerup NG and Rossi L (2022) Contrapuntal connectedness: Analysing relations between social media data and ethnography in digital migration studies. In: Sandberg M, Rossi L and Galis V (eds) *Research Methodologies and Ethical Challenges in Digital Migration Studies*. Cham: Palgrave Macmillan, 53–85.
- Sandberg M and Rossi L (2022) Caring for (Big) Data: An introduction to research methodologies and ethical challenges in digital migration studies. In: Sandberg M, Rossi L and Galis V (eds) *Research Methodologies and Ethical Challenges in Digital Migration Studies*. Cham: Palgrave Macmillan, 1–21.
- Siapera E, Boudourides M, Lenis S, et al. (2018) Refugees and network publics on Twitter: Networked framing, affect, and capture. *Social Media + Society* 4(1): 2056305118764437.
- Sîrbu A, Andrienko G, Andrienko N, et al. (2021) Human migration: The big data perspective. *International Journal of Data Science and Analytics* 11(4): 341–360.

- Taylor J and Pagliari C (2018) Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics* 14(2): 1–39.
- Tazzioli M (2020) Governing migrant mobility through mobility: Containment and dispersal at the internal frontiers of Europe. *Environment and Planning C: Politics and Space* 38(1): 3–19.
- Tomkinson S (2015) Doing fieldwork on state organizations in democratic settings: Ethical issues of research in refugee decision making. *Forum: Qualitative Social Research* 16(1). DOI: 10.17169/fqs-16.1.2201.
- Townsend L and Wallace C (2017) The ethics of using social media data in research: A new framework. In: Woodfield K (ed.) *The Ethics of Online Research*. Bingley: Emerald Limited, 189–207.
- UK Research and Innovation (2021) Research that may require a full ethics review. Available at: <https://beta.ukri.org/councils/esrc/guidance-for-applicants/research-ethics-guidance/ethics-reviews/research-that-may-require-full-ethics-review/> (accessed 6 July 2021).
- van Liempt and Bilger V (2012) Ethical challenges in research with vulnerable migrants. In: Vargas-Silva C (ed.) *Handbook of Research Methods in Migration*. Cheltenham: Edward Elgar, 451–466.
- Yuval-Davis N, Wemyss G and Cassidy K (2018) Everyday bordering, belonging and the reorientation of British immigration legislation. *Sociology* 52(2): 228–244.
- Zimmer M (2018) Addressing conceptual gaps in big data research ethics: An application of contextual integrity. *Social Media + Society* 4(2): 2056305118768300.