



Article

# Making up the predictable border: How bureaucracies legitimate data science techniques

new media & society

1–17

© The Author(s) 2023

Article reuse guidelines:

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

DOI: 10.1177/14614448231161276

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



**Burcu Baykurt** 

University of Massachusetts Amherst, USA

**Alphoncina Lyamuya** 

University of Southern California, USA

## Abstract

This article examines how claims to predictable borders via data science techniques are crafted in bureaucratic institutions. Through a case study of testing algorithmic systems at a transnational agency, we examine how humanitarian organizations reconcile the risks of predictive technologies with the benefits they claim to receive. Drawing on a content analysis of policy documents and interviews with humanitarian technologists, we identify three organizational strategies to justify working toward predictability: constantly seeking novel variables and data, maintaining ambiguity, and shifting models to adapt to changing circumstances. These strategies, we argue, sustain the claim that a predictable border is possible even when the technical reality of machine learning models does not live up to bureaucratic imaginaries. The so-called success of a predictable border does not solely derive from its technical capacity to estimate human mobility accurately but from creating a semblance of a predictable border inside an organization.

## Keywords

Algorithms, borders, data, migration, predictive analytics

---

## Corresponding author:

Burcu Baykurt, Department of Communication, University of Massachusetts Amherst, Amherst, MA 01003, USA.

Email: [bbaykurt@umass.edu](mailto:bbaykurt@umass.edu)

## Introduction

Over the last two decades, predictive technologies have grown to determine who is allowed at the border from visitors to immigrants to asylum seekers. With their credulous pledge to eliminate irregular entries, algorithmic systems appeal to state and non-state actors who turn to these technologies in the name of national security or efficient management of borders. Critics point to concerns around fairness, rights to seek asylum, and freedom of movement as these systems are quietly incorporated into everyday border management practices. Keeping these concerns in mind, this article examines how claims of predictable borders via algorithmic systems are crafted in bureaucratic institutions. We examine two UN Refugee Agency (UNHCR) projects, namely, Project Winter Cell and Project Jetson that use machine learning and data science techniques as a case study to examine how the UNHCR claims and performs predictability of border movements.

We define a predictable border as one that aims to use data science techniques from diverse sources to identify non-linear, complex relations to help predict human mobility. A predictable border promises to anticipate who might arrive and when by attempting to “translate humanitarian needs into data models” (UNHCR Innovation Service, 2021). The so-called “technological benevolence” (Benjamin, 2019) of the UNHCR’s predictable border, however, potentially perpetuates exclusionary mechanisms of border control by “expanding the scope for racialization” (Amoore, 2021: 2). From concerns around the ethics of consent for data collection to data security to racial and gender bias and discrimination in decision-making, predictive analytics in migration presents a range of significant risks and harms. It is also technically too complex to build and maintain a predictable model of mobility given the uncertainty and diversity of mechanisms resulting in forcibly displacing people worldwide. Against these risks of success and potential harm, we ask: how does a humanitarian agency keep the idea of a predictable border resilient when it does not work as intended or risks harming the people it aims to protect?

Drawing on publicly available documents and interviews with data analysts and managers involved in these two projects, we argue that the so-called success of a predictable border does not solely derive from its technical capacity to predict numbers accurately but also from creating a semblance of a predictable border within the organization. We identify three strategies that help humanitarian technologists justify and maintain the claim that a predictable border *is* possible: (1) constantly seeking novel variables and data, (2) continually maintaining ambiguity, and (3) quickly shifting models to adapt to changing circumstances. These strategies help humanitarian technologists and managers at the UNHCR protect their experiments with predictive analytics from internal and external pushback. They allow for continued efforts to deliver on the promise of a predictable border without having to determine clear metrics of success or failure.

This article contributes to social studies of digital technologies in migration and border control by focusing on how bureaucratic organizations build these predictive systems. We demonstrate how humanitarian agencies reframe their turn to these inherently uncertain, not to mention potentially harmful, technologies not solely through efficiency and better management but also through organizational strategies that legitimate a predictable border while trying not to sacrifice their humanitarian mandate.

## The idea and ideal of a predictable border

From the mass displacement of Syrians, Yemenis, and South Sudanese due to conflicts to the uprooting of millions of Venezuelans because of severe economic and political instability, the global number of refugees and asylum seekers has grown over the last two decades (International Organization for Migration, 2020a). In response, relatively wealthy states and those in stable regions have transitioned their territorial borders into what Petra Molnar (2020) calls “technological testing grounds.” From “fortress” Europe to North America, many wealthy states have turned to opaque and privately developed technologies to curtail a large inflow of border-crossers (Beduschi, 2020; Madianou, 2021; Molnar and Gill, 2018). These “untested technologies” (Madianou, 2021) at and across national borders have been justified as practical responses that help potential host governments deal with complex and protracted challenges in managing migration.

The turn to novel data-driven technologies complements prior border externalization practices by powerful states wherein borders are restructured into complex and exclusionary sites (Shachar, 2020), and the so-called migration management is outsourced to third countries (Bialasiewicz, 2012). Tech-centric interventions to supposedly curtail illegal border crossing and streamline asylum processes are also forming their own “rationality of governing in borderzones” (Aradau, 2020: 1). While humanitarian agencies such as the UNHCR insist on distinguishing between morally responsible uses of tech in contrast to privately-owned ones (Jacobsen, 2015; UN Global Pulse and UNHCR Innovation Service, 2017), a more extensive surveillance infrastructure regime shapes the design and uses of similar tools and techniques (Benjamin, 2019; Couldry, 2017; Sargsyan, 2016). By turning migrants and refugees into data points, this emerging assemblage of data-centric technologies in humanitarian organizations is likely to rematerialize existing practices of exclusion and discrimination in various permanent databases (Chouliaraki and Georgiou, 2022). Cloaked behind the veneer of objectivity and science, these new digital interventions tend to reinforce the growing overlap between humanitarianism and securitization due to the expansion of surveillance (Ticktin, 2014), not to mention to reproduce the power asymmetries derived from colonial legacies (Madianou, 2019a).

A “deep border,” to use Louise Amoore’s (2021) apt term, appears through machine learning models that attempt to seek different sets of features and clusters to grasp “representations from data” and to “generate meaning” from digital information (p. 2). Even if machine learning is not launched at the border, its logic and techniques might be applied to anticipate how many people arrive, when, and what kind of demographics. Amoore (2021) rightly suggests that bringing machine learning techniques into the border ultimately “forecloses and disavows” the politics of mobility. Not only do such predictive or smart borders become “more intensive sites of racialized surveillance and displacement” (Benjamin, 2021), but they also expand the reach of violence on racialized bodies beyond state institutions (Amoore, 2021; Beduschi, 2020; Bellanova et al., 2021; Molnar, 2021; Molnar and Gill, 2018).

Between the violent politics of so-called smart borders at the macro-level (Achieme, 2021; Aizeki et al., 2021; Molnar, 2020) and their socio-economic impact on the everyday life of refugees (Metcalf, 2022; Tazzioli, 2022; Ustek-Spilda and Alastalo, 2020), this article focuses on the work of designing and running predictive analytics projects at

a transnational organization such as the UNHCR. The Agency is mandated to protect, provide for, and find durable solutions for forcibly displaced individuals (UNHCR, 2013). Through its Innovation Service, the UNHCR, similar to other international organizations, explores the possibility of predicting future migration flows to help with humanitarian responses on the ground (International Organization for Migration, 2020b). We examine UNHCR analysts' strategies to enable the ongoing work around modeling border movements and sustain these predictive initiatives against legal and political pushbacks. We seek to understand how analysts at the Innovation Service incorporate predictive projects into existing workflows and how they modify or recalibrate their work to shield it against criticism. Our research aims to demonstrate how the efforts toward crafting a predictable border align with the politics of existing practices inside a humanitarian agency.

## **Case selection and methods**

As a case study, this article focuses on the UNHCR Innovation Service's two predictive projects (Project Winter Cell and Project Jetson). Data for our case studies come from publicly available documents on the UNHCR websites, including academic articles and online publications documenting how these projects came together and what they do. We also review publicly available records on the UN Global Pulse and UN OCHA Center for Humanitarian Data websites since these two offices work closely with the UNHCR Innovation Service and share similar objectives.

We complement the documentary analysis with in-depth interviews with four data analysts at the UNHCR Innovation team and two humanitarian technologists outside the UNHCR with field experience. We identified these technologists from the publicly available documents and initially contacted 10 interviewees via email. The interviews were conducted over Zoom between October 2021 and March 2022 and lasted an average of an hour. We discussed analysts' backgrounds, involvement in predictive analytics projects, justifications underlying the use of predictive analytics in border movements, and how they navigate the risks and problems that emerge in the testing of predictive technologies. Since Project Winter Cell was a precursor to Jetson, which was still in-progress during our interviews, most interview data focused on the latter. But our interlocutors also mentioned that Winter Cell provided a background to Jetson as part of their experimentation.

The data analysis drew on a grounded theory approach wherein we iterated between data collection and analysis (Charmaz, 2006). We focused on identifying how our interlocutors described and defended predictive analytics projects and aimed to trace those themes across the documents, transcripts, and memos we put together. The two authors routinely checked and challenged each other's interpretations, moving back and forth between data collection, memo writing, and identifying themes.

## **Innovation at the UNHCR**

Humanitarian agencies cite the growth in the "size and severity of humanitarian crises" as a rationale for embracing novel technologies and data science techniques (Earney and

Moreno Jimenez, 2019; Pham and Luengo-Oroz, 2022). The UNHCR's "unprepared response" in the Dollo Ado refugee camp (Ethiopia) in 2011, for example, was one of these cases, wherein the Agency's failure to anticipate the number of Somali refugees resulted in health problems and mortality rates at the camp (Hammond, 2014; Richardson et al., 2013). The UNHCR, in turn, sought to embrace a proactive model of monitoring and preparation to better meet the needs of refugees in humanitarian emergencies (Earney and Moreno Jimenez, 2019; UN Global Pulse and UNHCR Innovation Service, 2017). Turning toward innovation (and predictive technologies) was also a part of the UNHCR's "pragmatic" strategy to legitimate its role as a critical humanitarian actor and sustain its funding from the powerful states in the Global North (Adelman, 2001; Barnett, 2001; Chimni, 1998; Forsythe, 2001; Loescher, 1993, cited in Hamlin, 2021: 82).

The UNHCR's Innovation Service (formally known as the Innovation Unit) was established in 2012 as an "inter-divisional unit" aimed at better understanding innovation and its potential role in the UNHCR (Ambos and Tatarinov, 2019). They began with a small team of four, consisting of analysts and strategists tasked to "capture, harness, and reward innovation within the organization" (Ambos and Tatarinov, 2019: 33). During the 2014–2015 period, the Innovation Service reportedly adopted a lab approach—one that primarily relied on strong partnerships with other UNHCR units and outside NGOs—in identifying and testing innovative solutions for the Agency (Ambos and Tatarinov, 2019). We heard in interviews that the Agency's interest in innovation "goes deeper than a technological solution." Outside the UNHCR, humanitarian technologists have a history of tracking human mobility and resources with technological tools and anticipatory methods (Alburez-Gutierrez and Garcia, 2018; Jacobsen, 2015; Madianou, 2019a; UN Global Pulse, 2020). As one of our interlocutors remarked, humanitarian technologists believe they "may be able to reduce the number of lives lost and money [spent], so the funding is a lot better spent" with better preparation (humanitarian technologist outside the UNHCR, 20 September 2021). Technology or analytical techniques might be a means toward this goal.

The Innovation Service's first big, publicly known project, Project Winter Cell, began in 2015 in the middle of the so-called refugee crisis in the Mediterranean, which Luiza Bialasiewicz (2012) calls "the premier laboratory for experimenting creative solutions to the policing of EU borders" (pp. 847-848). To better understand—and prepare for—the number of arrivals and all the routes through the Mediterranean, the Innovation Service turned to what they called "non-traditional datasets" and searched for patterns in people's movements (data analyst at the UNHCR, 28 September 2021). The team first tried to harness vast weather data from the World Meteorological Organization, the Met Office in the United Kingdom, and the Met Service of Southeastern Europe (Karmi, 2016). They drew on tracking vessel movements and broadcast warnings to be able to predict the arrival of refugees and coordinate with country offices. According to the UNHCR's reporting, they also monitored social media and "coordinate[d] with a point person in the intelligence unit inside Sweden's Migration Agency to keep abreast of the latest refugee movements" (Karmi, 2016). The team also experimented with a bot examining xenophobia and racism against refugees online to help with advocacy strategies (Moreno, 2017). The Winter Cell was initially designed to be a short-lived experiment but was extended

into 2016. Once the team decided that the situation on the ground became more manageable, the project was folded in June 2016.

Project Winter Cell became a precursor to the next one, Project Jetson (UNHCR Innovation Service, 2019b). At the request of field officers in Dollo Ado, the Innovation Service sought to develop a tool that would help predict precisely how many refugees would arrive at a point in time. “Can we use basic machine learning models with historical data to anticipate things in an aggregated way? This is basically how Jetson was born,” recalled one of the team members (data analyst at the UNHCR, 28 September 2021). Following similar techniques from Winter Cell, data analysts first turned to weather data, conflict-related data, and historical records of border crossings in Somalia since the country has a long history of displacement. If Project Winter Cell relied on early warning systems by utilizing real-time data (including maritime data, mobile phone data, and meteorological data), Project Jetson primarily utilized the forecasting (modeling) approach by using longitudinal datasets for prediction. Once the modeling was complete, Project Jetson could reportedly estimate arrivals from “13 out of the 19 regions” it targeted “with 3–4 weeks in advance” (Earney and Moreno Jimenez, 2019: 116).

Efficiency, which ostensibly leads to better preparation in humanitarian sites, is an oft-cited motivation for developing predictive analytics in border movements. Humanitarian technologists also suggest that data science techniques are an extension of the data work they are already doing. These novel techniques are, however, developed within the parameters of contemporary data capitalism (Korkmaz, 2022; Sadowski, 2019; West, 2019). Data capitalism, also known as surveillance capitalism, not only aims to ubiquitously surveil and continually capture all data in the name of capital accumulation but also extracts and exploits data, often via invasive, opaque techniques (Madianou, 2019a; Sadowski, 2019; Zuboff, 2019). Instituting predictability within refugee mobility also presents a new set of risks and harms. For example, while fleeing genocidal violence from the Myanmar government, Rohingya refugees experienced intensive data collection at the UNHCR’s Cox Bazaar camp. It soon became clear that the UNHCR was sharing these data with the Bangladeshi government to aid the Agency’s efforts at “finding durable solutions for the refugees” (Human Rights Watch, 2021). It was revealed that between 2018 and 2021, the Bangladeshi government shared Rohingyas’ biometric and other identifiable data with the Myanmar government to facilitate the so-called “repatriation eligibility assessments” (Human Rights Watch, 2021). The UNHCR apologized for sharing the biometric data in this case, but its exploration with data collection and analysis continued (Amoore, 2021).

Against such mistakes and potential risks, the Innovation Service’s dedication to developing predictive projects presents a case of what Ruha Benjamin (2019) calls “technological benevolence.” Using the promise of better care and safety as a justification, humanitarian technologists experiment with machine learning models in forced displacement and claim to distinguish their work from data/surveillance capitalism. But collecting data for humanitarian ends or housing and analyzing data inside a non-governmental organization does not automatically translate into humanitarian accountability. On the contrary, the logics of securitization and capitalism continue to drive many predictive migration projects from those of the UNHCR to Frontex (Amoore, 2021; Madianou, 2019b). It is already hard, if not impossible, to develop an ethical and socially just machine learning model that anticipates forced displacement. From a statistical

perspective, trying to predict human mobility is also a kind of make-believe. Several scholars have argued that while the size of migration flows can be forecasted, human mobility as a phenomenon is “fundamentally stochastic” (Kulkarni et al., 2019; Smith et al., 2014; Song et al., 2010). Migration flows such as those of refugees have random patterns that may be analyzed statistically but not predicted with precision.

## **Building a predictable border**

The Innovation Service is not at all oblivious to such challenges. Our interlocutors often brought up these issues unprompted to showcase their awareness of and care about applying machine learning to humanitarian work. However, they also remained committed to developing these systems and remained firm in their beliefs that predictive technologies might be helpful in humanitarian settings. This section discusses three strategies data analysts use to render predictive analytics acceptable within the UNHCR. First, they frame these projects as an opportunity to discover new intelligence about refugee movements. Second, they carefully maintain an ambiguity around the content, use, and goals of the projects to protect them from hostile governments and internal criticism from the UNHCR. Finally, they continually shift the models under the celebratory banner of experimentation and try to adapt to changing field dynamics and external events. These strategies aim to decouple the idea of a predictable border from the existing predictive technologies and techniques developed by the Innovation Service. Rather than focusing on the success, failure, risks, or harms of existing projects, these justifications help data analysts sustain the possibility of a predictable border while shielding their experimental projects against immediate pushback.

## ***Seeking novel variables***

Predicting human movements is a highly complex task and depends on a wide range of datasets to at least make sense statistically. Data analysts thus seek new explanatory variables and track novel trends that could explain human mobility and approximate accuracy levels in their models. Project Jetson started with “enough historical data on arrivals,” one of our interlocutors explained, but then the team pulled in:

different data sources such as rainfall, commodity prices, or conflict data. The idea is basically to build a streamlined data set and then a bunch of different predictive models to test out whether we could forecast arrivals and how far in the future we could do this. (A data analyst affiliated with the UNHCR, 14 September 2021)

Members of the Innovation Service do not see this data work as a challenge. On the contrary, they frame using different sources of information as a way of capacity-building. Coordinating across various sources and agencies in and outside the UNHCR cultivates this capacity both epistemologically and organizationally. They seek access to publicly available data that may be directly or tangentially related to migration trends. In turn, at least in theory, the UNHCR could reportedly obtain better insights into the field dynamics even if machine learning models are unsuccessful.

Field officers and analysts in the humanitarian field are used to collecting and analyzing various data sources across different units to better prepare for refugee arrivals. What distinguishes a project like Jetson, some of our interlocutors asserted, is the ability to combine diverse datasets while exploring different relationships across several variables. One analyst remarked that “machine learning forced [them] to do [data collection] in a more structured way” and helped the organization track in a more systematic fashion (a data analyst affiliated with the UNHCR, 14 September 2021). Others mentioned that predictive analytics enabled tracking trends they might not have otherwise followed. Examples of such “novel” trends were goat prices for the Dollo Ado camp or bus schedules and tickets in the case of Brazil.

Priding itself on pursuing a “human-centered” approach, the Innovation Service tried to engage refugees and field officers in data discovery and collection during Jetson. A data analyst recounted flying to Dollo Ado to interview refugees with the help of a translator in the early days of Project Jetson. During one of those visits, the analyst found out about what came to be described as “the goat story” (UNHCR Innovation Service, 2019a). A refugee explained that following security threats, the family had to sell all their livestock (mainly goats) before they could flee to Dollo Ado because of the inability of goats to survive the perilous journey. To the data analyst, this detail implied that during a significant outflow of Somalis to Dollo Ado, market prices would also plummet, given that many families would be selling their goats simultaneously. Selling goats before departure thus became more than just one refugee’s story. It was adopted as a predictor of Somalis’ movements to the southern Ethiopian border and incorporated into Jetson’s model to estimate the size of future refugee flows in the region. The data analysts recognized that such novel variables were sensitive to local contexts and periods. They mentioned how bus fares, for example, replaced goat sales in another refugee site to acknowledge how specific these new insights were to locales.

Regardless of whether a predictive analytics project achieves the intended accuracy levels, members of the Innovation Service consider this data discovery process very useful. They believe these new insights and analytical techniques ostensibly strengthen the knowability of migration and displacement for the UNHCR. They also feel good about using such arguably low-stakes, publicly available data because they could avoid political questions of privacy or surveillance while praising new forms of information (Meissner and Taylor, 2021; Molnar, 2020). We find, however, little to no accountability regarding the responsibility of testing these forms of data, such as weather trends or goat prices, to predict human mobility. Recent evidence seems to counter the assumption that market prices can be solely associated with a large outflow of Somalis to neighboring countries (Khalif, 2020; Omar and Bearak, 2020). Data analysts do not question the impact of using unrelated and context-dependent data to inform high-stakes decision-making, such as preparing for the arrival of refugees at the border or profiling refugee populations. The veneer of complexity provided by machine learning techniques justifies blending these different data sources and their claims to be able to predict refugee mobility. Data analysts also do not consider that these information sources, legitimated by the UNHCR, might create market opportunities for data analytics start-ups (Taylor and Meissner, 2020), thereby creating a dependence on commercial firms (Lemberg-Pedersen and Haioty, 2020) in the future.



## *Maintaining ambiguity*

The Innovation Service likes to promote Project Jetson through multiple channels. Team members author academic publications, give presentations worldwide, and publicize their experiences on blogs and websites. Nevertheless, there is very little publicly available information about the details of this project. The stakes of revealing Jetson's content and results are high. Only select UNHCR officials have access to the model and datasets. Of course, data analysts coordinated with the officers at Dollo Ado to share their estimates and receive actual numbers to check the model's accuracy levels. But beyond that, they mentioned being very careful about the limits of circulating Jetson in and outside of the UNHCR. One interlocutor reported,

The Ethiopian team told [us] not to publish anything about Jetson. If the government of Ethiopia knew those are accurate predictions and there are so many people arriving, you are going to shoot yourselves in the foot; you're going to shut down the UNHCR in Ethiopia. (Data analyst at the Innovation Service, 28 September 2021)

Keeping the content and narrative of Project Jetson ambiguous is a crucial strategy for the Innovation Service. That is not total opacity. But there is careful selection and curation of what is to be shared and what is kept opaque. The team collaborates with academics worldwide and expresses an eagerness to share their knowledge with researchers if there is interest. (The list of these academic partnerships is, however, not publicly available.) When we asked our interlocutors what kind of risks they anticipate in a project like Jetson, for example, they immediately responded that the content of Jetson is only shared with a select few. They made sure to create an impression of a tightly kept environment wherein a project such as Jetson is tested and refined, thereby claiming to control potential risks and harm.

This ambiguity partly owes to the precarity of the UNHCR's relationship with host countries, as we see in the above quote. Predictive technologies contain multiple levels of opacity due to technical and social reasons (Burrell, 2016; Pasquale, 2015). Privacy concerns for one of the most vulnerable populations globally also call for extra care and safeguarding. The Innovation Service is aware of its duty to protect such projects from the encroachment of host governments. Getting the data of refugees (as we saw in the case of Cox Bazaar) or the predictions of Jetson in the hands of host governments presents a significant risk to the lives of the forcibly displaced. Recognizing that risk might also endanger continuing projects such as Jetson at the UNHCR.

Maintaining ambiguity about such projects is also derived from the internal politics of the UNHCR and the Innovation Service's attempt to keep the legitimacy of their work unchallenged within the organization. The Innovation Service is not a fundamental unit within the UNHCR. It is a relatively new division trying to prove its worth to the rest of the Agency. Team members are thus careful to present their projects as something which supplements existing workflows at field offices rather than supplant what they are already doing. The field office at Dollo Ado, for example, could not take any explicit action on Jetson's predictions because the UNHCR headquarters had no established policy that allows using "mathematical formulas" for preparation (an engineer at the Innovation Service, 18 October 2021). An analyst explained how they navigate the

ambivalent role between the field operations and what they call the policy arm of UNHCR as follows:

We don't want to be telling them from Geneva or New York like this is going to happen, and you need to reallocate. We want to say, here's what our models are showing, here's something you might want to investigate, you might want to think a little bit about. (Analyst at the Innovation Service, 28 September 2021)

Another reason the Innovation Service was cautious about circulating or suggesting actions on Jetson's results was the lack of guidelines regarding data ownership and algorithmic ethics across the UN.

Not being able to act on ("operationalize") Jetson's results is a contentious topic among the members of the Innovation Service. Some think Jetson was a failure because the field offices did not change any of their preparation based on predictions. Others suggest it was the higher levels of bureaucracy, such as the weak legitimacy of the Innovation Service or the lack of UN-wide guidelines around algorithmic accountability, which caused the missing link between operational responses and Jetson's numbers. In other words, Jetson was not a failure; the UNHCR was not yet ready to act on it. Either way, the team carefully maintained ambiguity on many levels, for example, among team members, between the Innovation Service and field offices, and between the UNHCR and members of the public. It aimed to protect projects such as Jetson from being questioned or attacked internally or externally.

### *Shifting models*

In our interviews with the Innovation Service members, we were often reminded that Project Jetson was a "proof of concept" and not a tool. Our interlocutors insisted this was not a model that was actively in use. Nor was it intended to provide a perfect model that works across all places. Jetson, they told us, strove to develop a working machine learning model for predicting refugee movements. However, the Innovation Service members were aware that their results were specific to the sites and periods wherein they conducted their work. They knew they could not generalize beyond the confines of their testbed. Moreover, they did not want to because, as the "goat story" demonstrated, it was more crucial to find "sensitive variables that can be indicators of movement, even if in an over-simplified way" (Earney and Moreno Jimenez, 2019). As discussed in the previous section, there were also regulatory and organizational obstacles to rolling out a predictive model across the UNHCR. The goal for the Innovation Service was then to show they *can* model and predict refugee movements while also identifying a diverse range of data sources (i.e. predictors) that the Agency could keep track of for the next wave of arrivals.

Migration scholars have already established that there is "no systematic way of estimating the future size of PoC groups because the causes of forced displacement are difficult to predict" (Alburez-Gutierrez and García, 2018). The Innovation Service members seemed to accept that a predictable border was a shifting target. "There are constantly emergent trends that do not always have the data," one team member explained, "so

you're trying to balance the constantly changing situation with the ability to build on what you have learned and what you do know" (a data analyst affiliated with the Innovation Service, 14 September 2021). Another member admitted that "if you were to run the model right now in Ukraine, the predictive displacement would be zero" (a member of the Innovation Service, 2 March 2022). Project Jetson thus aimed to provide estimates within minimal periods, such as 3 or 4 weeks in advance. Our interlocutors accepted that its predictive model must be continually reworked and retrained across different sites and at other times. The project website declares, "Jetson actively seeks new data sources, new narratives, and new collaborations to keep iterating and improving" (UNHCR Innovation Service, 2021).

Experimentality has become "a mode of governing in borderzones" to test new technologies or anticipation techniques but often in debilitating ways for migrants (Amoore, 2021; Aradau, 2020). The Innovation Service's central role across the UNHCR is to create spaces of experimentation within the organization while also promoting data science and applied artificial intelligence techniques in the humanitarian sector. Experimentality certainly affords the Innovation Service to pursue a more flexible approach as they tinker with different datasets or statistical models. It allows them to make mistakes in a more contained way without de-legitimizing the humanitarian work of the UNHCR. Even though the UNHCR promotes durable solutions for refugees as part of its core mandate, there is no clear endpoint in experimentation. The tinkering and iterating, especially in human mobility, could continue across different sites.

The experimentality behind shifting models also obscures the feedback loop between data analysts and field officers, wherein the latter's observations are not only fed into the models, but also their review ends up settling the accuracy levels. The Innovation Service aims to keep Jetson's predictions in a range of options rather than seeking precision. "[W]e didn't specifically give [field officers in Dollo Ado] a number," one of our interlocutors explained, "But we wanted to get the intervals between predictions so that they know that there will always be a maximum value that they had to have in mind" (an engineer at the Innovation Service, 18 October 2021). This cautious approach allows the team not to "extend beyond the models," thereby containing the impact of predictive analytics within the locales they were developed. But it also means a project like Jetson relies heavily on the assessment of field officers to prove its reliability. For example, suppose the field office finds some predictions unrealistic compared with what they had expected or observed. The Innovation Service then switches to other types of insights, such as trying to estimate the reasons behind fleeing. "With the operation, they have to tell you what is valuable out of the modeling," one analyst explained (an engineer at the Innovation Service, 18 October 2021). Otherwise, analysts continue working around projectors to produce something that the field office deems useful.

The ongoing experimentality and shifting models in line with the feedback of field officers and response to site-specific changes demonstrate how Jetson—or any predictive project in the field of migration—"selectively" confers different variables or predictors' reliability (Grommé, 2015). Rather than perceiving the borders and human mobility as the absolute sites of truth from which the predictive models could capture and learn, there is a complex network of UNHCR staff, competing interests, organizational limits,

and changing circumstances, which shape the design, results, and validity of predictability (Bourne et al., 2015). However, precisely this kind of selectivity and openness also sustains the idea of a predictable border. Continually shifting machine learning allows the Innovation Service to remain at the helm while proliferating its experiments across different sites.

## Discussion

Despite the complexities underlying predictive analytics and machine learning in predicting border crossing, UNHCR officials are still determined to expand projects such as Jetson in other regions (Aylett-Bullock et al., 2021). Those in favor of migration predictability maintain that despite their limitations, predictive technologies enable researchers and humanitarian technologists to “assemble the necessary data” useful to explain not only how the next wave of displacement “might unfold” but also “where the displaced would flee and when” (Edwards, 2008: 347). These proponents suggest that the ultimate objective of using predictive models in migration and border management is to capture the unfolding of border movements in advance, which helps actors be well prepared (Edwards, 2008; Pham and Luengo-Oroz, 2022). However, these technologies also carry a wide range of risks for the already-vulnerable populations and present scant evidence that they are effective.

To better understand the dynamics of a predictable border within bureaucratic organizations like the UNHCR, our analysis highlights how predictive systems in migration intersect with the larger politics of humanitarian organizations. We aim to counter the assumption that states are the “primary actor in assembling and analyzing migration data” (Meissner and Taylor, 2021) and spotlight the intertwined role both states and non-state actors play in reimagining migration and borders. Humanitarian organizations, unlike states or companies, have a distinct responsibility toward the forcibly displaced and members of the public. The UNHCR Innovation Service—and the rest of the organization—is quite clear about the risks and harms of instituting predictive analytics in their everyday work. The organizational strategies we identify help them separate the idea and ideal of a predictable border from the actual experiments they are running in the office or the field.

Through a case study, we identified three main strategies that UNHCR analysts pursue to sustain the possibility of a predictable border. First, by seeking novel variables, the Innovation Service claims to generate an organizational value for the UNHCR beyond the success of predictive analytics. Even when projects such as Winter Cell or Jetson are folded, these new datasets or correlative insights revealed through blending them remain relevant practices within the organization. Second, maintaining a carefully curated ambiguity around the content and consequences of predictive projects allows the Innovation Service to avoid internal and external pushback. It also creates an image of safety around the impact of predictive projects wherein the UNHCR—and its strategies of ambiguity—supposedly protect vulnerable populations from the consequences of predictions. Finally, the Innovation Service continually shifts predictive models in response to changing local conditions and the feedback from field offices. This constant experimentality

sustains the Innovation Service's leadership in predictive projects and their promise that a predictable border is possible to build.

These strategies show that the design and testing of predictive analytics within a humanitarian agency, such as the UNHCR, need to adapt to the politics of the organization (Reutter, 2022). Beyond the effectiveness or success/failure of a specific technology at hand, these relational strategies sustain the socio-technical imaginary of a predictable border. These organizational practices do not fully explain why humanitarian organizations insist on using algorithmic systems despite the overwhelming evidence that such data-driven technologies do not work as intended and worsen the conditions for the already-marginalized (Sánchez-Monedero and Dencik, 2022). But they demonstrate how these data-driven techniques are legitimated inside a bureaucratic organization. Even if a particular predictive project does not achieve its goals (as the so-called success of Jetson is contested even by the members of the Innovation Service), the idea of being able to predict the mobility of refugees remains intact. Drawing on *other* benefits of adopting predictive technologies (e.g. seeking new variables or experimentation), humanitarian technologists keep working on such projects, and the fitful implementation of predictive technologies continues apace.

## Conclusion

In this article, we have examined predictive analytics projects of the UNHCR Innovation Service to understand how a humanitarian agency attempts to adopt machine learning models in decision-making. We have argued that the resilience of the ideal of predictive analytics at the UNHCR depends on three organizational strategies—seeking novel variables, maintaining ambiguity, and shifting models—rather than the quantitative strength of analytics or predictions. We showed that these strategies help humanitarian technologists justify their efforts toward designing and testing a predictable border by decoupling the success of a specific project from other organizational practices that arguably help the UNHCR understand human mobility better. They also aim to create a sense of responsibility and security around the contents of these projects through a carefully maintained bureaucratic ambiguity.

Our inquiry would have certainly benefited from a participant-observation component to systematically monitor how these systems are built and tested on the ground. Since we could not secure access to the internal office of the Innovation Service or the field sites, we prioritized the reported accounts of the UNHCR staff in our analysis. What other interests—inside and outside the UNHCR—maintain the resilience of trusting data science techniques in migration? What would it take for the UNHCR to give up on developing predictive analytics? Future research should explore these questions by studying the perspectives of field officers and other units within the UNHCR to understand better the durability of these organizational strategies against the valid critiques of algorithmic systems in the field of migration. The stakes of challenging the premise of such projects—that border movements can be predicted—are high as ongoing experimentation in refugee sites and at humanitarian agencies continues to amplify existing power asymmetries across the Global North and South, exclusionary mechanisms of borders, and biases in decision-making.

## Acknowledgements

Earlier versions of this article were presented at the *Critical Borders: Radical (Re)visions of AI* conference at the University of Cambridge, at the *STS-MIGTEC* workshop, Migration Working Group at UMass Amherst, and *Digital Technologies and Migration Regimes* workshop at the University of Oxford. The authors would like to thank the organizers and participants of these meetings, the Journal's anonymous reviewers, and Dr. Scott Timcke for their helpful comments.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Burcu Baykurt  <https://orcid.org/0000-0002-1690-4237>

Alphoncina Lyamuya  <https://orcid.org/0000-0003-3945-4328>

## References

- Achiume E (2021) Digital racial borders. *AJIL Unbound* 115: 333–338.
- Aizeki M, Boyce G, Miller T, et al. (2021) *Smart Borders or Humane World?* Immigrant Defense Project's Surveillance, Tech & Immigration Policing Project, and the Transnational Institute. Available at: [https://www.tni.org/files/publication-downloads/smart\\_borders\\_humane\\_world\\_2021.pdf](https://www.tni.org/files/publication-downloads/smart_borders_humane_world_2021.pdf)
- Alburez-Gutierrez D and García C (2018) *The UNHCR Demographic Projection Tool: Estimating the Future Size and Composition of Forcibly Displaced Populations*. Available at: <https://www.unhcr.org/5ae9ee747.pdf>
- Ambos TC and Tatarinov K (2019) Seed, scale, structure: how international organizations shape innovation. Available at: [https://www.unige.ch/gsem/files/2816/2193/7363/Report\\_Seed\\_Scale\\_Structure.pdf](https://www.unige.ch/gsem/files/2816/2193/7363/Report_Seed_Scale_Structure.pdf)
- Amoore L (2021) The deep border. *Political Geography* 25: 102547.
- Aradau C (2020) Experimentality, surplus data and the politics of debilitation in borderzones. *Geopolitics* 27: 26–46.
- Aylett-Bullock J, Cuesta-Lazaro C, QueraBofarull A, et al. (2021) *Operational Response Simulation Tool for Epidemics within Refugee and IDP Settlements*. UN Global Pulse. Available at: <https://www.unglobalpulse.org/document/operational-response-simulation-tool-for-epidemics-within-refugee-and-idp-settlements/>
- Beduschi A (2020) International migration management in the age of artificial intelligence. *Migration Studies* 9: 576–596.
- Bellanova R, Irion K, Lindskov Jacobsen K, et al. (2021) Toward a critique of algorithmic violence. *International Political Sociology* 15(1): 121–150.
- Benjamin R (2019) *Race after Technology: Abolitionist Tools for the New Jim Code*. Medford, MA: Polity.
- Benjamin R (2021) Professor Ruha Benjamin—Obert C. Tanner Lecture on AI and Human Values 2021. Available at: <https://www.youtube.com/watch?v=e3MQum7HrVM>
- Bialasiewicz L (2012) Off-shoring and out-sourcing the borders of Europe: Libya and EU border work in the Mediterranean. *Geopolitics* 17(4): 843–866.
- Bourne M, Johnson H and Lisle D (2015) Laboratizing the border: the production, translation and anticipation of security technologies. *Security Dialogue* 46(4): 307–325.

- Burrell J (2016) How the machine ‘thinks’: understanding opacity in machine learning algorithms. *Big Data & Society*. Epub ahead of print 6 January. DOI: 10.1177/2053951715622512.
- Charmaz K (2006) *Constructing Grounded Theory*. London: SAGE.
- Chouliarakis L and Georgiou M (2022) *The Digital Border: Migration, Technology, Power*. New York: NYU Press.
- Couldry N (2017) Surveillance-democracy. *Journal of Information Technology & Politics* 14(2): 182–188.
- Earney C and Moreno Jimenez R (2019) Pioneering predictive analytics for decision-making in forced displacement contexts. In: Salah A, Pentland A, Lepri B, et al. (eds) *Guide to Mobile Data Analytics in Refugee Scenarios*. Cham: Springer, pp. 101–119.
- Edwards S (2008) Computational tools in predicting and assessing forced migration. *Journal of Refugee Studies* 21(3): 347–359.
- Grommé F (2015) Turning aggression into an object of intervention: tinkering in a crime control pilot study. *Science as Culture* 24(2): 227–247.
- Hamlin R (2021) The United Nations high commissioner for refugees. In: Hamlin R (ed.) *Crossings: How We Label and React to People on the Move*. 1st ed. Stanford, CA: Stanford University Press, pp. 70–92.
- Hammond L (2014) *Somali Refugee Displacements in the Near Region: Analysis and Recommendations*. UNHCR. Available at: <https://www.unhcr.org/55152c699.pdf>
- Human Rights Watch (2021) *UN Shared Rohingya Data without Informed Consent*. Human Rights Watch. Available at: <https://www.hrw.org/news/2021/06/15/un-shared-rohingya-data-without-informed-consent>
- International Organization for Migration (2020a) *IOM Global Report 2019: Operations and Emergencies*. Available at: <https://www.iom.int/sites/g/files/tmzbd1486/files/documents/IOM-Global-Report-2019-Operations-and-Emergencies.pdf>
- International Organization for Migration (2020b) Future migration trends. *Migration Data Portal*. Available at: <https://www.migrationdataportal.org/themes/future-migration-trends>
- Jacobsen KL (2015) *The Politics of Humanitarian Technology: Good Intentions, Unintended Consequences, and Insecurity*. London: Routledge.
- Karmi O (2016) *As a Big Chill Threatens Refugees, Unhcr’s “Winter Cell” Responds*. UNHCR. Available at: <https://www.unhcr.org/news/latest/2016/1/56a20b3c6/big-chill-threatens-refugees-unhcrs-winter-cell-responds.html>
- Khalif A (2020) Why a quiet Hajj is hurting Somalia? *Foreign Policy*. <https://foreignpolicy.com/2020/07/29/hajj-restrictions-somalia-livestock-goats-humanitarian-disaster/>
- Korkmaz DEE (2022) Understanding surveillance capitalism from the viewpoint of migration. *International Migration* 60(2): 256–260.
- Kulkarni V, Mahalunkar A, Garbinato B, et al. (2019) Examining the limits of predictability of human mobility. *Entropy* 21(4): 432.
- Lemberg-Pedersen M and Haioty E (2020) Re-assembling the surveillable refugee body in the era of data-craving. *Citizenship Studies* 24(5): 607–624.
- Madianou M (2019a) Technocolonialism: digital innovation and data practices in the humanitarian response to refugee crises. *Social Media + Society*. Epub ahead of print 26 July. DOI: 10.1177/2056305119863146.
- Madianou M (2019b) The biometric assemblage: surveillance, experimentation, profit, and the measuring of refugee bodies. *Television & New Media* 20(6): 581–599.
- Madianou M (2021) Nonhuman humanitarianism: when “AI for good” can be harmful, information. *Communication & Society* 24(6): 850–868.
- Meissner F and Taylor L (2021) *Migration information infrastructures: power, control and responsibility at a new frontier of migration research*. Available at: <https://doi.org/10.31235/osf.io/gkzmq>

- Metcalf P (2022) Autonomy of migration and the radical imagination: exploring alternative imaginaries within a biometric border. *Geopolitics* 27(1): 47–69.
- Molnar P (2020) *Technological Testing Grounds: Migration Management Experiments and Reflection from the Ground Up*. *European Digital Rights*. Available at: <https://edri.org/wp-content/uploads/2020/11/Technological-Testing-Grounds.pdf>
- Molnar P (2021). *Artificial borders? The digital and extraterritorial protection of “Fortress Europe.”* Center for Human Rights and Global Justice, November 12. Available at: <https://www.youtube.com/watch?v=4YeyQyvILsM&t=2s>
- Molnar P and Gill L (2018) *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada’s Immigration and Refugee System*. Toronto, ON, Canada: International Human Rights Program and the Citizen Lab, University of Toronto.
- Moreno R (2017) *Teaching a “Robot” to Detect Xenophobia Online*. UNHCR Innovation Service. Available at: <https://www.unhcr.org/innovation/teaching-robot-detect-xenophobia-online/>
- Omar F and Bearak M (2020) With drastically smaller hajj, Somalia’s livestock industry goes from “boom to doom.” *Washington Post*. [https://www.washingtonpost.com/world/africa/hajj-somalia-livestock-exports/2020/07/28/10c984e6-d03a-11ea-826b-cc394d824e35\\_story.html](https://www.washingtonpost.com/world/africa/hajj-somalia-livestock-exports/2020/07/28/10c984e6-d03a-11ea-826b-cc394d824e35_story.html)
- Padoan L (2011) *Camps in East Ethiopia Struggle to Cope with Influx of Somali Refugees*. UNHCR. Available at: <https://www.unhcr.org/news/latest/2011/8/4e3abe716/camps-east-ethiopia-struggle-cope-influx-somali-refugees.html>
- Pasquale F (2015) *The Black Box Society*. Cambridge: Cambridge University Press.
- Pham KH and Luengo-Oroz M (2022) Predictive modeling of movements of refugees and internally displaced people: towards a computational framework. *Journal of Ethnic and Migration Studies* 49: 408–444.
- Reutter L (2022) Constraining context: situating datafication in public administration. *New Media & Society* 24(4): 903–921.
- Richardson L, Partner N, Bush A, et al. (2013) An independent review of UNHCR’s response to the Somali Refugee Influx in Dollo Ado. *Ethiopia* 201: 87.
- Sadowski J (2019) When data is capital: datafication, accumulation, and extraction. *Big Data & Society* 2019: 1–12.
- Sánchez-Monedero J and Dencik L (2022) The politics of deceptive borders: “biomarkers of deceit” and the Case of iBorderCtrl. *Information, Communication & Society* 25(3): 413–430.
- Sargsyan T (2016) Data localization and the role of infrastructure for surveillance, privacy, and security. *International Journal of Communication* 10: 17.
- Shachar A (2020) *The Shifting Border: Legal Cartographies of Migration and Mobility* (Ayelet Shachar in Dialogue). Manchester: Manchester University Press.
- Smith G, Wieser R, Goulding J, et al. (2014) A refined limit on the predictability of human mobility. *IEEE International Conference on Pervasive Computing and Communications (PerCom)* 2: 88–94.
- Song C, Qu Z, Blumm N, et al. (2010) Limits of predictability in human mobility. *Science* 327(5968): 1018–1021.
- Taylor L and Meissner F (2010) A crisis of opportunity: market-making, big data, and the consolidation of migration as risk. *Antipode* 52(1): 270–290.
- Tazzioli M (2022) Governing refugees through disorientation: fragmented knowledges and forced technological mediations. *Review of International Studies* 48: 425–440.
- Ticktin M (2014) Transnational humanitarianism. *Annual Review of Anthropology* 43(1): 273–289.
- UN Global Pulse & UNHCR Innovation Service. (2017). *Social Media and Forced Displacement: Big Data Analytics & Machine-Learning*. Available at: <https://www.unhcr.org/innovation/wp-content/uploads/2017/09/White-Paper-Social-Media-4.pdf>



- UN Global Pulse (2020) *Understanding Population Movement from Venezuela to Brazil Related to COVID-19 Border Closures*. <https://www.unglobalpulse.org/project/understanding-population-movement-from-venezuela-to-brazil-related-to-covid-19-border-closures/>
- UNHCR (2013) *The mandate of the High Commissioner for Refugees and his Office*. The UN Refugee Agency. Available at: <https://www.unhcr.org/en-us/protection/basic/526a22cb6/mandate-high-commissioner-refugees-office.html>
- UNHCR Innovation Service (2019a) *A Goat Story*. UNHCR Innovation Service. Available at: <https://medium.com/unhcr-innovation-service/a-goat-story-3ed6bdd2b237>
- UNHCR Innovation Service (2019b) *Is It Possible to Predict Forced Displacement?* UNHCR Innovation Service. Available at: <https://medium.com/unhcr-innovation-service/is-it-possible-to-predict-forced-displacement-58960afe0ba1>
- UNHCR Innovation Service (2021) *Predicting the Unpredictable: Preparing for Potential Future Scenarios*. Available at: <https://medium.com/unhcr-innovation-service/predicting-the-unpredictable-preparing-for-potential-future-scenarios-1b22cd7f8da2>
- Ustek-Spilda F and Alastalo M (2020) Software-sorted exclusion of asylum seekers in Norway and Finland. *Global Perspectives* 1: 12978.
- West SM (2019) Data capitalism: redefining the logics of surveillance and privacy. *Business & Society* 58(1): 20–41.
- Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

### Author biographies

Burcu Baykurt is an assistant professor of media studies at the University of Massachusetts, Amherst and a faculty associate at the Berkman Klein Center for Internet & Society at Harvard University.

Alphoncina Lyamuya is a PhD student in communication at the Annenberg School for Communication and Journalism, University of Southern California. Her research interests lie at the intersection of emerging digital technologies, transnational governance and inequality. She holds a Master of Public Policy and Administration degree with a graduate certificate in data analytics and computational social science, and a BA in legal studies from the University of Massachusetts, Amherst.