The University of Chicago

Place Matters: The Possibilities and Pitfalls of Machine Learning Algorithms for Refugee Resettlement

By Stephanie Achugamonu

A project essay submitted for partial fulfillment of the requirements
for a Bachelor of Arts degree in Public Policy

Paper presented to:

Instructional Professor, Chad Broughton

Instructional Assistant, Saliem Shehadeh

Undergraduate Program in Public Policy

February 28, 2025

Abstract

The global refugee population is dramatically increasing, posing a challenge for over-saddled refugee agencies who ultimately determine where these individuals should be placed. A machine learning algorithm called GeoMatch positions itself as a solution to this crisis by using historical data to identify communities where refugee families are most likely to succeed. In this paper, I aim to examine how the algorithm functions and has been implemented by resettlement agency personnel. Moreover, I pit GeoMatch's purported benefits against its risks, ethical concerns, and unaddressed limitations. The methodology includes analyzing quantitative research findings, supplementary documentation on the algorithm, and a previously published interview in Stanford's Momentum magazine with a senior resettlement agency staff member. I find that although GeoMatch has branded itself as a cutting-edge technology that simplifies the resettlement process, substantial concerns over data quality, modeling techniques, limited testing, and a hyper-fixation on maximizing shorter-term employment must raise alarm bells. The developers of GeoMatch need to address these shortcomings in order for GeoMatch's potential gains to translate to measurable, real-world improvements to refugee integration outcomes.

Keywords: GeoMatch, algorithms, refugee resettlement, integration outcomes

Place Matters: The Possibilities and Pitfalls of Machine Learning Algorithms for Refugee Resettlement

Consider two refugees, Hasan and Omar. They grew up together in the same town in Afghanistan, attended the same schools, possessed similar skills, and pursued the same profession. They both fled the same conflict and came to the United States with the hope of creating a better life. However, during this transition, the journey took them in different directions. Hasan and Omar were resettled in different cities and had disparate outcomes. Hasan moved to a city with a large Muslim community, quickly capitalized on his skills, landed a job, and learned English through his friendships with coworkers. On the other hand, Omar relocated to a city that did not offer any job prospects or the chance of developing a social network. He struggled to find his feet here. The case of Hasan and Omar, one that symbolizes how the trajectories of incoming refugees can vary widely, is not unique. There is a growing voice out of the refugee resettlement community that asserts that place matters; where a refugee is resettled within a host country can dictate their ability to successfully integrate into their new communities and thrive.

The United Nations High Commissioner for Refugees (UNHCR) reported that at the end of June 2024, 122.6 million individuals around the world were forcibly displaced "as a result of persecution, conflict, violence, human rights violations or events seriously disturbing public order" (United Nations High Commissioner for Refugees 2024, 2). Upon isolating those who were not internally displaced, UNHCR specifies that there were 43.7 million refugees and 8 million asylum seekers (United Nations High Commissioner for Refugees 2024, 2). The numbers confirm the existence of a large cohort of people in desperate need of international resettlement. What is even more concerning is that this cohort is only increasing in size.

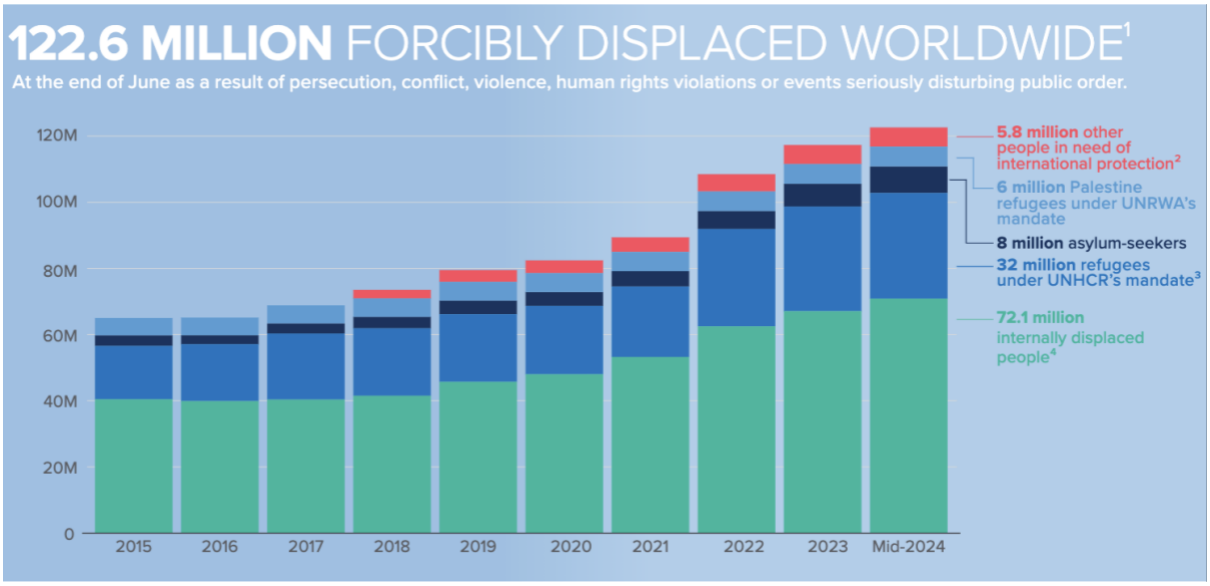

Figure 1: Bar chart showing the growing number of forcibly displaced persons worldwide including those categorized as refugees (United Nations High Commissioner for Refugees 2024, 2).

Taking into account the surging figures of forced displacements around the world, where exactly should the millions of refugees and asylum seekers longing for safe haven be placed? This question has overwhelmed governments, resettlement agencies, and their team of placement officers who are responsible for deciding which community new arrivals are assigned to. Placement officers must consider a wide array of case factors, making their job onerous. For example, staff at Global Refuge, a prominent resettlement agency, analyze refugee information such as country of origin, spoken language, ties to family and friends already living in the United States, and known medical conditions (Global Refuge, n.d). The work doesn't end there. Next, staff must examine attributes specific to the locations in their network of affiliate sites including job opportunities, services, and accessible infrastructure (Immigration Policy Lab 2023b, 1). Finally, placement officers must manually assign each refugee to the location that meets their unique needs without overcrowding any individual community (Immigration Policy Lab 2023b, 1). In fiscal year 2024 alone, 100,034 refugees resettled in the United States which is the most

we've seen since the late 90s (Chishti, Bush-Joseph, and Green 2024). Recognizing a growing resettlement population, it would be impractical and inefficient for staff to perform every step of this complex analysis for each case.

The GeoMatch algorithm was developed to address these challenges and help match refugees to placement locations. The creators of GeoMatch argue that the algorithm can efficiently identify the best resettlement location for refugees and asylum seekers while prioritizing designated integration outcomes. How exactly does the algorithm work? Has it proven itself useful when deployed in real world situations? How should we interpret its reported successes? What are the risks and ethical considerations associated with GeoMatch? These are questions I seek to explore in this paper. I find that although GeoMatch has helped streamline the resettlement process for placement officers, caution must be raised about GeoMatch's overstated potential impact, its limited testing, and various unaddressed concerns that emerge as early as its modeling stage.

Origins of GeoMatch

The machine learning tool GeoMatch was the product of a collaboration between the Immigration Policy Lab's (IPL) branches at Stanford University and ETH Zurich, a leading public university in Switzerland. Using "supervised machine learning and optimal matching techniques," IPL designed an algorithm to help government officials and placement officers pair refugees with the optimal area for them to live within a host country (Ozkul 2023, 60). IPL also boasts that GeoMatch is a "first-of-its kind algorithm" because a recommended location represents the area where a refugee is "most likely to succeed" with respect to an outcome of interest (Immigration Policy Lab 2023a). The data fed into the algorithm includes gender,

language spoken, age, highest level of educational attainment at arrival, and country of origin. GeoMatch tries to unearth interactions between personal characteristics and resettlement locations. IPL gives a brief overview on how the algorithm works:

1. Historical data about past refugees and their integration outcomes is used as an input into the algorithm
2. The algorithm develops models to predict integration outcomes based on trends and patterns in the historical data
3. The algorithm receives new data on the personal characteristics of incoming refugees
4. The algorithm predicts how well these newcomers will do in various locations
5. These predictions are used to generate location recommendations for placement officers (Immigration Policy Lab 2023c)

What do the developers mean by when they say the algorithm is trained on historical data? This still remains a bit unclear. Here is what we know. IPL's researchers used "de-identified data from one of the largest resettlement agencies for working-age refugees" (Bansak et al. 2018a, 327). More specifically, this data was on 33,782 refugees between the ages of 18 and 64 who were assigned to U.S. cities from 2011 to 2016 (Bansak et al. 2018a, 327). The algorithm was built using data restricted to refugees who entered the United States between 2011 and the second quarter of 2016 (Bansak et al. 2018a, 327). This dataset was harnessed to train supervised learning models for each location that predicted how likely a refugee was to find a job based off on their background traits (Bansak et al. 2018a, 327). The models were applied to unseen data of refugees who came to America in the third quarter of 2016 to forecast the expected employment success for each newly arrived refugee at each potential placement location (Bansak et al. 2018a, 325-326). The researchers recognized that many refugees come to the United States as part of a larger unit (e.g., family unit) and refer to this unit as a case. As a result, the research team focused on the predicted probability that at least one individual in the unit would obtain employment at a given location 90 days after arrival in the United States

(Bansak et al. 2018a, 326). Cases are then assigned to a location where the average of this metric was maximized. Moreover, researchers attempted to mimic the real-world constraints on the number of refugees that can be sent to each location. In the algorithmic-based assignment process, each location could not accommodate more cases than the number it actually received during the defined time period (Bansak et al. 2018a, 327).

The research team also tested the algorithm in Switzerland where asylum seekers were assigned to one of 26 cantons. The algorithm was trained on data from the Swiss State Secretariat for Migration (SEM), the authority in charge of asylum placement. The method of placement in Switzerland is interesting because SEM relies on "proportional random assignment of cases to locations, and tracks employment outcomes for several years after asylum seekers' arrival" (Bansak et al. 2018a, 327). Due to the long-term observation of asylum seekers, the research team was excited at the opportunity to assess the algorithm's ability to optimize a longer-term employment metric – the employment of a refugee at the conclusion of their third year in Switzerland (Bansak et al. 2018a, 327). They employed data of the 22,159 working age asylum seekers who had been granted subsidiary protection status (Switzerland's biggest refugee category) and came to Switzerland between 1999 and 2012 (Bansak et al. 2018a, 327). The test set were refugees who arrived in 2013. Again, constraints were imposed on the algorithm so that each canton received the same number of cases as it did in actuality, which the law dictates must be proportional to the canton's population (Bansak et al. 2018a, 327).

The Why in the Where

Before we continue with a critical analysis of GeoMatch, it is important to first address if where a refugee is resettled has any significance. Does place matter? Dr. David T Lardier Jr. and

his colleagues investigated the impact of post-resettlement stressors on refugee health outcomes and would argue that it does. They interviewed 290 refugees from Afghanistan, Iraq, Syria, and the Great Lakes Region of Africa who had resettled in the U.S. between 2021 and 2023 (Lardier Jr. et al. 2023, 2). They found that social and economic post-resettlement stressors were associated with high rates of mental and physical health issues (Lardier Jr. et al. 2023, 2). Post-migration stressors include difficulty accessing quality healthcare, "acculturation, language barriers, low socioeconomic status, poor living conditions, and limited social support" (Lardier Jr. et al. 2023, 2). His team also found that mental and physical health episodes increase immediately after resettlement and can even linger several years after (Lardier Jr. et al. 2023, 3).

Depending on the resettlement location, refugees can be exposed to a varying degree and range of post-resettlement stressors. For example, if a refugee is afraid of violence in the community, unable to afford basic resources, and incapable of rebuilding social networks, then the proliferation of stress will lead to poor health outcomes (Lardier Jr. et al. 2023, 3). For instance, the refugees interviewed by Lardier Jr. reported conditions such as cardiovascular dysfunction, weight loss, sleep challenges, headaches, and exhaustion (Lardier Jr. et al. 2023, 7). Furthermore, anxiety due to unemployment and lack of financial assistance in the face of mounting expenses was observed (Lardier Jr. et al. 2023, 7).

A factsheet shared by the U.S. Department of Housing and Development states that "all refugees are eligible for federally funded assistance from the US Department of State and/or the US Department of Health and Human Services' Office of Refugee Resettlement" (U.S. Department of Housing and Development 2024, 1). Examples of included services are cash assistance, job training and placement services, and English language lessons (U.S. Department of Housing and Development 2024, 1). Also, refugees and those granted asylum seeker status

can access healthcare by enrolling in Medicaid and Children's Health Insurance Program (CHIP), (U.S. Department of Health and Human Services, n.d). In addition, local resettlement agencies, community organizations, and faith-based groups are tasked with helping reduce challenges for refugees in their community (U.S. Department of Housing and Development 2024, 1). An Iraqi woman interviewed by Lardier Jr. and his team mentioned that her rent had initially been covered by the federal government (Lardier Jr. et al. 2023, 8).

However, these services are not indefinite. This aid is designed to be short-term with the end goal that refugees become self-sufficient. The Iraqi woman stated that her financial assistance was discontinued after three months of being in the U.S. Her battle to pay rent resulted in emotional distress and countless sleepless nights. Despite the government's assurances of support, other interviewees described how they experienced the denial of government assistance like disability benefits (Lardier Jr. et al. 2023, 8). The researchers urge for the mitigation of post-migration economic and social stressors in order to reduce the risk of refugees developing health issues. GeoMatch focuses heavily on improving employment outcomes and aims to reduce the likelihood of unemployment, a key stressor. The algorithm hopes to intervene before the refugee's arrival, guiding them to a location that provides them with a better chance of obtaining a job.

Implementing GeoMatch

Before we dive into how GeoMatch works when deployed for refugee resettlement, let's first consider how refugee resettlement has primarily been decided in the United States. Refugee Council USA explains that after a rigorous pre-arrival process, the choice of where a refugee is placed in the United States is adjudicated by one of ten U.S. resettlement agencies and their

expansive networks of local partners (Refugee Council, n.d.). The ten national resettlement agencies are Global Refuge, US Conference of Catholic Bishops, Episcopal Migration Ministries, Ethiopian Community Development Council, Church World Service, International Rescue Committee, Bethany Christian Services, HIAS, US Committee for Refugees and Immigrants, and World Relief (Refugee Council, n.d.). The Department of State also collaborates with these agencies. Individual cases are allocated to one of these agencies at "a weekly draft with a randomized order, held in Washington DC" (Bansak et al. 2018b, i). These organizations are bound by constraints when assigning cases across their network of offices. For instance, if refugees have relatives already living in the U.S. who want to sponsor them, resettlement agencies are obligated to place them in close proximity (Bansak et al. 2018b, i). Also, resettlement agencies must ensure that refugees with significant health conditions will be able to receive sufficient medical care. The remaining batch of cases "are assigned weekly on a case by case basis to affiliate sites" with "the smallest proportion of their yearly capacity currently filled" (Bansak et al. 2018b, i).

Helen Pursel, the assistant director for Refugee Pre-Arrival Services at Global Refuge, has been very vocal about how time-consuming and difficult it is to assign refugees to one of Global Refuge's affiliates in over 50 U.S. cities. She has been in this business for 22 years. For a long time, this process involved a group of staff members gathering weekly to examine new cases and their complex requirements, pouring over multiple spreadsheets to isolate a suitable location, and spending many hours manually pairing refugees to different affiliate sites (Immigration Policy Lab, 2023b). The administrative burden was immense. However, she claims that after Global Refuge started using GeoMatch in 2023, her job has become much easier. She starts by describing how inefficient the process used to be:

> I have seen from the beginning, refugee placement starting from faxing over biographical data…I've gone from standing at the fax machine and waiting and then looking at cases and trying to decide where to place them…It would take hours and hours to prepare beforehand and then hours and hours to actually meet and place the cases (Pursel 2024).

Staff extensively review case details about the refuge's primary language, housing or medical needs, and their existing familial ties in the U.S. They ultimately try as best as they can to connect them to the most suitable affiliate sites — all while ensuring that no single area accepts more refugee cases than they can handle.

After Global Refugee began using GeoMatch, Pursel was amazed by how streamlined and quick the previously time-consuming placement process became:

> We are at a new age where we can give all of the information to IPL, have them convert it into an algorithm, and then boil it down for us…Having our daily processes that we go through explained and then entered and refined into an algorithm that helps place people has been truly amazing. I am able to have everything all in one dashboard. Everything I need is right there…I sign in and it takes just a matter of minutes to place cases...I would recommend it to everyone (Helen Pursel 2024).

GeoMatch helps save Helen and her colleagues time and effort. GeoMatch provides recommendations that are essentially instantaneous compared to what was possible before. Its feedback can then help Global Refuge staff and other resettlement agencies to finalize the placement decision. Pursel explains that GeoMatch is able to integrate with Global Refugee's database, analyze case dynamics, and keep track of the locations that currently have the capacity or specialized (health) services for new cases. GeoMatch doesn't replace human decisions but supports staff in their decision-making methods and reduces manual components.

Purported Successes and Benefits of GeoMatch

In the end, the researchers found that algorithmic assignment through GeoMatch undisputedly "increased expected refugee employment" when compared to actual assignment (Bansak et al. 2018a, 327). They stated that even after using different functions and training data from various time periods, "these tests all show considerable gains" (Bansak et al. 2018a, 327). They declared that in the United States, GeoMatch predicts that the median refugee would be twice as likely to obtain a job 90 days after arrival if they had been placed by the algorithm compared to actual assignment (Bansak et al. 2018a, 327). They also observed that algorithmic assignment led to higher employment rates in nearly all sites. According to them, the average employment rate across all locations was 34% without the algorithm and 48% with algorithmic assignment, "which means that the optimized assignment would increase the employment rate above the baseline by roughly 41%" (Bansak et al. 2018a, 327). In Switzerland, they found that the third-year employment rate was 15% without the algorithm and 26% with optimized assignment, and these results "suggest that the data-driven assignment has the potential to increase third-year employment in the Swiss context by about 73%" (Bansak et al. 2018a, 328).

The developers of GeoMatch also released a factsheet that emphasizes GeoMatch's potential long-term impacts. One benefit is that when asylum seekers and refugees get a job and integrate into their community, tax revenue increases while costs associated with social services decrease (Immigration Policy Lab 2023a, 2). In addition, GeoMatch promotes refugee's "financial independence, entrepreneurship, and social mobility" (Immigration Policy Lab 2023a, 2). Furthermore, improved integration translates to the social cohesion required to combat inequality, conflict, and extreme political divides (Immigration Policy Lab 2023a, 2). IPL argues that GeoMatch is particularly innovative because its recommendations are driven by big data and

are efficient. This sounds revolutionary if one considers the traditional approach to deciding where a refugee is resettled. GeoMatch is described as low cost, adaptable to existing systems, easily customizable to any quantifiable integration outcome, complementary to the expertise of placement authorities, universal, and dynamic (Immigration Policy Lab 2023a, 1).

One shouldn't confuse these statistics as evidence of GeoMatch's proven and notable gains. The findings in the main research paper by Bansak et al. must be contextualized. These statistics represent potential gains — projected improvements in employment rate. IPL chooses to save this important clarification for the final three sentences of the last paragraph of a four-page document about GeoMatch:

> Once sufficient data is collected, the research team will be able to determine whether refugees placed through GeoMatch actually fare better than those placed under the previous manual system, as a rigorous impact evaluation is built into the implementation plan. But for now, the tool has greatly simplified the placement decision-making process and ensured that overstretched refugee resettlement agencies like [Global Refuge] are able to obtain immediate, reliable, and data-driven recommendations on where to locate refugees to maximize their chances of success (Immigration Policy Lab 2023b, 4).

In other words, it is important to make a distinction between the processing efficiencies that GeoMatch provides and the notion of improved refugee integration outcomes. The staff at refugee resettlement agencies value GeoMatch's ability to simplify and speed up the task of assignment. These sentiments do illustrate the powerful combination of human decision-making and the streamlining possibilities of artificial intelligence. However, GeoMatch's processing efficiencies don't necessarily mean improved refugee integration outcomes like employment, but we can hope that the former is met with the latter. Time and future evaluations will tell. At this moment, we must honor this difference and proceed with caution.

Concerns, Doubts, and Risks of GeoMatch

There are additional reasons as to why GeoMatch's statistics and claims of improved expected employment outcomes should be called into question. It is still unclear who was represented in the training data. The background characteristics of refugees and asylum seekers included country of origin, language skills, gender, and age. Was the training data skewed towards refugees coming from a particular country? Was the training data skewed towards refugees who already possessed English language skills and/or are highly educated? These questions are left unanswered throughout IPL's research briefs. My hesitancy lies in the serious issue that unrepresentative datasets pose for machine learning algorithms. Bias is introduced when a machine learning algorithm is trained on data in which particular demographic cohorts such as gender and country of origin are heavily underrepresented or overrepresented (Franklin et al. 2024, 2). Answering these questions regarding who was represented in the historical data is key to ensuring that GeoMatch does not reproduce biases entrenched in the training data and translate to suboptimal assignment for marginalized subgroups of refugees.

Moreover, the researchers don't delve into how site-specific data is factored into the equation. On one hand, they assert that geographical context is a key factor that affects refugee integration, providing the example of how a French-speaking refugee experiences greater benefits when assigned to French-speaking cantons compared to those assigned to German-speaking cantons (Bansak et al. 2018a, 325). Bansak et al. emphasize that their models allow "for the discovery of refugee/location synergies" (Bansak et al. 2018a, 326). However, the list of data variables and measures used during the modeling phase does not include relevant location attributes (Bansak et al. 2018b, xii-xiii). Information one would assume to be pertinent to refugee placements such as city data on housing, job sectors, median wages, or census data on proportion

of residents by race, age cohort, and religious affiliation is absent. The small caveat is that GeoMatch does incorporate location-specific limits on the number of cases and distinguish locations that cannot accommodate cases with significant medical or language needs (Bansak et al. 2018a, 328).

To illustrate my concerns, imagine that you are helping your best friend find their perfect match using a dating app. Instead of reviewing X's profile to gauge if their interests and values align with your friend's, you base your decision on the bios of people who have previously matched with X. You focus on how well those past relationships turned out. You don't examine X's profile to learn more about his career, hobbies, or personality traits. While this approach might capture some indirect patterns, it completely overlooks information that could significantly influence whether your friend and X would be a good pair. You would feel like you were missing an important part of the puzzle. Similarly, the GeoMatch algorithm focuses on refugee characteristics, historical data, and a chosen outcome but overlooks important factors and dynamics within the assigned locations that could have a significant impact on integration success in the real world.

Furthermore, the researchers measure the algorithm's potential impact through backtesting; that is, using historical data to evaluate how well the algorithm would have performed in the past. Under this method, the algorithm's predictions are compared to the actual known outcomes from that period. However, backtesting introduces a host of concerns. Present-day conditions don't necessarily mimic past conditions. It is hard for backtesting to address all the factors that determine an algorithm's performance and implementation today or further down the line. If GeoMatch were to be utilized in the future, the expected employment rate predicted by the algorithm could deviate from the backtests. As a result, there is a risk of overshooting or

undershooting the actual employment success rate for newly arrived refugees should they yield to GeoMatch's recommended assignment (Ferwerda 2020, 17). Backtests just suggest that gains from GeoMatch's implementation are possible (Ferwerda 2020, 18).

Therefore, the algorithmic tool must be tested prospectively in order to confirm or deny the demonstrated potential for improvements to employment rate. This can be done through prospective randomized-controlled trials. A randomized-controlled trial would assign new arrivals through the GeoMatch tool and allow for a more rigorous and reliable evaluation of its effectiveness on employment outcomes (Ferwerda 2020, 18). Again, the goal of GeoMatch is to maximize a specified metric (e.g., short term employment). Decisions are based on what past immigrants with similar profiles have experienced. This does not guarantee that the refugee candidate will successfully obtain employment in this location.

In both cases of the U.S. and Switzerland, the research team focused on employment, specifically employment status 90 days after arrival in the U.S. IPL designed GeoMatch to be customizable and able to "optimiz[e] any integration outcome" (Immigration Policy Lab 2023a, 1). The developers of GeoMatch express that GeoMatch is young and still in its early stages. Piloting and testing began in 2020 and IPL launched a prototype of GeoMatch in 2022 (Immigration Policy Lab 2023c, 2). While promising that more is yet to come, the developers praise GeoMatch for its ability learn from and adjust to new data on refugee integration outcomes. However, the research team hyper-fixates on employment status 90 days arrival in the U.S. throughout its reports. Is this the metric we should aim to optimize as a signal of a refugee's successful integration? I argue no.

I understand the convenience of using this benchmark given that U.S. resettlement agencies are required to track and report it. Nevertheless, this narrowly defined integration

outcome does not address other elements such as a refugee's long-term employment, physical and mental health, educational attainment, or earnings. These factors are also important, if not more, in exploring a refugee's ability to confidently integrate into a new community. GeoMatch's overemphasis on employment overlooks the significance of social network building for refugee integration. Interpersonal relationships foster a sense of belonging and emotional support for refugees. Social networks also provide practical support that manifests as help with childcare, transportation, and financial struggles (Boateng et al. 2024, 3). By preventing isolation, social networks play a vital role in mitigating many post-resettlement stressors and improving refugee experiences. Recommended locations may even negatively impact longer-term or additional outcomes. This metric also doesn't inform us about what types of jobs refugees are getting. It is not clear if they are well-paid or if they maximize the refugee's skillset.

If space is limited, who makes it through to their personalized promise land? GeoMatch imposes restrictions to simulate real-world capacity constraints and aims to create a balanced distribution across resettlement locations. But what if the models truly showed that there is one location that a diverse group of refugees would just thrive in? In other words, the algorithm would have spit out the same city for many refugees. After imposing algorithm constraints, it is unclear who the chosen people who make the cut and get the coveted location assignment would be.

There are clear ethical considerations that must be made when proposing the use of machine learning and artificial intelligence tools for refugee resettlement and integration. One example is the issue of consent. Bansak's team mentions that GeoMatch was trained using de-identified data from a resettlement agency. Past and newer refugees must be informed explicitly about how their data will be used and shared. They must have the right to retain control over

their data, how it is utilized, and which parties have access to it. Obtaining free and informed consent from refugees to be a part of this algorithm system should be the expectation. Refugees or asylum seekers awaiting placement must be informed about the use of an algorithmic tool for placement suggestion. De-identification is not a flawless privacy protection method. When de-identified datasets are merged with other datasets, especially large ones, there is a risk for re-identification (Miller 2021).

Conclusion

While the developers of GeoMatch talk about its potential and its promises to help government and resettlement agencies grapple with the magnitude of the present-day refugee crisis, we cannot be too hasty to take their claims at face value. Throughout my investigation into available information on GeoMatch, several questions were left unanswered. Hursel's testimony does help illustrate GeoMatch's power to help overburdened placement officers with cumbersome aspects of the placement process through the delivery of efficient assignment recommendations. Even so, it is too early to signal GeoMatch's measurable impact on refugee integration outcomes. The developers of GeoMatch must incorporate more testing, prospective randomized-controlled trials, and a more inclusive set of variables.

Algorithmic decision-making systems are becoming more widespread and are "increasingly used by public administrations to assist human decision-making processes in public policy—including migration and refugee policy" (Bither and Ziebarth 2021, 5). Algorithmic tools are often presented as a data-driven, technological remedy to make existing systems more efficient. Yet, stakeholders frequently don't know how these systems work. They rarely understand how these tools were trained, tested, or implemented and the limitations

associated with each step in their development. There must be a concerted effort to interrogate the validity of their impact assessments including questions about the data, potential bias, and the metric that will be optimized.

Given the high stakes associated with the use of AI for refugee resettlement and integration, it is important to look beyond statistics which reflect what a metric related to refugee integration outcomes could be under algorithmic assignment based on historical data. Claims of potential gains should be approached with diligent scrutiny, especially when the algorithm has not been tested prospectively. The deployment of technologies like GeoMatch has real implications for refugees, a highly vulnerable population. Therefore, it is crucial that the developers of such tools are held accountable for addressing questions about the tool's data sources and design.

References

Bansak, Kirk, Jeremy Ferwerda, Jens Hainmueller, et al. 2018a. "Improving Refugee Integration through Data-Driven Algorithmic Assignment." *Science* 359, no. 6373 (2018): 325–29. https://www.jstor.org/stable/26401311.

Bansak, Kirk, Jeremy Ferwerda, Jens Hainmueller, et al. 2018b. "Supplementary Materials for Improving Refugee Integration through Data-Driven Algorithmic Assignment." *Science* 359, no. 6373: 325–29. https://www.jstor.org/stable/26401311.

Bither, Jessica, and Astrid Ziebarth. 2021. "Automating Decision-Making in Migration Policy: A Navigation Guide." Migration Strategy Group on International Cooperation and Development. https://www.bosch-stiftung.de/sites/default/files/publications/pdf/2021-11/Automating%20Decision-Making%20in%20Migration%20Policy_Bither%20and%20Ziebarth.pdf

Boateng, Godfred O, Karin Wachter, Roseanne C. Schuster, Tanya L. Burgess, and Mary Bunn. 2024. "A Scoping Review of Instruments Used in Measuring Social Support among Refugees in Resettlement." *International Journal of Environmental Research and Public health* 21(6): 805. doi:10.3390/ijerph21060805

Chishti, Muzaffar, Kathleen Bush-Joseph, Madeleine Greene. 2024. "How the Rebuilt U.S. System Resettled the Most Refugees in 30 Years." Migration Policy Institute. https://www.migrationpolicy.org/article/rebuilt-us-refugee-resettlement-biden#:~:text=The%20100%2C034%20refugees%20resettled%20in,the%20U.S.%20refugee%20program's%20history

Ferwerda, Jeremy, Nicholas Adams-Cohen, Kirk Bansak, et al. 2020. "Leveraging the Power of Place: A Data-Driven Decision Helper to Improve the Location Decisions of Economic Immigrants." Working Paper No. 20-06. https://arxiv.org/pdf/2007.13902

Franklin, Gillian, Rachel Stephens, Muhammad Piracha, et al. 2024. "The Sociodemographic Biases in Machine Learning Algorithms: A Biomedical Informatics Perspective." *Life (Basel, Switzerland)* 14(6): 652. doi:10.3390/life14060652

Global Refuge. n.d. "Understanding the Long Welcome." Accessed February 27, 2025. https://www.globalrefuge.org/what-we-do/refugee-resettlement/how/

Pursel, Helen. 2024. *Momentum*, Stanford, February 1. Video. https://momentum.stanford.edu/stories/using-machine-learning-to-help-refugees-succeed

Immigration Policy Lab. 2023a. "GeoMatch: Connecting People to Places." https://immigrationlab.org/content/uploads/2023/07/Factsheet-GeoMatch-2023.pdf

Immigration Policy Lab. 2023b. "Harnessing AI for Refugee Resettlement." https://immigrationlab.org/content/uploads/2023/10/GeoMatch-Case-Study-2023.pdf

Immigration Policy Lab. 2023c. "Implementing the Algorithm." Accessed February 27, 2025. https://immigrationlab.org/geomatch/streamlining-refugee-resettlement-geomatchs-partnership-with-global-refuge/

Lardier, David T., Jr., Julia Hess, Lucas Winter, and Jessica R Goodkind. of Post-Resettlement Stressors and Access to Healthcare on Health Outcomes in Recently Resettled Refugees in the United States". *The American Journal of Orthopsychiatry* 93 (6): 516-531. doi: 10.1037/ort0000697.

Miller, Katharine. 2021. Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/news/de-identifying-medical-patient-data-doesnt-protect-our-privacy

Ozkul, Derya. 2023. "Automating Immigration and Asylum: The Uses of New Technologies in Migration and Asylum Governance in Europe." Oxford: Refugee Studies Centre, University of Oxford. https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf

Refugee Council USA. n.d. "Resettlement Process." Accessed February 27, 2025. https://rcusa.org/resources/resettlement-process/

United Nations High Commissioner for Refugees. 2024. "Mid-Year Trends 2024." Accessed February 27, 2025. https://www.unhcr.org/mid-year-trends-report-2024

U.S. Department of Housing and Development. 2024. "Fact Sheet on Renting to Refugees and Other Eligible Newcomers." Accessed February 27, 2025. https://acf.gov/sites/default/files/documents/orr/landlord-guide-factsheet.pdf

U.S. Department of Human and Health Services. n.d. “Coverage for lawfully present immigrants.” Accessed February 27, 2025. https://www.healthcare.gov/immigrants/lawfully-present-immigrants/