

AI translation is jeopardizing Afghan asylum claims

Cost-cutting translations are introducing errors and putting refugees at risk.

By ANDREW DECK

19 APRIL 2023

TRANSLATE ▾

- Neural machine translation tools are on the rise, and increasingly being used by for-profit government contractors and aid organizations working with refugees and migrants.
- Machine translations of Pashto and Dari, in particular, are riddled with errors that have introduced confusion into already complex immigration processes, and led to the rejected asylum claim of at least one Afghan refugee.
- Advocates for refugees spotlight the importance of human attentiveness in translation work, and caution against relying on machine translations when working in high-stakes scenarios, including asylum court.

In 2020, Uma Mirkhail got a firsthand demonstration of how damaging a bad translation can be.

A crisis translator specializing in Afghan languages, Mirkhail was working with a Pashto-speaking refugee who had fled Afghanistan. A U.S. court had denied the refugee's asylum bid because her written application didn't match the story told in the initial interviews.

In the interviews, the refugee had first maintained that she'd made it through one particular event alone, but the written statement seemed to reference other people with her at the time — a discrepancy large enough for a judge to reject her asylum claim.

After Mirkhail went over the documents, she saw what had gone wrong: An automated translation tool had swapped the "I" pronouns in the woman's statement to "we."

Mirkhail works with Respond Crisis Translation, a coalition of over 2,500 translators that provides interpretation and translation services for migrants and asylum seekers around the world. She told *Rest of World* this kind of small mistake can be life-changing for a refugee. In the wake of the Taliban's return to power in Afghanistan, there is an urgent demand for crisis translators working in languages such as Pashto and Dari. Working alongside refugees, these translators can help clients navigate complex immigration systems, including drafting immigration forms such as asylum applications. But a new generation of machine translation tools is changing the landscape of this field — and adding a new set of risks for refugees.

Machine translation has been on the rise since the introduction of neural network techniques, similar to those used in generative artificial intelligence. In 2016, Google launched its first neural machine translation system. Today, when subtitling films for streaming companies or drafting documents for law firms, some of the most established global translation companies use neural machine translation in their workflow in an effort to cut costs and boost productivity. But like the new generation of AI chatbots, machine translation tools are far from perfect, and the errors they introduce can have severe consequences.

Companies working in this space generally recognize the danger of pure automation, and insist that their tools be used only under close human supervision. “Machine-learning translations are not yet in a place to be trusted completely without human review,” said Sara Haj-Hassan, chief operations officer of Tarjimly, a nonprofit startup that connects refugees and asylum seekers with human volunteer translators and interpreters, to *Rest of World*. “Doing so would be irresponsible and would lead to inequitable opportunities for populations receiving AI translations, since mistranslations could lead to the rejection of cases or other severe consequences.”

The unmet demand, however, is undeniable. Tarjimly, which currently works with over 250 language pairs, saw a fourfold increase in requests for Afghan languages in 2022, according to the organization's impact report.

Similar concerns have been raised over generative AI tools. OpenAI, the company that makes ChatGPT, updated its user policies in late March with rules that prohibit the use of the AI chatbot in “high-risk government decision-making,” including work related to both migration and asylum.

“Machine-learning translations are not yet in a place to be trusted completely without human review.”

The stakes for getting translations right can be grave for asylum seekers filling out applications. “One of the things that we see frequently is pointing to small technicalities on asylum

applications,” Ariel Koren, the founder of Respond Crisis Translation, told *Rest of World*. “That’s why you need human attentiveness. The machine, it can be your friend that you use as a helper, but if you’re using that as the ultimate [solution], if that’s where it starts and ends, you’re going to fail this person.”

That is particularly true for work with Afghan refugees who speak Pashto and Dari — languages native to tens of millions of Afghans around the world. The United Nations High Commissioner for Refugees (UNHCR) estimates that over 6 million Afghans were displaced by the end of 2021 alone, including those displaced following the U.S. withdrawal from Afghanistan and the Taliban’s return to power. At the same time, AI language tools for Pashto have lagged behind more dominant languages like English and Mandarin. The latter are considered “high-resource” languages, with a large amount of texts available online compared to a language like Pashto.

It is difficult to say how prevalent machine translation is in the immigration system, but there’s clear evidence it is being used. In 2019, ProPublica reported that U.S. Citizenship and Immigration Services (USCIS) officers were instructed to use Google Translate to vet the social media accounts of asylum applicants. Major translation companies like LanguageLine, TransPerfect, and Lionbridge have contracts with U.S. federal immigration agencies, some totaling millions of dollars. Each of these companies advertises machine translation in its suite of services. Ultimately, it is up to each agency and department whether they opt in or out of these tools in their day-to-day operations.

6 million

The estimated number of Afghans displaced by the end of 2021.

UNHCR

At the same time, providers are actively pitching refugee organizations to integrate machine translation into their work. International Refugee Assistance Project (IRAP), a nonprofit that offers legal support to refugees in Afghanistan and Pakistan, received multiple solicitation emails from a for-profit government contractor concerning machine translation.

One of those emails, sent by U.K.-based translation company The Big Word, pitched WordSynk: the company’s signature product, described on its website as “utilising Machine Translation, AI, and translation memory to leverage high-quality, cost-effective outcomes.” IRAP never responded to The Big Word’s sales pitch, but the company lists the U.S. Department of Defense, the U.S. Army, and the U.K. Ministry of Justice among its clients. An internal document, reviewed by *Rest of World*, lists Pashto and Dari among The Big Word’s “core language” offerings for government customers.

The Big Word did not respond to *Rest of World's* request for comment.

Whether automated or not, translation flubs in Pashto and Dari have become commonplace. As recently as early April, the German Embassy to Afghanistan posted a tweet in Pashto decrying the Taliban's ban on women working. The tweet was quickly ridiculed by native speakers, with some quote tweets claiming that not a single sentence was legible.

"Kindly please don't insult our language. Thousands [of] Pashtun are living in Germany but still they don't hire an expert for Pashto," posted one user, researcher Afzal Zarghoni. The German Embassy later deleted the tweet.

Seemingly trivial translation errors can sometimes lead to harmful distortions when drafting asylum applications.

"[Machine translation] doesn't have a cultural awareness. Especially if you're doing things like a personal statement that's handwritten by someone," Damian Harris-Hernandez,

"Kindly please don't insult our language."

co-founder of the Refugee Translation Project, told *Rest of World*. "The person might not be perfect at writing, and also might use metaphors, might use idioms, turns of phrases that if you take literally, don't make any sense at all."

Based in New York, the Refugee Translation Project works extensively with Afghan refugees, translating police reports, news clippings, and personal testimonies to bolster claims that asylum seekers have a credible fear of persecution. When machine translation is used to draft these documents, cultural blind spots and failures to understand regional colloquialisms can introduce inaccuracies. These errors can compromise claims in the rigorous review so many Afghan refugees experience.

Dari and Pashto are currently Refugee Translation Project's most frequently requested languages, according to Harris-Hernandez. Despite the demand, the organization refuses to use automated translation tools, relying exclusively on human translators.

"There's not really a lot of advantage to [machine translation]. The advantage comes in if you don't know the language and you're trying to translate something for a customer," Harris-Hernandez said, explaining that the incentives look different for his organization compared to many for-profit providers. "The only thing that matters is the money that comes in."

Muhammed Yaseen, a member of the Afghan team at Respond Crisis Translation, told *Rest of World* that organizations are banning the use of machine translation for good reason. He claims the machine tools he's tested are unable to translate certain words, such as the terms for some

relatives in Dari dialects, and specialized words like military ranks that can be vital to the asylum applications of former U.S.-allied soldiers.

“If we use machines for Afghans, I think we would be unfair to them,” Yaseen said. “I really feel that if we rely on machines, I [am] expecting at least 40% of our decision making on the asylum applications for refugees would be incorrect.”

Andrew Deck is a reporter at Rest of World.