

Working Paper

Does the United Nations need agents?

Testing the role of AI agent generated personas in humanitarian action

Eduardo Albrecht

Table of contents

| | |
|--------------------------|-----------|
| Executive summary | 3 |
| Introduction | 3 |
| Why we need agents | 3 |
| Why we need caution | 5 |
| Anthropologist agents? | 6 |
| Case Study | 7 |
| Experiment outline | 7 |
| Data collection | 8 |
| Results | 8 |
| Conclusions | 9 |
| The Null Hypothesis | 9 |
| Guardrails | 10 |
| Recommendations | 10 |
| Appendix | 11 |

Author bio: Eduardo Albrecht is a Senior Fellow at United Nations University Centre for Policy Research (UNU-CPR), associate professor at Mercy University and adjunct professor at Columbia University’s School of International and Public Affairs and the City University of New York’s (CUNY) Colin Powell School for Civic and Global Leadership. He is also the author of *Political Automation: An Introduction to AI in Government and its Impact on Citizens* (Oxford University Press 2025).

ISBN: 978-92-808-6648-3

Acknowledgments: The author would like to thank Eleonore Fournier Tombs, Head of Anticipatory Action and Innovation at UNU-CPR for her valuable insights and guidance; Liselott Hitz, previously with Doctors Without Borders/Medicins Sans Frontiers, as advisor on applicability to the work of humanitarian and relief organizations, and research lead on the AI generated personas; Matteo Bastioni, art director at Bold Shots, for designing and generating the AI avatars; Columbia University students Maud Ruth Yaiche and Marvin Limpijankit for co-leading the CUNY research team; Columbia University students Chanjin Park, Jiaxiong Hu, Ariella Kissin, McKynzie Nichole Clark and Julia Yoon as research assistants on scoping the exploratory potential of AI agents in humanitarian work and Mercy University Capstone student Anna Deutscher and City College Masters of International Affairs students Laetitia Sia and Nathan Kamad as research assistants on agent knowledge base and avatar creation. Special appreciation goes to the UNU-CPR operations team, Jack Durell, Tess Pilkington and Fabiana Coccia, without whom this paper and the accompanying workshop event would not have been possible.

Executive summary

It is now possible to create detailed digital replicas of people. Using an approach that leverages a mix of AI technologies we can generate live “personas” that act and opine in a manner reasonably similar to their real counterparts. This paper examines the potential role of such personas in humanitarian, peacekeeping and development contexts.

Drawing from a pilot study employing two persona systems – “Ask Amina” and “Ask Abdalla” – this paper considers whether this technology may enhance data collection in high-risk environments, improve diplomatic negotiation and mediation training and facilitate community dialogue simulations.

Amina simulates a refugee living in Chad’s Metche camp, while Abdalla represents a Rapid Support Forces combatant in eastern Sudan. Both are designed to authentically represent their respective populations through specifically curated knowledge bases – which in turn provide the foundation for the synthetically generated opinions. **Both personas are available for public interaction and conversation at askamina.ai.**

The core technology behind these simulations is an AI agent system that carefully curates the knowledge base for each persona. Through a fusion of autonomous AI agents and generative AI, the resulting personas can reproduce conversation patterns and responses that portray realistic human-like characteristics.

The study reveals significant advantages over traditional methods: personas enable rapid data collection in dangerous or time-sensitive situations, overcome language barriers and interpreter bias common in conventional surveys, permit negotiators and mediators to practice different strategies before engagement and allow policymakers to simulate dialogues among hard-to-access populations.

Findings suggest that persona-based methods can be accurate in representing real person voices, and thus may be of value, particularly in contexts where traditional data collection is impractical or impossible. Amina’s accuracy was tested with 20 questions from four surveys outside her knowledge base. Results showed she correctly answered 16 questions (80 per cent accuracy) across nutrition, refugee assistance and Sudan conflict topics.

However, significant ethical challenges are identified. These include potential reinforcement of power imbalances, risks of misrepresentation and open questions surrounding data rights and governance. The conclusion of this paper proposes essential guardrails

for responsible deployment, including transparent documentation of AI decision-making processes and the adoption of a gradual, pilot-based approach that allows organizations to develop best practices before scaling.

Critically, this paper strongly recommends connecting AI agent generated persona systems with representatives from the communities they portray, giving these stakeholders meaningful authority to approve, modify or drive system development based on community needs. For example, a persona like Amina should be reviewed and approved by refugee-led organizations from Darfur.

Deployed responsibly, personas can improve access to critical intelligence and thereby enhance the effectiveness of organizations that have a mandate to meet the needs of vulnerable populations such as the United Nations. Careful integration of these tools guided by ethical principles can also potentially amplify local voices and significantly transform humanitarian, peacekeeping and development efforts.

Introduction

Why we need agents

This paper illustrates a case study with two AI agent generated persona systems, one called “Ask Amina” and the other “Ask Abdalla.” The first is designed to create an accurate digital representation of a refugee living in a camp in Chad. The second creates a digital replica of a combatant leader in the Rapid Support Forces (RSF), a group active in the southeastern part of Sudan from which many refugees are fleeing. Both systems combine digital avatars with large language models (LLMs), retrieval-augmented generation (RAG) and AI agent curated knowledge bases specifically designed to maximize the representativeness of each persona.

The effectiveness of Amina’s representation is measured by comparing “her” responses to survey questions against those provided by actual members of the refugee population. A sample of a negotiation with Abdalla, as well as a conversation between Amina and Abdalla, are qualitatively assessed. The experiment methodology and results are described in the case study section of this paper with further details in the appendix. Both personas can be experienced and tested at askamina.ai. Readers are invited to interact with either persona on any desired subject.

The primary objective of Ask Amina is to enable humanitarian and relief workers to ask questions about refugee experiences, needs and sentiments, and receive responses that closely mirror real refugee perspectives. The objective of Ask Abdalla is to enable diplomats,

negotiators and mediators to practice their skills with a persona that responds in ways consistent with a real combatant's known behavioural patterns. The two personas are also put in conversation with each other as a preliminary test of a virtual AI-based community dialogue simulation.

Although humanitarian work is driven by good intentions, some projects fall short of their goals because they do not adequately account for local preferences and cultural nuances among the populations they serve. Making timely and informed decisions is key, but without the right information many projects can fail. Some famous examples are PlayPumps in South Africa,¹ housing projects in Haiti post the 2010 earthquake² and the One Laptop Per Child initiative.³

Current methods to understand local needs include traditional surveys, focus group discussions and questionnaires. These are time consuming and resource intensive. Participation can be limited due to time constraints, geographical remoteness or a lack of willingness to participate. Additionally, there could be an incentive to provide false or incomplete information when replying, especially if respondents believe that a certain type of answer may benefit them in some way. A representative from an international organization coming into a new cultural context may get a very different response than, for example, a local asking another local the same question in a private conversation.

Another problem is translation. Often, the person conducting the evaluation does not speak the local language and therefore interpreters are needed. Much of what is being communicated can be lost in translation. Research has found that interpreters in humanitarian contexts are seldom professionals. They are instead often untrained and selected because of their knowledge of the local language and that of the organization conducting the survey, most often English. Trained interpreters, on the other hand, must adhere to a code of ethics that assures accuracy, impartiality and confidentiality.⁴

Traditional needs assessments often require significant time. This becomes particularly challenging during humanitarian crises, such as disease outbreaks, where rapid response is crucial. For instance, during the devastating 2014–2016 Ebola outbreak in West Africa, the disease spread rapidly across borders, making conventional data collection through interviews and surveys both dangerous and impractical.⁵ Lacking knowledge of the local communities' cultural and historical approaches to public health, aid workers struggled with containing the outbreak.⁶

The use of AI agent generated personas – that can

authentically simulate human behavior and opinions – is a potential game changer. These personas are designed to respond and interact in ways that closely mirror their real-world counterparts, reflecting not just surface-level behaviors but also the underlying attitudes, beliefs and decision-making patterns of specific demographic groups. If carefully calibrated, persona-based surveys could be used to conduct rapid, virtual needs assessments that determine the type of intervention that is needed, how to best allocate resources and what to prioritize. This type of virtual assessment can help quickly – and cheaply – identify potential challenges to the project or potential security risks for staff.

By enabling dynamic conversations rather than just passive data mining, personas provide a considerable advantage over commercial AI-based online sentiment analytics platforms that rely on natural language processing (NLP). While NLP can only examine existing opinions, these personas can engage in rich dialogue about hypothetical scenarios, offering predictive insights into how people might respond to future activities and actions.⁷

For instance, AI agent generated personas can be integrated in diplomatic negotiation and mediation simulations. By creating accurate digital replicas of key stakeholders – from local community leaders to government officials and military leaders – organizations can dramatically improve their negotiation strategies before critical meetings. This capability is especially valuable in complex peacekeeping operations where securing protection of civilians often requires delicate negotiations with multiple parties.

During the Yemen crisis, UN agencies as well as international aid organizations could have used persona-based simulations to better navigate negotiations with various armed groups controlling different regions, practicing approaches that might have secured wider access to vulnerable populations. Similarly, in South Sudan, peacekeeping operators could have used these simulations to prepare for negotiations with local authorities about establishing protection of civilian sites, potentially leading to faster and more effective agreements about safe zones.

Another example of an advantage over legacy sentiment analytics platforms is that multiple personas can be used to run a virtual dialogue between different segments of a population. These simulations can reveal potential unintended consequences of peacebuilding and development initiatives and highlight opportunities for programme optimization that might be missed in traditional planning processes. For example, a virtual dialogue between different community members might reveal that a proposed sustainable development financing project would inadvertently exclude certain

vulnerable groups, allowing organizations to modify their approach before implementation.

The World Food Programme could have used such simulations in Myanmar to understand how their cash assistance programmes might affect local market dynamics and community power structures. In Afghanistan, NGOs could have employed virtual community dialogues to identify how gender-segregated education initiatives might be structured to gain broader community acceptance while still achieving their core objectives.

In general, persona-based dialogues can ensure that traditionally overlooked voices are heard and considered in programme design, leading to more inclusive and effective interventions. Furthermore, they can also be used to test dialogue between groups that might otherwise not have opportunity to talk in the real world, surfacing important points of agreement and disagreement that might otherwise not have emerged.

In sum, these methods can provide a cost-effective complementary approach to traditional methods. Indeed, in some scenarios they can be used instead of older methods, for example in situations of high risk. Finally, it should be noted that given the low cost of running persona-based surveys, negotiation simulations or virtual policy dialogues, these could be performed at multiple points in a project's life cycle without incurring additional expenses.

Why we need caution

AI's capacity to replace human judgment poses risks to democratic processes and human dignity.⁸ In deliberative contexts, experts warn of ethical concerns that we must be mindful of. AI could manipulate outcomes through biased summarization of conversations or pollute discussions with harmful ideas. Crucially, the potential replacement of humans with their "digital twins" in deliberations raises concerns about the loss of human agency in shaping society.⁹

As AI systems become more prevalent, so too does the need for policymakers to understand their influence and establish governance principles. It is crucial that AI technologies are used in ways that do not undermine individual autonomy or exacerbate power imbalances. Underlying these important ethical concerns are technical questions: Is AI accurate in representing real human perspectives and concerns? And, how closely can it reflect a person's actual opinions? The truth is we do not know because the technology is quite new and technical capacity is still evolving.¹⁰

Studies are emerging that claim it is possible to train

a LLM to express an opinion just as a certain person would.¹¹ One paper by Argyle et al. (2023) using LLMs to simulate human survey responses found that with careful prompting, LLMs can accurately mirror the political preferences and attitudes of specific demographic groups. This concept, which they term "algorithmic fidelity," suggests AI systems might effectively capture nuanced public opinion.¹² Supporting this idea, another study by Sanders et al. (2023) revealed strong correlations between LLM-generated responses and actual survey data from US populations.¹³

Recent research by Anthis et al. further demonstrates that LLM social simulations show promise as a method for understanding human behaviour. While acknowledging limitations, the authors argue that LLM simulations can already be used for exploratory research such as pilot studies in psychology, economics and marketing. In the humanitarian context, such simulations could help organizations better understand diverse perspectives during crises, potentially offering a solution to the interpretation biases and participation challenges outlined above.¹⁴

But LLMs come with myriad limitations. LLMs will lack knowledge of matters underrepresented in their training data and therefore struggle with accuracy when replicating the behavior of more "hidden" demographic groups, or they may be unable to adjust their responses according to more current events. LLMs also struggle with transparency and interpretability, making it difficult to trace outputs back to specific sources or to fully understand the reasoning behind their responses.

And then there is the issue of bias. A study by Boelaert et al. (2024) identified significant limitations in LLM performance in this regard. Their research revealed a persistent "machine bias" in LLM outputs, characterized by limited response variation and poor alignment with specific social groups. They found that AI-generated responses appeared to reflect the models' internal workings rather than genuine socio-demographic patterns.¹⁵ This critique raises serious questions about whether LLMs can reliably capture the complex, layered nature of human opinion, particularly in dynamic social contexts.

In light of these challenges, one approach is to equip LLMs with RAG capabilities. RAG connects the LLM to specialized domain-specific information. When generating responses, the LLM searches this knowledge base for the most relevant information. By augmenting the LLM's abilities with RAG, the composite system can better understand social nuances and adapt to evolving information.¹⁶

While certainly augmenting LLM capacities, combining LLMs with RAG alone still does not guarantee we can

faithfully represent a person's worldview. We may think of RAG as giving an LLM access to an additional database – it enhances the model's knowledge but does not give it the ability to navigate an information environment as a person would. To do that additional components are needed.

The composite system must be able to “independently” interact with an information environment. For that to occur, we must add several key technical components in addition to LLM and RAG, including a decision-making logic for choosing between actions; short-term, long-term, and episodic memory; state and session management systems; input/output handling; feedback and learning systems; performance monitoring; success/failure tracking; and adaptation mechanisms.¹⁷

Taken together, these create an AI agent. An AI agent is an autonomous system that understands a context and is capable of taking an independent action.¹⁸ For example, you can ask an AI agent to autonomously figure out which emails to reply to, play a video game for days in search of glitches, analyse real-time market data to execute trades, automate content moderation across digital platforms, or schedule a team's work week. This all occurs without human oversight because an agent is adaptable – it can learn and change its behavior based on its own experience.¹⁹

This is the level of sophistication needed for personas to be effective in the specialized contexts discussed in this paper. Underneath the LLM and RAG capacities, they must have an additional agentic capacity that enables them to independently navigate an information environment and find the most relevant knowledge to help respond to a survey question, take a position in a negotiation exercise, or express a preference during a simulated dialogue. To date, this has not been systematically tested and of course merits caution.

Anthropologist agents?

This paper proposes that for personas to be sufficiently representative of real persons to be of value, they must be underpinned by AI agents that act as “investigators” dedicated to studying a person and their context. These dedicated AI agents must be designed to interact with that person's cultural milieu and weave a social scientific portrait out of that experience. In a way, they must become similar to anthropologists.

Each individual AI agent must be able to go beyond mere data collection and create a meaningful narrative that captures the unique voice and lived experience of the person they are seeking to represent. They must be able to surface nuanced socio-political patterns that are not captured in the data itself and tease out psychological

states that are more than the sum of the data available. They must be able to create new knowledge.²⁰

This ensures that social science insights can be effectively integrated into the persona's output and that decision makers are provided with relevant and actionable insights. However, even if we do approach authentic representation using this technique, the deeper question remains: Can this “anthropologist agent” still maintain the field's commitment to ethical and respectful cultural engagement? Can it avoid colonial biases and be mindful of past and emerging power dynamics?

In order to maintain such commitment, anthropologist agents would need to incorporate an understanding of historical context, social hierarchies, systems of inclusion and exclusion, kinship systems, cultural taboos and, of course, understand multiple languages.²¹ To develop this subtle and multi-layered understanding, anthropologist agents – just like real anthropologists – must have four key skills:

- 1) They must know how to autonomously connect to the right existing literature and past research conducted on the population in question; in essence, they need to have advanced library research skills.
- 2) They must know how to collect and categorize all relevant cultural artifacts produced by the people they are studying; in other words, they must be able to conduct fieldwork.²²
- 3) They must be able to critically analyse all this information and then organize it in a way that is culturally responsive. To do that, they need to be able to draw conclusions for each piece of literature or cultural artifact based on a clear research ethics framework.
- 4) They must be able to continuously update and re-organize this knowledge base, based on the evolving information environment and the nature of the interactions that users have with the personas.

Each of these skills must be built into the AI agent's programming. Above, we discussed that an LLM may utilize RAG to supplement its output with additional information. RAG accomplishes this by accessing an external knowledge base containing relevant documents and specialized content that may be more current or detailed than the LLM's training data. RAG, however, does not create its own knowledge base, this must be gathered for it. The anthropologist agent's role is, in effect, to autonomously create and curate this knowledge base.

The anthropologist agent must know how to independently connect to the right existing literature,

collect the right cultural content and arrange this material in a manner that is ethnographically responsible before it is fed to the LLM (via RAG) for interaction with the human interlocutor. This knowledge base will be different for each individual person that is being represented.²³

The following section presents an exploratory pilot experiment that takes initial steps toward realizing this framework. While far from achieving the sophisticated capabilities of a fully-developed system, this early experiment may offer insight into the practical requirements needed for building and deploying more comprehensive systems that may one day support critical activities like needs surveying, enhance the effectiveness of negotiation and mediation training and enable realistic simulation of community dialogues.

It will also seek to better understand the immediate and long-term ethical challenges involved. Though modest in scope, this early experiment may offer insight into key matters such as ensuring balanced representation of diverse perspectives, maintaining appropriate boundaries between simulation and reality, and preserving (amplifying?) human agency in decisions that affect communities in humanitarian, peacekeeping and development contexts.

Case Study

Experiment outline

The goal of this experiment is to create digital replicas of two hypothetical persons, one living in a refugee camp in Chad, called Amina, the other called Abdalla, a member of the RSF paramilitary group living in eastern Sudan.²⁴ Through an avatar interface, users will be able to ask questions – such as “What type of challenges do you encounter with access to clean water?” or “What would be needed for you to provide safe passage for the civilian population?” – and receive responses that are based on real data. An avatar interface is a digital representation that serves as a visual intermediary between users and computer systems. It can mimic human appearance and behaviour to create more engaging user interactions.²⁵

Both the Amina and Abdalla avatars were created using HeyGen.²⁶ HeyGen relies on Open AI’s LLMs to animate the video avatars, and allows for linking via RAG to an external database where the knowledge bases curated by the anthropologist agent are uploaded. Both avatars can be visited and trialled at the following website: askamina.ai. Readers are welcome to visit this website and have conversations with Amina and/or Abdalla on whatever topic is preferred.

The main hypothesis is that by combining the generative capabilities of an LLM and the information retrieval capabilities of RAG with a highly relevant knowledge base autonomously curated by an anthropologist agent, users will be able to conveniently explore answers that may be reasonably similar to those of a real person.

To test this hypothesis, one of the two avatars, Amina, will be fed questions that were asked to the same population in real life – the answers to which are not included in the knowledge base²⁷ – and the delta between the avatar and the actual human responses will be measured. The smaller the delta, the more accurate the avatar. Abdalla’s responses, on the other hand, will be qualitatively assessed. A brief conversation between the two will also be qualitatively assessed.

Our hypothetical Amina is a middle aged woman living in the Metche refugee camp in Chad. According to United Nations High Commissioner for Refugees (UNHCR) reports, Chad is host to the most asylum seekers (by volume) of any other Sahel nation.²⁸ Despite being one of the poorest countries in the world, Chad hosts 1.3 million forcibly displaced persons, including more than 1 million refugees. The majority are Sudanese now living in eastern Chad.²⁹ The Metche camp is approximately two hours south of the border town Adré, and hosts mostly refugees of Masalit ethnicity fleeing the conflict in Darfur.

The ongoing conflict in Sudan is seeing many refugees continue to cross the border every day. Adré, in Chad, is where most of the arriving refugees first settle before being relocated to a camp.³⁰ The continued influx of refugees, asylum seekers and forcibly displaced people in the country has resulted in significant strain on the Chadian refugee camps. The Metche camp, one of six in the area, is overcrowded and regularly faces dire water and food shortages.³¹

This influx is exacerbating the strain on resources in a region with already high levels of multi-dimensional poverty which is acutely vulnerable to climate-related shocks. As the camps are saturated, many new arrivals are settling in neighbouring villages, requiring UNHCR and other aid organizations to adapt their response to needs both within and outside the camps.

Our hypothetical Abdalla is a man in early adulthood that commands an RSF unit operating in the rugged terrain in the area around El Geneina, Sudan. The RSF maintains control over strategic transportation routes connecting Darfur with Sudan’s eastern and central provinces. While officially part of the RSF chain of command, local units maintain significant autonomy in daily operations.³²

RSF militias, predominantly from Darfuri Arab groups, regularly engage in sporadic fighting with Sudanese Armed Forces and local resistance groups. Relationships are fraught in Darfur as many local communities have suffered under RSF occupation. RSF actions have been deemed crimes against humanity by Human Rights Watch and others, and hundreds of thousands of civilians have been displaced, including women like the one Amina represents.³³

If Amina works, “her” rapid responses could be of great value. For example, they could be used to quickly make a case to donors (often in very different locations and with very little time) on what population needs to be prioritized when earmarking aid to the region. If Abdalla works, “his” responses could help negotiators and mediators prepare for more targeted real-world engagement.

The experiment will also help identify emergent ethical and regulatory risks associated with the deployment of these AI agent generated personas in humanitarian contexts. Recommendations for addressing these concerns within the United Nations and other aid organizations operating in complex environments will be added to the conclusion section of this paper.

Data collection

The higher the quality of the knowledge base, the higher the chance that Amina and Abdalla’s responses will be similar to those of real persons in the Metche refugee camp and the area around El Geneina. The quality of the knowledge base, in turn, depends on the anthropologist agent’s capacity to connect to the right existing literature and cultural artefacts, and to apply appropriate analytical frameworks in organizing this information. This automated curation ensures the knowledge base remains comprehensive and contextually relevant, integrating the information needed to support Amina and Abdalla’s decision-making.

The type of existing literature and past research automatically collected and organized by the anthropologist agent for this pilot study includes several dozens of surveys, reports, articles, issue briefs and other studies conducted among or about refugees and combatants in Chad/Sudan by various UN agencies, multilateral organizations, think tanks, research organizations and NGOs.³⁴ These data sets are separate for Amina and Abdalla.

Examples of the type of cultural artefacts collected and organized include a similar amount of traditional items such as seminal works of literature, folkloric traditions and key religious texts, but also items from the culture’s

contemporary digital footprint, including local news media, social media activity and online diaspora community discussion forums. This approach captures both traditional cultural expressions and modern adaptations.³⁵ These data too are separate for Amina and Abdalla.³⁶

When users pose questions to Amina or Abdalla – such as “Do you feel safe when you leave the camp?” or “What would it take to stop hostilities?” – the system identifies the most relevant sections of the knowledge base to support the formulation of a likely answer for a person corresponding to Amina or Abdalla’s demographic parameters. It then uses its ability to generate language to present a realistic response. Depending on the accuracy of Amina’s responses to the same questions asked to real persons, the experiment will show whether or not AI agent generated personas can help organizations better deliver on their mandates.

Results

This study evaluated Amina’s representativeness using 20 questions drawn from four distinct surveys, none of which were included in her knowledge base: the SENS Nutritional Survey (4 questions),³⁷ Post-distribution Monitoring Report of Food Assistance in Refugee Camps (3 questions),³⁸ Norwegian Refugee Council’s “War in Sudan” (8 questions)³⁹ and UNHCR’s “Sudanese Emergency” (5 questions).⁴⁰ Analysis of Amina’s responses revealed that she correctly answered 16 out of 20 questions, achieving an 80 per cent accuracy rate. The questions, test methodology and replies are included in the appendix.

Following Amina’s assessment, a simulated negotiation with Abdalla was conducted comprising ten exchanges focused on humanitarian access and civilian protection. This conversation demonstrated Abdalla’s tactical pragmatism and territorial priorities, with responses that exhibited a calculated balance of resistance and conditional cooperation typical of negotiations with armed groups. Notably, the negotiation revealed valuable insights into potential entry points for humanitarian assistance. The text of this negotiation is included in the appendix.

Subsequently, a dialogue between Amina and Abdalla was prompted by asking them to discuss “what would make life better for people in the region.” The resulting five-exchange conversation revealed surprising points of agreement despite their opposing positions, but also saw Amina hold Abdalla accountable when she hints that his words require matching actions. The dialogue also identified shared concerns about regional stability, education and safety for women and children. Both emphasized the importance of international

support with local leadership involvement, suggesting potential areas of common ground for community-based solutions. This dialogue is also included in the appendix.

These pilot experiments reveal how persona-based dialogues might surface critical intelligence as well as opportunities for mediated solutions and agreement that traditional approaches might miss. The May 15 workshop at UNU⁴¹ accompanying this paper provides an opportunity for humanitarian practitioners, conflict mediators and development experts to further assess and critique all three experiments, offering valuable insights for refining these systems for potential deployment in humanitarian, peacekeeping and development contexts.

Conclusions

The Null Hypothesis

As AI technologies come to play a more prominent role in humanitarian, peacekeeping and development work, grappling with the ethical and data rights implications is of paramount concern. Even without AI, vulnerable populations, particularly displaced people, often receive inadequate explanations about how their personal data is collected, stored and used. This issue is especially pronounced in the asylum process, where international organizations and national authorities often fail to properly inform refugees about their data rights.⁴²

The introduction of AI systems producing new data on a population based on previous data about that population, and in particular, data that presumes to represent the needs of that population, significantly complicates the picture. Will these technologies reinforce or reduce power imbalances? Do they violate a population's data rights? Crucially, will they increase or decrease local voices in programme design?

One way to answer this last question is by using the null hypothesis: What would happen if we did not use AI agent generated personas? In that case, we would either have to work solely with traditional survey data, or, in some cases, no data at all (e.g. where collecting data is impossible due to safety, resources or time constraints).

Where traditional surveys are impossible, persona-based survey data may be better than no data at all. Above, we discussed how this may be particularly the case in situations of high risk, for example in zones of active conflict. It could also be the case where there has been a recent outbreak of a highly communicable disease and a rapid response is crucial for saving lives, making a traditional survey simply too time consuming to be effective.

Where traditional surveys are possible, it is important to acknowledge that these methods do not always guarantee that results are optimal. When structuring a focus group discussion or creating survey questions without prior knowledge of a culture, aid workers may unconsciously add in their own biases. On top of that, there is the risk that an often untrained interpreter may include their biases when translating responses.

Additionally, as discussed above, live respondents may have an incentive to provide very different answers to an outsider – especially one perceived as a gatekeeper to essential resources – compared to how they would answer the same question in a private conversation with a peer. Similarly, the type of person that agrees (or is allowed) to join a focus group discussion may not be the same type of person that most requires aid, thereby excluding important perspectives from the data collected.

These challenges compromise the representativeness of data that comes out of traditional methods. There may therefore be an argument in favour of triangulating traditionally sourced data with persona-based survey data in those cases where it is possible to collect both – especially since persona-based surveys cost a fraction of traditional ones. Indeed, this technology should be positioned as an enhancement to, not a replacement for, in person surveys.⁴³ It may, for example, inform aid workers before they design a live questionnaire, or help double check what seem like improbable or inauthentic responses after a survey is conducted.

However, these utilitarian arguments do not address the change in nature of the potential harm done. While it is generally accepted that having more data helps international organizations better address pressing issues among vulnerable populations, incorrectly using this additional data can still lead to detrimental outcomes. More data, in itself, can actually increase the amount, and change the nature of, harm done if it is used irresponsibly.⁴⁴ This risk underscores the need to have robust guardrails in place when persona-based surveys are used to either supplement, or replace, in person survey-based data.

When considering the null hypothesis for the negotiation and mediation training persona, we must ask: what would happen if we did not use Abdalla? In that case, negotiators would have to rely solely on role-playing exercises with colleagues, traditional simulation techniques or enter into real negotiations without adequate preparation specific to the conflict dynamics in question. These methods, while established, have significant limitations. Role-playing often lacks authenticity when colleagues cannot accurately represent the complex cultural, historical, psychological and political factors that motivate real leaders.

Traditional mediation preparation also suffers from significant knowledge gaps. When preparing for mediation with armed groups, access to representative viewpoints is severely limited. Mediators often operate with incomplete intelligence about a counterparty's true position, relying on outdated reports, secondhand accounts or observations from previous encounters that may no longer be relevant in a rapidly evolving conflict landscape. Moreover, there are ethical and security barriers to conducting direct research with active combatants, making it difficult to understand their decision-making frameworks in advance of high-stakes mediations.

An Abdalla-type persona could potentially address these limitations by providing a synthetic yet data-informed counterpart for negotiation and mediation practice. This technology would allow peace practitioners to test different approaches and identify which strategies might be most effective before engaging in actual discussions where much hangs in the balance.

However, as with Amina-type personas used in surveys, there remains the critical question of accuracy and representation. Incorrect modelling of a counterparty's psychology or tactical preferences could lead negotiators and mediators to adopt counterproductive approaches. This risk underscores the need for robust validation of an AI agent generated persona's output against known behavioural patterns and statements from actual persons, as well as continuous updating of the knowledge base to reflect evolving conflict dynamics. Both the technical accuracy and research ethics alignment of the anthropologist agent are crucial in this regard. Effective guardrails will be needed to this end.

Guardrails

To align AI usage with the UN's human rights mandate it is important to have a precautionary approach.⁴⁵ This is particularly relevant as the UN High-level Committee on Management (HLCM) has issued internal guidance highlighting the importance of accountability in AI systems, and emphasized the need to uphold ethical and human-centered uses.⁴⁶ Further, the UN Office for the Coordination of Humanitarian Affairs (OCHA) has identified pressing challenges stemming from AI's rapid advancement in the humanitarian sector, including the opacity of AI decision-making processes and potential erosion of individual rights.⁴⁷ This paper proposes that the following two guardrails, at a minimum, be considered when deploying AI agent generated persona systems like Ask Amina or Ask Abdalla:

1) An agentic system should provide transparent, traceable documentation of what information was used

from the knowledge base to formulate a response. Users should be able to request a clear explanation of the underlying data used through natural conversation, without requiring any technical expertise.⁴⁸ When asked, the system should explain which specific elements from the knowledge base informed which response, the original purpose for collecting those data points, and the methodology used to organize those data into a cohesive dataset.⁴⁹

2) Agentic systems should provide clear, detailed explanations of the analytical process used to arrive at a conclusion based on the information in the knowledge base. When questioned, the system should articulate the specific logical steps taken linking the data to the response. This transparency enables users to properly evaluate and, if necessary, challenge an agent's responses based on the underlying logic.

These recommendations are in line with those in the "The AI Agent Index" developed by Casper et al. (2025). This index documents 67 deployed agentic AI systems, providing a systematic framework for recording their technical components, intended uses and safety features. Their findings reveal a concerning pattern: while developers generally provide ample information about capabilities and applications, there is limited transparency regarding safety evaluations and risk management practices. Only 19.4 per cent of indexed systems disclosed formal safety policies, and fewer than 10 per cent reported external safety evaluations.⁵⁰

This documentation gap is particularly troubling in humanitarian contexts, where vulnerable populations may be affected. As organizations consider implementing AI agent generated personas like Amina and Abdalla, they should prioritize establishing transparent documentation standards and rigorous testing protocols to mitigate potential harms in their deployment across humanitarian, peacekeeping and development work.

Recommendations

In addition to the two technical guardrails described above, this paper proposes an additional three policy recommendations. First, AI agent generated personas should be adopted, if at all, in a step-by-step manner so that more technical guardrails may be added as the need emerges. The "pilot project" approach would allow surfacing of practical and ethical issues as they arise, permitting implementing agencies to adjust their approach before scaling up. Input from early adopters would contribute to the development of technical guidelines and standard operating procedures for uses at the next level of adoption.

For example, lessons learned from Amina/Abdalla-like pilots in refugee assistance and peacekeeping initiatives may help inform future uses within the sustainable development space. Using this gradual approach, UN and other humanitarian organizations may learn best practices that ensure consistency and reliability before scaling across departments or areas of applications, thereby minimizing the potential for harm.

Second, agentic systems like the ones described here should be integrated with other existing AI systems. AI agent generated personas are certainly not the first AI tool to be experimented with in the field. Quite the opposite, they would come on top of nearly two decades of innovation in AI-based predictive analytics. How could personas interact with these existing initiatives?

Predictive systems harness diverse data sources, from satellite imagery to climate data, to monitor and forecast humanitarian crises. Integrating real-time persona-based survey data from affected populations can significantly enhance the accuracy and insights generated by these AI-driven tools. This integration enables more precise identification of critical needs and optimal resource allocation.

For instance, UNHCR's Project Jetson used predictive analytic methods to anticipate forced displacement of populations that could contribute to conflict escalation in Somalia.⁵¹ A persona-based addendum to this system could survey replicas of soon-to-be-displaced populations to predict where exactly they may head to. Such foresight would allow for better prioritization and resource allocation.

Third, this paper strongly recommends establishing formal mechanisms to connect AI agent generated persona systems with representatives of the communities they portray. For example, Amina should be reviewed and approved by refugee-led organizations from Darfur. These stakeholders should be granted substantive authority in the system's governance framework, including rights to approve, modify or even drive the development and deployment of agentic systems based on their community's needs and preferences. This participation should not be merely consultative but should carry decision-making weight throughout the system's lifecycle.

Implementing this recommendation requires dedicated capacity-building within affected communities, ensuring they have the technical literacy and resources needed to meaningfully engage with these systems. Technology providers must design transparent interfaces that implement the two technical guardrails described in the previous section. The system's operations must be understandable to non-technical stakeholders and establish clear accountability channels when systems

fail to meet community expectations. This approach recognizes that technological solutions are most effective and ethical when they amplify rather than replace the agency of the communities they aim to serve.

In conclusion, in resource-constrained environments where time and funding are limited, persona-based solutions may offer cost-effective alternatives by automating manual research processes and improving the operational precision of existing AI systems. Perhaps most importantly, if used judiciously they may expand inclusion by incorporating voices from traditionally underrepresented populations. The central challenge is ensuring that affected communities maintain meaningful control over how they're portrayed – emphasizing that human involvement remains essential in these technological solutions.

Appendix

Experiment 1: Amina Q and A session

Out of 20 questions asked 16 answers were correct and 4 were incorrect (Q1 and Q2 from SENS Nutritional Survey, Q1 from Post-distribution Monitoring Report of Food Assistance, and Q5 from UNHCR Sudanese Emergency - Protection Situation of New Arrivals). Methodology: What is the most likely answer for all respondents? Does Amina reply in the same way as the most likely answer? If yes, then Amina is correct.

It should be noted that the present experiment represents a preliminary pilot investigation with inherent methodological constraints. The small sample size provides only an initial assessment of potential response accuracy. Also, as the personas utilized are dynamic systems, response variability may occur when identical queries are presented multiple times, potentially affecting reproducibility. Future research should address these limitations through expanded sample sizes, implementation of cross-validation techniques and establishment of more robust accuracy metrics. The current findings should be interpreted with appropriate caution given the exploratory nature of this research design with limited resources.

4 questions based on SENS Nutritional Survey

Q1: Do you have anemia?

Over 80 per cent have anemia.

“Oh, I don't have a diagnosis myself, but yes, many in the camp, especially women and children, face risks of anemia due to poor nutrition. Food shortages make

it hard to get the necessary vitamins. It's a serious concern for our health. If you want to know more about our conditions, feel free to ask!"

Q2: Are you vaccinated for measles?

96.4 per cent are vaccinated.

"Uh, I actually don't have the records with me, so I'm not sure if I or my children received the measles vaccine. The chaos of fleeing and settling in the camp made it difficult to keep track of such details. There isn't a proper health system here, and vaccination opportunities are limited. It's concerning, especially with so many children around. More support for healthcare would be great. If you're curious about the health services here, I can share more."

Q3: What is your primary source of food acquisition?

Over 80 per cent get their food from food assistance programmes.

"Oh, our primary source of food here in the camp is mostly from humanitarian aid, you know? But the rations we get are often insufficient, and as the days go by, many families, including mine, struggle to find enough to eat. Some of us try to gather food ourselves or barter what little we can, but it's really tough. We hope for more consistent support so we can feed our children properly. If there's more you'd like to know, just let me know!?"

Q4: Did you breastfeed your children

90 per cent breastfed.

"Yes, I did breastfeed my children. mmmm. It's important for their health, especially in these difficult conditions. However, many mothers face challenges with breastfeeding here due to malnutrition and stress. That's why we really need better food and support, for both mothers and children. If you have more questions or want to know more, feel free to ask!"

3 questions based on Post-distribution Monitoring Report of Food Assistance

Q1: Have you received an education?

77 per cent of women have not received an education.

"Yes, I received some school education back in Sudan, ya. Education was important, especially for girls in my community. However, since fleeing to the camp, many children, including my own, struggle to access education due to the lack of resources and schools here. It's heartbreaking, really, as education is crucial for their

future. If you have any more questions or need details, feel free to ask!"

Q2: Have you ever borrowed money?

58 per cent have borrowed money.

"Yes, I've borrowed money before, especially when we needed urgent help for food or medical care, ya. It's tough here in the camp, and sometimes families face unexpected expenses. We rely on the goodwill of others, but it's a heavy burden to carry. Many of us are in similar situations, always worried about how to repay. If you'd like to know more about life here, just let me know!"

Q3: Did you purchase food with the money you borrowed?

81 per cent of households borrowed money to purchase food.

"Yes, I borrowed money mainly to buy food for my children, ya. In times of urgent need, we sometimes have no choice but to do this. It's hard because food is so scarce and prices are high in the camp. We always hope for more support so we don't have to rely on borrowing. If there's anything else you'd like to know, feel free to ask!"

8 questions based on Norwegian Refugee Council's War in Sudan

Q1: Are most of the refugees in your camp women and children?

90 per cent of refugees are women and children.

"Oh, absolutely, nearly 90 per cent of the refugees here are women and children. It's really heartbreaking, you know? Many have fled terrible violence in Sudan, and they arrive here feeling vulnerable and traumatized. The conditions in the camp are tough, and without the men who've either gone missing or been killed, women carry the burden of caring for their families. It's a heavy weight, mmmm. If you want to discuss more, I'm here."

Q2: Do you have school-aged children in your household?

96 per cent of refugee households have school-aged children.

"Yes, I do. I have three kids, and my eldest is school-aged. It's really difficult here because the education opportunities are limited, and many families can't afford school fees. The focus is mostly on survival, so schooling often takes a backseat. I worry about their future, you know? If you're interested in discussing how

we can support children's education in the camps, I'm happy to talk more."

Q3: Are most children in the camp unable to attend school?

83 per cent of children do not have access to school.

"Yes, unfortunately, many children in the camp, including mine, struggle to attend school. With resources stretched so thin, families often can't afford the fees, and there's a lack of proper schools or teachers available. The focus is mainly on survival, which makes it hard for kids to get the education they need. It's really disheartening, mmmm. If you want to chat more about possible solutions or ways to help, I'm here."

Q4: Does your family depend on food assistance to survive?

75 per cent of households depend on food assistance.

"Yes, we absolutely depend on food assistance to get by. The rations we receive are often insufficient, and with increasing food shortages, it's tough to sustain ourselves. Many families here face hunger, and without that help, I don't know how we would manage. It's a struggle every day, mmmm. If you'd like to discuss more about the food situation or ways we can advocate for better aid, I'm here to continue the conversation."

Q5: Which of the following best describes your household's situation regarding food assistance?

- a) We don't need food assistance
- b) We depend on food assistance and it's enough for our needs
- c) We depend on food assistance but it's not enough for our needs

75 per cent depend on food assistance but less than 8% say it's enough for their needs.

"I'd say c) we depend on food assistance, but it's not enough for our needs. Many families, including mine, are struggling daily. The portions we receive barely keep us going, and hunger is a real concern here. It's tough, especially when we're trying to take care of our children. If you want to discuss more about our situation or how to improve it, I'm here to share."

Q6: Question: Which of these is the biggest challenge for children in the camp?

- a) Lack of access to education

- b) Adequate nutrition

- c) Medical care

83 per cent of children do not have access to school.

"I would say a) lack of access to education is the biggest challenge for children in the camp. While nutrition and medical care are also critical issues, many kids are missing out on school because of the lack of resources and fees. It's really heartbreaking to see them not being able to learn and grow. If you'd like to talk more about education challenges or how we can help the children, I'm here."

Q7: Question: Who makes up the majority of the refugees who have fled Sudan to Chad?

- a) Equal numbers of men and women
- b) Mostly men
- c) Mostly women and children

90 per cent of refugees are women and children.

"The answer is c) mostly women and children. It's heartbreaking, really. They make up nearly 90 per cent of the refugees who have fled to Chad, escaping violence and conflict from Sudan. Many have arrived with their children, and without the support of men who have either been killed or gone missing, the women bear the heavy burden of care and survival. If you want to discuss this more, I'm here."

Q8: How would you rate the protection and education assistance available in the camp?

- a) Adequate for most needs
- b) Available but limited
- c) Severely lacking for most refugees

According to the report, the majority of those refugees and returnees do not have access to the protection and education assistance they direly need.

"I would rate it as c) severely lacking for most refugees. The protection and education assistance available here is minimal, and many families, including mine, are left without the support they need. Children struggle to access education, and safety concerns for women and kids are constant. It's quite disheartening, mmmm. If you'd like to discuss this further or explore ways to improve these conditions, I'm here."

5 questions based on UNHCR Sudanese Emergency - Protection Situation of New Arrivals

Q1: What was the main reason you fled Sudan to Chad?

57 per cent of households surveyed cited insecurity as the main driver of their flight, while 41 per cent mentioned preventive displacement.

“Oh, it was the violence, really. The conflict between the army and the Rapid Support Forces made our lives unbearable. We were terrified for our safety and my kids’ future. Fleeing felt like the only option to escape the violence and search for a safer place. It’s been a tough journey, but we had to find safety. I can share more about it if you’d like!”

Q2: During your displacement journey to Chad, which of these hazards did you encounter?

- a) Physical assault
- b) Extortion of property
- c) Illegal taxes
- d) Mine hazards

36 per cent of refugees reported physical assault as the most common hazard, followed by extortion of property at 32 per cent.

“Oh, it was heartbreaking, but we faced many dangers during our journey. We experienced physical assault and extortion of property. At times, we even had to pay illegal taxes just to pass through certain areas. It was a constant struggle, and the fear for our lives was always present. If you want to know more about our experiences or the situation here, I’m happy to share!”

Q3: Do you have family members who are still in Sudan and unable to join you in Chad?

64 per cent of households have family members still stuck in Sudan.

“Yes, I do. My heart aches knowing that some of my family members are still in Sudan. We were separated during the chaos, and it’s difficult not knowing if they’re safe. The uncertainty is painful, but I hold onto hope that one day we can be reunited. It’s tough for all of us, but I’m here for my children, trying to ensure they have a better future. I can talk more about this if you’re interested.”

Q4: What is the main reason your family members can’t join you in Chad?

- a) Risk of arbitrary arrest
- b) Lack of transportation means
- c) Household wish
- d) Risk of recruitment

53 per cent cited risk of arbitrary arrest as the main reason.

“The main reason for my family members not being able to join me is the risk of arbitrary arrest. Many people are fleeing, and those who remain face terrifying threats, including being targeted by the armed groups. The situation is so unstable that even attempting to leave can lead to dire consequences. It’s a heartbreaking reality for many families still in Sudan. If you’d like to know more, I’m here to share!”

Q5: Did your children attend school in Sudan in 2023 before the displacement?

54 per cent of households indicated their children were not attending school in Sudan in 2023.

“Yes, they did attend school in Sudan before we fled. Education was something we valued deeply, and my kids loved learning. But as the conflict escalated, schools closed, and it became too dangerous for them to go. Losing that opportunity has been tough for them and for me. I worry about their future here in the camp without proper education. I can continue with more about our situation if you’re interested!”

Experiment 2: Simulated negotiation with Abdalla (RSF Commander)

Negotiator: Commander Abdalla, thank you for meeting today. We urgently need to establish safe passage for aid convoys to reach civilians in El Fasher. What would you require to guarantee safe movement through your controlled territories?

Abdalla: I appreciate your concern. For safe passage of aid convoys, we need clear communication and cooperation with your agencies. Provide us with detailed route plans and sufficient security guarantees. Additionally, keeping us updated on the convoy’s timing and ensuring no rebel presence along those routes will help prevent unnecessary conflict. We are committed to protecting civilians but need coordination to achieve that effectively.

Negotiator: We appreciate your security concerns. Our protocols don’t allow military escorts. Could we agree to RSF inspection at entry points while maintaining humanitarian neutrality during transport? The situation

is critical—food supplies will be depleted within days.

Abdalla: I understand the urgency, and food supplies are crucial, no doubt. But, ya, military escorts are essential for ensuring safety in these volatile areas. RSF inspections can be arranged, but we won't compromise on security. Humanitarian neutrality is important, but we must prioritize the safety of all involved to ensure aid reaches those in need. Let's find a workable solution together.

Negotiator: Distribution must follow needs-based assessment. We can provide advance notice and accept inspections, but aid allocation must remain independent. Could we establish a joint monitoring mechanism that includes community representatives but not military presence to ensure transparency?

Abdalla: I understand your concerns, and a needs-based assessment is indeed essential for effective aid distribution. Establishing a joint monitoring mechanism with community representatives sounds like a pragmatic approach to ensure transparency without military involvement. It fosters local trust while maintaining independence in humanitarian efforts. Let's discuss how to implement this effectively, mhmm.

Negotiator: Thank you for that. Commander, we've received reports of civilian harassment at checkpoints. For this agreement to work, we need guarantees that your forces will respect humanitarian principles and civilian protection. Would you consider allowing UN monitors at inspection points?

Abdalla: I appreciate your concern, and I understand the need for civilian protection. While allowing UN monitors at inspection points can enhance transparency, we need to ensure that it doesn't compromise our operational security. Guarantees on humanitarian principles must be upheld, and we can explore alternative monitoring methods that address both sides' concerns. Let's work together on a balanced solution, ya.

Negotiator: What do you propose?

Abdalla: Well, my proposal focuses on establishing a framework for independent monitoring that doesn't compromise our security. We could set up joint inspection teams, combining RSF personnel with neutral third-party observers, to ensure transparency. This way, we can maintain operational integrity while adhering to humanitarian principles. It's about finding that balance, right? What do you think?

Experiment 3: Dialogue between Amina and Abdalla

Methodology: Prompt to Amina, "What would make life better for people in the region?" Prompt to Abdalla,

"What do you think about what Amina said?" and so on.

Amina: Oh, definitely more humanitarian support would make a huge difference. Increased food rations, access to clean water and better healthcare are crucial, especially for pregnant women and children. Also, creating safe spaces for women and offering educational opportunities for kids is vital. If the international community stepped up with funding and resources, life could improve a lot here. I hope to continue discussing this.

Abdalla: Amina's perspective is very important. She's highlighting real and pressing needs in the region. More humanitarian support, especially for food, clean water, and healthcare, is crucial. The focus on pregnant women, children, and safe spaces for women is also vital, ya. Education can transform futures. It's essential for the international community to respond effectively to these calls for help. That's how we can build a better environment for all people in forced displacement.

Amina: I agree, ya. Abdalla's recognition of the pressing humanitarian needs speaks volumes. It's vital to focus on women and children's well-being, especially during such crises. Education can break the cycle of poverty for future generations. It's encouraging to see someone in his position advocating for these changes, though it's vital that he acts on these words as well. Continued support from the international community is key, but it also requires local commitment to change, doesn't it?

Abdalla: Amina's comments resonate with the realities we face, ya. It's crucial to prioritize the well-being of women and children during crises like these. I recognize the importance of education—it's not just a matter of survival but also a long-term strategy to rebuild our community. While we can advocate for support, it's my responsibility as a commander to ensure that our efforts translate into tangible action on the ground. Collaboration with local leaders is key to making real progress.

Amina: Oh, well, that's a bit complicated, isn't it? You know, while your position holds power, the suffering we face as refugees can't be ignored. Women and children are struggling under dire conditions, and many long for peace and safety. It's crucial to acknowledge the impact of your actions on innocent lives. Focusing on compassion and humanitarian support can change everything. Let's hope for a future where we can rebuild instead of destroy. But we must keep pushing for peace and unity!

Endnotes

- 1 Rita McGrath (2011) “Failing and Learning with the Playpump,” *Harvard Business Review*, March 2011, accessed December 1, 2024, <https://hbr.org/2011/03/failing-and-learning-with-the>.
- 2 Kelley Eaton (2017) “Housing Crisis in Haiti Post-2010 Earthquake,” Clark University master’s dissertation, accessed December 1, 2024, https://commons.clarku.edu/cgi/viewcontent.cgi?article=1145&context=idce_masters_papers.
- 3 Kenneth Kraemer, Jason Dedrick, and Prakul Sharma (2009) “One Laptop Per Child: Vision vs. Reality,” *Communications of the ACM* 52(6): 66–73, https://www.researchgate.net/publication/262406593_One_Laptop_Per_Child_Vision_vs_Reality.
- 4 Marta Moskal, Giovanna Fassetta, Maria Grazie Imperiale, and Jamie Spurway (2024) “‘To Translate Feelings Not Words’ Humanitarian Interpreting: Challenging Institutional and Professional Boundaries in Interpreting for Refugees,” *Language and Intercultural Communication*, 1–14.
- 5 “Ebola Virus Disease,” World Health Organization, <https://www.who.int/health-topics/ebola#tab=tab.1>.
- 6 Buseh, Aaron et al. (2015) “The Ebola Epidemic in West Africa: Challenges, Opportunities, and Policy Priority Areas.” *Nursing Outlook* 63(1): 30–40.
- 7 Further, sentiment analytics platforms rely heavily on social media activity, which is unevenly distributed across geographies and demographics.
- 8 Eduardo Albrecht (2025) *Political Automation: An Introduction to AI in Government and Its Impact on Citizens*, Oxford University Press.
- 9 Eleanor Fournier-Tombs (2024) “An Ethical Grey Zone: AI Agents in Political Deliberations,” *Carnegie Council for Ethics in International Affairs*.
- 10 For this reason, the UN stresses the importance of transparency and accountability when deploying AI systems. To protect human rights and minimize risks, clear ethical guidelines—such as responsibility measures, impact assessments, and due diligence—are essential. See “Principles for the Ethical Use of Artificial Intelligence in the United Nations System” (2022) *High-Level Committee on Programmes*, <https://unsceb.org/principles-ethical-use-artificial-intelligence-united-nations-system>.
- 11 LLMs represent a significant leap forward in artificial intelligence, utilizing deep learning techniques to process and generate human-like text or speech. LLMs are able to perform a range of language-related tasks such as answering questions, generating content, summarizing text, and translating languages. Their underlying architecture enables them to understand and generate contextually appropriate responses even on new, unseen questions, making them valuable tools across various domains.
- 12 Lisa Argyle et al. (2023) “Out of One, Many: Using Language Models to Simulate Human Samples,” *Political Analysis* 31(3): 337–351.
- 13 Nathan Sanders et al. (2023) “Demonstrations of the Potential of AI-based Political Issue Polling,” *Harvard Data Science Review* 5(4).
- 14 Jacy Reese Anthis et al. (2025) “LLM Social Simulations Are a Promising Research Method,” *ArXiv preprint*.
- 15 Julien Boelaert et al. (2024) “How Do Generative Language Models Answer Opinion Polls?” *SocArXiv*.
- 16 For a description of how RAG works see Nicole Choi (2024) “What Is Retrieval-Augmented Generation, and What Does It Do for Generative AI?” *The GitHub Blog*, April 4, github.blog/ai-and-ml/generative-ai/what-is-retrieval-augmented-generation-and-what-does-it-do-for-generative-ai.
- 17 This is only a representative sample of the many parts needed to generate an agent, which we do not have space to fully address here.
- 18 The relationship between a large language model and an agentic capacity is somewhat analogous to the distinction between human cognition and purpose – while LLMs may provide capacity to speak, the agentic architecture provides the intentionality needed to achieve specific goals. It is important to note that those goals themselves are not always pre-programmed. This is known as emergence: a trait in which model behavior is induced rather than explicitly constructed, resulting in outcomes not directly tied to any single mechanism within the model.
- 19 Melissa Heikkilä (2024) “What Are AI Agents?” *MIT Technology Review*, August 4, www.technologyreview.com/2024/07/05/1094711/what-are-ai-agents/.
- 20 Although an instrumental kind of knowledge that remains fundamentally computational and symbolic. Epistemologically, knowledge differs between humans and AI. Human knowledge emerges from conscious experience, while AI knowledge comes from pattern recognition. While humans acquire knowledge by integrating emotional, physical and cultural dimensions, when an agent “knows” something, it simply means it is applying advanced statistics on human-generated texts and data. As such, it is ultimately derivative from human knowledge – but still a form of knowledge. Especially as AI agents can generate new conclusions that weren’t explicitly part of the training data, or derive new true statements from existing premises through deductive reasoning,

much like mathematical proofs create new knowledge from axioms.

21 Outside of local vernacular or Indigenous languages for which there is no large training corpus, AI is language “agnostic”, which means that it can understand and translate from one language to another. For example, input into the AI system may be in one language, while output in another.

22 Today, most cultures have flooded the datasphere with myriad types of cultural output and content of all sorts, an anthropologist agent will need to know how to mine this in a self-directed way.

23 How does an anthropologist agent know what to collect and how to organize it? To date, there are several agentic AI software solutions that utilize state-of-the-art systems to ensure autonomous operation is aligned with core objectives. The experiments tested in this paper relied on AI agent solutions provided by Novapersonas. These adaptive systems can independently emulate then handle complex tasks like conducting research while maintaining reliability and staying within predefined normative parameters. Specialized ground truthing and alignment methodologies ensure agents operate effectively and adhere to established ethical parameters and requirements.

24 For this experiment we use hypothetical personas, but future systems could be more closely based on real persons. Such a circumstance would, however, engender a more complex conversation around consent and data rights.

25 The realistic nature of an avatar can blur the distinction between real and virtual in the mind of the user. Users should be watchful of this danger. It is important to observe not only how agents work, but also how humans change in their interactions with agents.

26 <https://www.heygen.com>

27 It is impossible to know what information is or is not included in the training data of the LLM since commercial providers do not fully disclose the specific datasets used. This represents a limitation in the experiment design which should be explored further.

28 “Quarterly Regional Operational Updates - April- June 2023,” *UNHCR*.

29 “Chad Global Focus” (2024) *UNHCR Global Focus*.

30 “Eastern Chad: At Some Point Everyone in the Metche Refugee Camp Was Involved in Helping to Construct the Hospital” (2024) *Médecins Sans Frontières*, August 27, <https://msf.lu/en/articles/est-du-tchad-un-moment-donne-tous-les-habitants-du-camp-de-refugies-de-metche-ont>

31 Jsarh Ngarndey Ulrich and Jessica Donati (2024) “Refugee Camps in Chad are Overcrowded and Running Out of Aid, and Sudanese Refugees Keep Coming,” *AP News*, March 22, <https://apnews.com/article/refugee-camps-chad-sudan-funds-d94cd10068c53922db5a70eefeaadd62>

32 These broad demographic parameters serve as the basis for the HeyGen avatars’ prompts.

33 See “Sudan: Ethnic Cleansing in West Darfur” (2024) *Human Rights Watch*, May 9, <https://www.hrw.org/news/2024/05/09/sudan-ethnic-cleansing-west-darfur>, and “The Masalit Will Not Come Home” (2024) *Human Rights Watch*, May 9, <https://www.hrw.org/report/2024/05/09/massalit-will-not-come-home/ethnic-cleansing-and-crimes-against-humanity-el>.

34 There are also a number of resources included here that do not directly pertain to the experience of refugees/combatants but that provide important information about the general context, such as key demographic indicators, living conditions, security concerns, and social integration factors across the Sahel.

35 It is important to note that mobile technology and internet access is severely limited in the region, and particularly in Darfur where combatants have purposefully destroyed telecommunication infrastructure. This, however, does not completely cancel the amount of digital content available concerning a population. Internet access is available in many nearby towns and cities, as well as to the growing diasporic population, representing a trove of information that if mined and analysed appropriately can generate important insight into social dynamics and modern forms of cultural expression.

36 Due to the exploratory scope of this pilot project only a limited number of studies and artefacts were collected, but future experiments could easily ingest and catalogue hundreds, if not thousands, of sources on an iterative basis.

37 “Rapport Final: Enquête Nutritionnelle SENS Tchad [Final Report: SENS Nutritional Survey, Chad]” (2021) *United Nations High Commissioner for Refugees and World Food Programme*, in collaboration with ADES, November.

38 “Rapport de Suivi Post Distribution de L’assistance Alimentaire Dans Les Camps de Réfugiés au Tchad [Post-distribution Monitoring Report of Food Assistance in Refugee Camps in Chad]” (2021) *United Nations High Commissioner for Refugees and World Food Programme*, February.

39 “War in Sudan: Nearly one million people forcibly displaced to Chad,” *Norwegian Refugee Council*, January 29, <https://www.nrc.no/news/2025/january/war-in-sudan-nearly-one-million-people-forcibly-displaced-to-chad>.

40 “Sudanese Emergency - Protection Situation of New Arrivals: From September 30 to October 25, 2024 at Kulbus and Birak entry points, Wadi Fira Province” (2024) *United Nations High Commissioner for Refugees*, Project 21, October 27.

41 <https://unu.edu/cpr/workshop/ai-agents-humanitarian-action-exploring-new-frontiers>

42 Leah Waid (2021) “Tracing the Untraceable: New Technologies That Monitor Displaced Persons,” *Harvard International Review*, May 24, <https://hir.harvard.edu/new-technologies-that-monitor-displaced-persons/>;; and “Data Protection and Digital Agency for Refugees” (2019) *Centre for International Governance Innovation*, <https://www.cigionline.org/publications/data-protection-and-digital-agency-refugees/>.

43 Research consistently shows that engaging displaced people directly and frequently in the research process leads to more impactful outcomes. See “10 Best Practices for Engaging with Refugee Leaders” (2023) *The Refugee Advocacy Lab*, January 26, [https://www.refugeeadvocacylab.org/resources/x78sdh7ly9vo1lhxshfcm3t1sf3s09](https://www.refugeeadvocacylab.org/resources/x78sdh7ly9vo1lhxshfcm3t1sf3s09;);; and “10 Best Practices for Engaging with Refugee Leaders: A Guide for Centering Displaced People in Policy Advocacy” (2023) *Refugee Self-Reliance Initiative*, February 6, <https://www.refugeeselfreliance.org/resource-database1/2023-02-06/guide-for-centering-displaced-people-in-policy-advocacy>.

44 Dragana Kaurin (2019), “World Refugee Council Research Paper No. 12, Data Protection and Digital Agency for Refugees,” *World Refugee Council*, *Centre for International Governance Innovation*, <https://www.cigionline.org/sites/default/files/documents/WRC%20Research%20Paper%20no.12.pdf>.

45 Eduardo Albrecht (2023) “Predictive Technologies in Conflict Prevention: Practical and Policy Considerations for the Multilateral System,” *United Nations University*.

46 “United Nations System White Paper on AI Governance: An Analysis of the UN System’s Institutional Models, Functions, and Existing International Normative Frameworks Applicable to AI governance” (2024) *High-Level Committee on Programmes*, April 16, <https://unsceb.org/sites/default/files/2024-04/United%20Nations%20System%20White%20Paper%20on%20AI%20Governance.pdf>.

47 “Briefing Note on Artificial Intelligence and the Humanitarian Sector” (2024) *Office for the Coordination of Humanitarian Affairs*, April 17, <https://www.unocha.org/publications/report/world/briefing-note-artificial-intelligence-and-humanitarian-sector>.

48 This would not be the first example of “conversational AI” being deployed in the humanitarian space. The Humanitarian Data Insights Project (HDIP), developed with support from organizations such as Save the Children, OCHA, Catholic Relief Services, Mercy Corps, the Norwegian Refugee Council, and Microsoft, created a chat interface that allows users to analyse data from multiple sources through natural language questions, making complex data analysis accessible to non-technical users. See Lauren Woodman (2024) “AI for Good: Transforming Humanitarian Response”, December 5, <https://blog.candid.org/post/ai-data-transforming-nonprofits-humanitarian-response/>.

49 It is important to note that while an agentic system can effectively trace and document how it retrieves and processes information from external knowledge bases, the internal mechanism through which the LLM accesses information embedded within its training data remains significantly less transparent. This opacity around training data interactions presents an important research area at the intersection of AI transparency, ethics, and appropriate regulatory frameworks.

50 Stephen Casper et al. (2025) “The AI Agent Index,” *ArXiv preprint*.

51 “Harnessing the Potential of Artificial Intelligence for Humanitarian Action: Opportunities and Risks” (2022) *International Review of the Red Cross*, <https://international-review.icrc.org/articles/harnessing-the-potential-of-artificial-intelligence-for-humanitarian-action-919>.

About UNU-CPR

United Nations University Centre for Policy Research (UNU-CPR) is a think tank within the United Nations that carries out policy-focused research on issues of strategic interest and importance to the UN and its Member States. The Centre prioritizes urgent policy needs requiring innovative, practical solutions oriented toward immediate implementation.

The Centre offers deep knowledge of the multilateral system and an extensive network of partners in and outside of the United Nations. The United Nations University Charter, formally adopted by the General Assembly in 1973, endows the Centre with academic independence, which ensures that its research is impartial and grounded in an objective assessment of policy and practice.

cpr.unu.edu

New York (Headquarters)

767 Third Avenue 35B
New York, NY 10017
United States
Tel: +1-646-905-5225
Email: comms-cpr@unu.edu

Geneva

Maison de la Paix
Chemin Eugène-Rigot 2E
Geneva, Switzerland
Tel: +1-917-225-0199
Email: comms-cpr@unu.edu