

THE SENTINEL PROJECT FOR GENOCIDE PREVENTION:

Countering Misinformation In South Sudan and
the Democratic Republic of the Congo



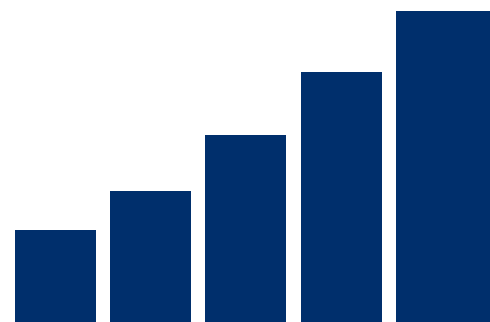
Table of Contents

1. Introduction	3
2. The Sentinel Project	5
3. Impact	8
Growth & Scale	12
Challenges and Opportunities	15
Lessons Learned	17
The Future	19
4. Conclusion	21

Creating Hope in Conflict: A Humanitarian Grand Challenge (CHIC) is a partnership of the U.K. Foreign, Commonwealth & Development Office and Global Affairs Canada, and is implemented by Grand Challenges Canada. CHIC identifies and accelerates innovations that save and improve the lives of those living in humanitarian crises caused by conflict.

This study was authored by Anahi Ayala Iacucci in collaboration with the CHIC team (Zainah Alsamman, Nabeel Chudasama and Nabeela Jivraj and) with the support of The Sentinel Project team.

Photos courtesy of the Sentinel Project team.





INTRODUCTION

As dawn broke over Yambio County, South Sudan, Achan—a local health worker—received a chilling message: “Town X is under attack by armed youth. Flee now!” Within minutes, panic swept through nearby villages. Families hid, markets shuttered, and humanitarian services stalled. But when Achan reached Town X, he found no violence—only routine morning bustle. The warning had been a false alarm, a rumour that spiralled into chaos.

Stories like this are not rare in conflict zones like parts of South Sudan and the Democratic Republic of Congo. In places where access to accurate news is limited, whispers spread like wildfire. Misinformation, disinformation, and hate speech—collectively known as MDH—can tear apart communities and obstruct critical aid. In fragile environments, where trust is scarce, MDH can turn fear into violence and falsehoods into fatal decisions.

Across global hotspots—from Ethiopia to Ukraine—rumours travel through the same channels that humanitarian workers use to deliver lifesaving messages. MDH can derail aid convoys, reignite ethnic tensions, or spread conspiracy theories during disease outbreaks. In response, agencies are stepping up their efforts. The Red Cross issues digital risk alerts, CDAC partners with local media on fact-checking, and the WHO promotes ethical guidance on “infodemic management.”

To further develop this response, Grand Challenges Canada commissioned a [landmark Scoping Study](#) on Misinformation, Disinformation and Hate Speech in 2024[1]. This report maps current interventions, identifies critical gaps, and offers tools for donors and implementers to build community-based strategies for countering harmful narratives before they escalate.

Since 2020, the Creating Hope in Conflict initiative[2] – GCC’s flagship humanitarian innovation portfolio[3] – has supported a range of innovations designed to stop MDH at its roots. One such idea came from the Sentinel Project[4], a small non-profit organization based in Toronto. Between 2020 and 2022, they were funded by GCC to test the bold belief that local knowledge combined with smart technology could interrupt the spread of false information and keep communities safe.

[1] Anahi Ayala Iacucci, [Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict](#), Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

[2] Find additional information [here](#).

[3] The UK Foreign, Commonwealth & Development Office and Global Affairs Canada, with support from Grand Challenges Canada, partner on the “Creating Hope in Conflict: A Humanitarian Grand Challenge”. Through this Grand Challenge, they identify and support groundbreaking solutions that engage the private sector and draw from the experiences of affected communities in order to significantly improve – and in many cases, save – the lives of vulnerable people affected by conflict.

The project, called “Countering Misinformation In South Sudan and the Democratic Republic of the Congo”, used the WikiRumours platform to crowdsource rumour reports and then fact-checks them with support from on-the-ground, trained moderators. By serving as a source of verified facts, the platform wanted to provide accurate, unbiased information to community members, humanitarian actors, and local authorities to improve knowledge of human rights violations and inform appropriate decisions and actions to counter major threats[5].

The objectives established for the Sentinel Project under its Grand Challenges Canada grant were threefold: (i) to enrol 5,000 new subscribers across the DRC and South Sudan, (ii) to refine the platform’s technology and operational model to enable faster replication in other humanitarian crisis zones, and (iii) to demonstrate the platform’s value to donors and private sector partners through a robust monitoring and evaluation (M&E) framework.

As of the close of the 2020–2022 funding period, the project was evaluated as successfully completed, with over 27,000 subscribers reached and key software updates deployed. Work on the third objective is ongoing, with early evidence demonstrating increasing donor interest and community-level impact.



[4] Find more information [here](#).

[5] Grand Challenges Canada, [The Sentinel Project for Genocide Prevention](#), August 18, 2020

A NOTE ON METHODOLOGY

This case study draws on multiple sources, including internal progress and results reports from the Sentinel Project team, a third-party site visit report, a key informant interview with project leader James Kataliko, and supplementary feedback and information provided by executive director Christopher Tuckwood. Additional insights were drawn from a desk review of published blog posts, surveys, and partner communications.

THE SENTINEL PROJECT

Back in Yambio, that belief took shape as WikiRumours[6]—a simple, open-source platform where anyone could report suspicious stories via SMS, WhatsApp, and other trusted communication channels. Local moderators would vet each report, verifying it using trusted sources. Within 24 hours, verified information—either reassurance or warning—was broadcast back through the same networks that once fuelled panic.

A 2025 survey[7] in South Sudan involving over 1,800 respondents revealed that 86% believed that misinformation was a significant problem. Nearly 41.4% admitted to spreading unverified conflict-related rumours and 49.1% felt ill-equipped to verify information. These findings illustrate both the demand and the gaps in trust that WikiRumours is designed to address.

This is where the Sentinel Project's story takes root: at the crossroads of fear and resilience. Here, communities do not surrender to lies—they reclaim the narrative. With Grand Challenges Canada's support beginning in 2020, the Sentinel Project began scaling their prototype into a robust, locally grounded operation focusing on three pillars: people, technology, and community.

Rather than deploy outside experts, the Sentinel Project trained around 400 community ambassadors in South Sudan and the DRC. Equipped with basic phones, solar chargers, and print guides, these ambassadors collected tips and relayed updates using toll-free numbers. Two key software updates improved verification workflows and introduced early-warning analytics. Local leaders and authorities—

[6] See [here](#).

[7] Sentinel Project, [Survey Findings: Critical Insights on Communication, Trust, and Misinformation in South Sudan](#), Jun 3, 2025

deeply trusted by their communities—were trained as system users and information sources, thus reinforcing credibility.

“

In the DRC, the team developed a community anticipation mechanism: as soon as they identified a potential rumour source (for example, an alarming post about elections or a claim that a medicine would “prevent” people from voting), they immediately prepared and sent a preventive message to the community. They used both SMS groups and social media to dispatch these verified information “postcards.”

Project Leader James Kataliko - DRC

”

As the subscriber base grew, the team engaged radio stations, NGOs, and media partners to amplify verified content. In total, it is estimated that WikiRumours reached more than 2 million people between 2020 and 2022[8]. By 2025, it was active across three provinces in eastern DRC with over 28,000 subscribers and more than 200 trained moderators and approximately 64,000 subscribers across six project areas in South Sudan.

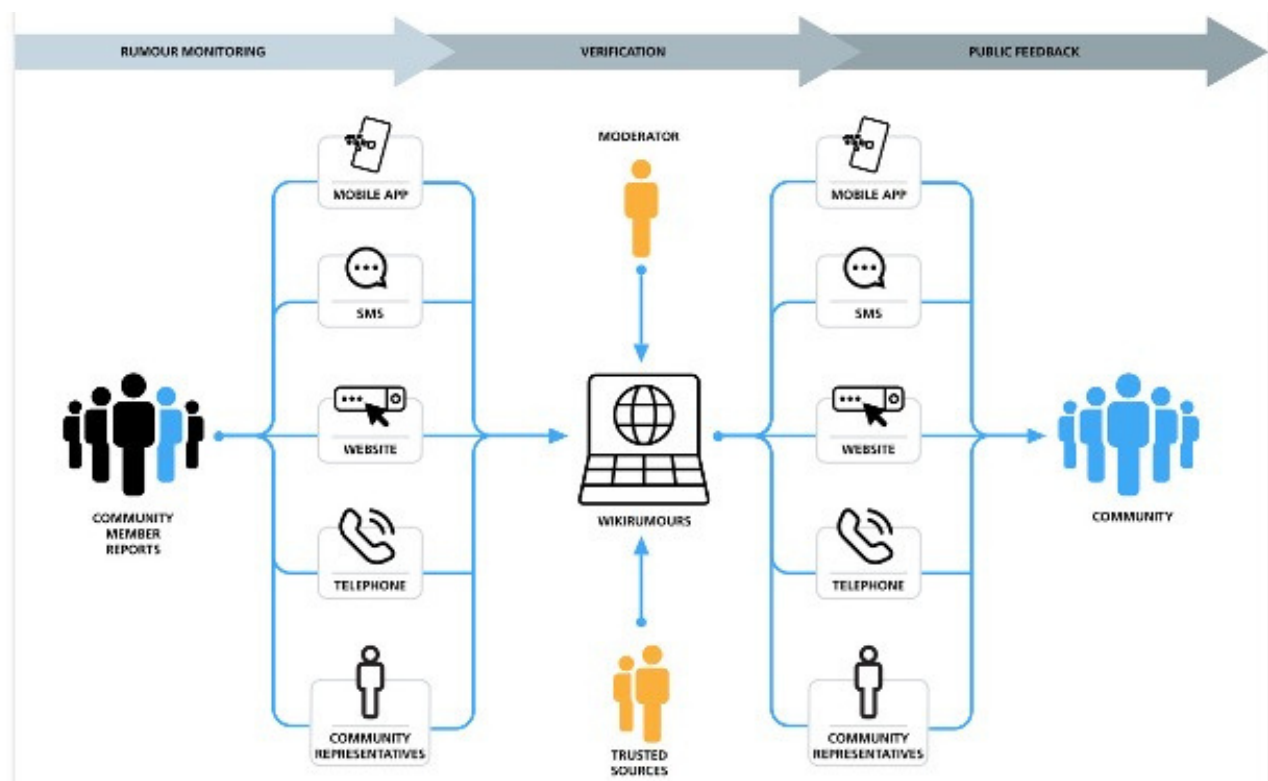
User feedback reported by a Sentinel Project community ambassador and project lead in the DRC demonstrated real impact: users felt safer, more informed, and less prone to rumour-driven fear. “Peace spaces”—community meetings for sharing and verifying information—helped defuse tensions before elections or outbreaks of conflict. In Bunia, a rumour threatening to disrupt voting was swiftly countered with anticipatory SMS messages, thus enabling peaceful participation.

Monitoring and evaluation were integral throughout: weekly data tracked the speed of rumour resolution subscriber sentiment while town halls and radio segments—co-created with local people—ensured that counter-messages resonated. This locally rooted, two-way approach proved essential in places where top-down interventions often fail.

[8] The Sentinel Project typically measures the reach of their projects using direct beneficiaries, which are the subscribers to the WikiRumours platform, and two tiers of indirect beneficiaries. The first tier is estimated based on the average number of people who live in subscriber households, and the second tier is based on the populations of the areas where a given project is active. This number relates to the overall sum of the direct beneficiaries and the second tier of indirect beneficiaries.

The Sentinel Project's success lies in its ability to turn bystanders into information stewards. Guided by the ABCD lens—focusing on Actors, Behaviours, Content, and Drivers—the team uses tools like Benesch's "Countering Dangerous Speech" and Brown's "Defusing Hate" frameworks to design timely, tailored responses[9]. Together, these approaches formed a powerful strategy for building resilience, not just correcting falsehoods.

In regions scarred by conflict and misinformation, the Sentinel Project's model stands as a blueprint for community-driven innovation: deeply embedded, rigorously verified, and fundamentally rooted in trust.



Workflow of the WikiRumours platform, showing how reports from communities are received via SMS or web, entered into the central database, verified by moderators and outreach personnel in coordination with trusted local leaders, and then returned to the community with accurate information through various channels, enabling rapid, hands-on verification and intervention.

[9] Anahi Ayala Iacucci, Navigating the Information Landscape: Misinformation, Disinformation, and Hate Speech in Humanitarian Contexts. Grand Challenges Canada, Creating Hope in Conflict initiative. 13 November 2024

IMPACT

In North Kivu and the outskirts of Juba, every rumour report flagged a community in need. According to user data provided by the Sentinel Project, over two years, more than 2,900 unique rumours were submitted by WikiRumours subscribers, generating over 4,400 sightings^[10], confirmations or denials from others. Moderators verified 94.4% of these within 24 hours, preventing false alarms that could trigger panic or flight.

What started with modest ambitions grew into a network of over 27,000 subscribers—17,000 in the DRC and 9,000 in South Sudan—far exceeding initial targets. These subscribers did not just receive updates; they became informal information hubs, relaying verified messages to families, neighbours, and beyond. In some refugee settlements, one SMS could reach an entire household of eight—often stopping a rumour in its tracks.



When several rumours were circulating around Goma, people from remote areas messaged the platform claiming that rebels or the army were retaking a neighbourhood in the city. Thanks to three “peace spaces” —including two groups of indigenous pygmy ambassadors — the team spent the night verifying these claims. By around 1AM, they were able to confirm that the reports actually referred to attack statistics in the nearby national park, not a military incursion into town.

Project Leader James Kataliko - DRC



[10] Sightings are individual reports of a given rumour. For example, if a rumour is reported by ten different people, it would be recorded as ten sightings.

The baseline survey administered by the project team revealed that just 41% of respondents ever checked a rumour. Post-intervention, 85% of subscribers believed WikiRumours had helped prevent a harmful rumour from spreading—unmistakable evidence of changing behaviours and growing trust in the platform.[11]

That trust showed up not just in statistics but in lived experience. “Before WikiRumours, reporting critical information could backfire,” said one of the community ambassador. “Now, people know we verify facts and build peace.” That sense of safety surfaced in everyday signs: fewer school absences, calmer markets, and more open conversations.

But trust did not come from simply launching an app. It came from listening.

The Sentinel Project embedded their efforts through local ambassadors, respected neighbours who knew the cultural dynamics. In South Sudan, engagement events invited open dialogue. In the DRC, recurring “peace spaces” of people gathered to discuss and verify information safely.

The Sentinel Project also engaged traditional authority figures—elders, religious leaders, even police—realizing that they often hold more sway than outside media. These trusted actors helped carry verified messages and calm fears. The project also adapted for inclusivity, reaching marginalized groups like indigenous pygmy communities and adjusting training to suit local needs.

These efforts built something powerful: a trusted space for grievances, updates, and dialogue. In volatile regions, this subtle shift toward trust and collaboration laid the foundation for a more resilient information ecosystem.



[11] The Sentinel Project team administered a baseline survey across the project sites, randomized to the extent possible in this context while aiming to achieve a cross-sectional/stratified sample of the population to reflect its diversity (gender, age, religion, ethnicity, languages spoken, etc.). The team also conducted baseline focus group discussions, with up to 15 people in each focus group. Collection methods included an KoboToolbox surveys, SMS survey, and SMS data collected over the course of the project, and interim focus group discussions. A final survey and focus group discussions were conducted to compare results across the project lifecycle.

Still, the journey was not smooth. Early resistance, digital literacy gaps, and even safety risks—like, for example, armed groups accessing moderator numbers – were considered and evaluated as serious challenges. Some communities feared reprisals or doubted the system’s relevance. The Sentinel Project countered this with expanded training, stronger feedback loops, and partnerships with radio stations and local leaders to build credibility.

Over time, the platform’s proven value for doing things like defusing election rumours and correcting dangerous health-related misinformation earned acceptance. However, the team remains vigilant: trust, they say, is something you earn daily.

By blending strong data with real human stories, the Sentinel Project showed what is possible when technology amplifies local voices. What began as small pilot projects—Hagiga Wahid in South Sudan and Kijiji Cha Amani in the DRC—has grown into full-fledged programs shaping national information flows through SMS, radio, and community leadership.



In the DRC, the Kijiji Cha Amani platform became a linchpin during the 2023 elections, where rumour control was vital. False claims about IDP voting locations in Bunia city were neutralized within hours through coordinated SMS and social media updates, averting unrest. Similarly, misinformation about voting delays in Minova town was quickly addressed through public alerts, thus ensuring calm.

Project Leader James Kataliko - DRC



This locally grounded, community-driven approach offers a replicable model for other fragile settings: one where accurate information becomes not just a tool but a form of protection. By the end of its two-year funding cycle with GCC, the Sentinel Project’s WikiRumours platform had evolved from a smaller-scale project into a trusted lifeline for tens of thousands navigating the uncertainties of conflict.

Beyond verifying thousands of rumours, it reshaped fragile information environments, fostering more resilient, interconnected ecosystems.

In areas like North Kivu and the outskirts of Juba, community ambassadors served as trusted and respected liaisons between civilians and authorities. In matter of a year, through town halls, SMS alerts, and co-hosted dialogues with women's groups, rumours gave way to trust, and speculation turned into shared understanding. These informal gatherings not only dispelled false alarms but also strengthened social ties.

Several lasting systemic impacts emerged:

- **Early Warning Systems:** Drawing from real-time platform data, the Sentinel Project developed a localized early warning prototype with the goal of empowering communities to flag risks—from militia movement to health threats—before they escalate.
- **Media Integration:** Partnerships with five radio stations extended the platform's reach beyond SMS, embedding verified updates into daily broadcasts for over 12,000 listeners.
- **Community Ownership:** Inclusive strategies ensured that women, youth, and marginalized groups shaped—not just received—the flow of information, paving the way for sustainable, locally governed systems.

The humanitarian impacts were tangible. Communities gained faster access to accurate information, thus helping to avert violence sparked by misinformation. In real crises, from armed group activity to intercommunal clashes, WikiRumours improved situational awareness, thus enabling people to make safer, more informed decisions[12].

Humanitarian organizations also benefited. In places like Yei, South Sudan, NGOs began consulting the WikiRumours team for ground-level insight before conducting fieldwork. While technology powered the platform, it was the methodology—local engagement, two-way communication, and responsiveness to community needs—that made it indispensable.

In short, WikiRumours did not just counter rumours; it rebuilt trust, shifted behaviours, and opened new channels for collaboration in two of the world's most volatile regions.

[12] Final Report to donor, Sentinel Project, 2023

GROWTH & SCALE

When WikiRumours first launched in North Kivu and on the outskirts of Juba, few could have foreseen how quickly the model would expand. By the end of its initial funding cycle, what began as a localized pilot had evolved into a scalable blueprint for community-driven rumour management—drawing global attention.[13]

Kenya & Uganda Pilots:

Prior to Grand Challenges Canada’s 2020 investment, the Sentinel Project had already tested the concept in Kenya, where its Una Hakika pilot enrolled 18,000 SMS users, showing that rural populations could and would act on locally verified updates[14]. That early success spurred extensions into northern Uganda’s Rhino Camp, where 540 people used on the platform[15] during a security scare[16] on the Katiku–Ocea Road.

South Sudan

The Sentinel Project has operated the Hagiga Wahid platform (an instance of WikiRumours) in South Sudan since 2019, implementing a small-scale expansion of the Hagiga Wahid project that they had originally started in Uganda, using locally verified SMS updates to combat misinformation. In 2025, the project team conducted a survey of 1,808 respondents across six regions to better understand communication practices, trust dynamics, and misinformation patterns. These insights will inform the next phase of Hagiga Wahid, focusing on strengthening rumour verification, enhancing early-warning systems, and expanding inclusive access to accurate information.

Democratic Republic of the Congo:

In the DRC, the Sentinel Project initially focused on building a network of trained community ambassadors. That groundwork had since expanded into an operational program with ambitions to further strengthen its broader public reach.

Recognizing the need for locally relevant scaling, the Sentinel Project launched additional efforts in Lubero, Uvira, Bukavu, and beyond.

[13] Ryan Flanagan, [This Canadian group wants to stop genocides before they happen](#), CTV News, February 12, 2020

[14] Brenna Daldorph, [Kenyans worry ‘fake news’ might spark more election violence](#), The World, August 5, 2017

[15] WikiRumours records (Confidential)

[16] Samuel Okiror, [Uganda calls in troops as violence flares between refugees and locals](#), The Guardian, 15 Sep 2020

The Kijiji Cha Amani platform laid the groundwork for scaling by creating community focal points—human nodes responsible for summarizing field intelligence and relaying it to humanitarian clusters. With this model, the Sentinel Project aims to reach 50,000 subscribers across eastern DRC.

To improve reach and relevance, peace theatre and songs were added to public radio programming. “Open days” and community festivals complemented the traditional training and rumour verification model in order to foster emotional engagement and cultural resonance.

The Sentinel Project is also exploring how children are targeted in misinformation efforts. Through a forthcoming educational track—“Mtoto bila ujeuri” (“Child without arrogance”)—they will counter recruitment propaganda and religious extremism through local storytelling and school-based dialogue.

Global:

Recognizing rising global demand for tools to combat digital misinformation, the Sentinel Project identified new countries—including Burkina Faso and Benin—as expansion targets, based on infrastructure readiness and communal tension risks. The pilot in Kenya grew well beyond the pilot phase, and had a national presence focused on the last general election in 2022. In Uganda, the project continued for a while under local leadership, marking a planned transition rather than withdrawing altogether.

To support replication, the Sentinel Project created a formal partner track offering free access to the WikiRumours code, remote training, and technical support. By mid-2022, three NGOs in Africa and South Asia had begun piloting the platform independently—evidence that peer-led scaling can be faster and more sustainable than direct deployment.

Infrastructure & Partnerships

The Sentinel Project also developed a new early-warning module: a heat map dashboard that visualizes rumour clusters for local safety committees[17]. This shifts the platform from reactive rumour verification towards proactive conflict prevention. This predictive analytics approach fills existing gaps in the MDH space, as documented in the MDH Scoping Study[18], and offers a preventive solution that should be able to feasibly measure its impact in order to improve effectiveness and pivot accordingly.

[17] Find more information [here](#).

[18] Anahi Ayala Iacucci, *Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict*, Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

Partnerships with mobile network operators—like MTN in South Sudan—opened doors to integrating rumour management into corporate social responsibility (CSR) portfolios. Talks with Airtel, Orange, and Vodacom suggest potential telecom-led sponsorship of such systems in the future.

The Sentinel project continued to attract funding from Global Affairs Canada to scale up their project in South Sudan and expand their work to Burkina Faso and Benin. The New Zealand government has also supported their work in South Sudan. To keep pace with scale, the Sentinel Project continues to work on embedding automation into the platform—adding keyword detection, bulk-verification tools, and basic natural-language sorting. These features can help reduce moderator burden and prepare the system for internet-era rumour tracking. This work is still conceptual, and the projects remain highly dependent upon human effort and judgement.

Through it all, WikiRumours has grown by design, not just demand. The Sentinel Project expands only where trusted local networks exist, tailors tools to emerging needs, and equips partners to lead. As it looks ahead, this adaptive, community-first model offers a sustainable roadmap for tackling misinformation in fragile contexts worldwide.



CHALLENGES AND OPPORTUNITIES

The Sentinel Project's journey with WikiRumours was shaped not just by successes but by challenges that demanded constant adjustment and learning.

Initially, the team focused on building a network of community ambassadors to verify rumours. Once activities began, it became clear that ambassadors alone were not enough to curb misinformation. The Sentinel Project enhanced their strategies, reaching out to local authorities, UN agencies, and NGOs—adopting a broader ecosystem approach that brought in new voices and expanded reach.

Another pivot came in the team's recruitment strategy. While the original plan called for training as many ambassadors as possible, experience showed that fewer, more deeply engaged individuals were more effective. This led to a more diversified network—including health workers, journalists, and local leaders—all helping extend the platform's reach and credibility.

Surveys revealed that residents placed more trust in figures like police, religious leaders, and elders than in radio or other media. In response, the Sentinel Project launched in-person discussion groups with these authorities—ensuring that verified messages were delivered by voices that communities already believed, thus enhancing both legitimacy and reach. This aligns with the idea that there is a pressing need for humanitarian actors to bridge online and offline modalities in addressing MDH and that technology alone cannot address the problem - social and behavioural dynamics play a crucial role in reinforcing MDH patterns[19].

“

Local media has been crucial. When an area in Beni was threatened by the ADF, community ambassadors there linked with local stationsm Radio Communautaire Bashu, Radio Moto Ocha, and in Beni, Radio Oasis. At one point, Radio Oasis's editor-in-chief joined the team for six months to help them cover that zone.

Project Leader James Kataliko - DRC

”

[19] Anahi Ayala Iacucci, [Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict](#), Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

Operationally, the team also had to manage funding hurdles, especially within the context of the COVID-19 pandemic and the implementation challenges it caused, which risked losing momentum in the field. During the COVID-19 pandemic, Grand Challenges Canada offered some grant flexibility—through no-cost extensions and reallocation approvals—which enabled the Sentinel Project to keep the project moving while utilizing those funds to leverage additional donors.

To ensure continuity, the Sentinel Project built a three-month activation buffer into its fundraising strategy and pursued multiple funding sources. They recommend that future funders also include flexible activation periods and streamlined approval processes to prevent fieldwork from stalling.

Another challenge came from potential partners seeking self-sustaining business models, which are difficult to align with the long-term nature of behaviour change. The Sentinel Project reframed its value by highlighting early wins (such as resolving 94.4% of rumours in 24 hours) and mapped out multi-year impact roadmaps. They also embedded rumour-verification training in journalism curricula to strengthen local capacity.

Meanwhile, the technology's real power lay in the ecosystem built around it. By offering open-source licensing and remote technical support, the Sentinel Project enabled peer NGOs to launch their own versions, thus turning one pilot project into a broader network of rumour management hubs.

Unexpectedly, the platform's geolocated rumour data began to reveal patterns that signalled impending violence or displacement. This insight led to a 2021 Early Warning/Early Action workshop[20], transitioning the system from reactive correction to proactive prevention[21].

Finally, the Sentinel Project developed a gender equality strategy to ensure that women and marginalized groups were not just recipients of information but co-creators of counter-messaging. This inclusive design broadened participation, even in low-literacy environments.

[20] See the recording of the workshop [here](#).

[21] The project team participated in several events both during and after the project which gave it an opportunity to raise the profile of the innovation, share knowledge generated during the project, and build relationships with other humanitarians and peacebuilders who might be able to replicate the innovation or otherwise benefit from it in their own contexts. Source: Final Progress Report, WikiRumours

These adaptations highlight a core lesson: innovation in fragile settings requires more than great technology. It demands patience, inclusive partnerships, and flexible funding models that turn friction into fuel. Through this approach, the Sentinel Project did not just solve problems—it built a replicable model for others tackling MDH in humanitarian crises.

LESSONS LEARNED

As the Sentinel Project scaled WikiRumours to full field deployment[22], they quickly learned that innovating in humanitarian crises requires more than technology. Navigating complex funding, trust, and policy ecosystems was just as essential. From this journey, four key themes emerged—offering guidance for future MDH initiatives.

1. Rumours Spread Quickly and Can Be Deadly: In fragile settings, unverified claims—from militia attacks to fake health alerts—can trigger panic and displacement. To date, WikiRumours has verified nearly 11,834 distinct rumours with around 80% resolution rate[23], proving that rapid, community-based verification is a critical form of early warning. Future projects should embed verification teams with access to fast communication channels like SMS and radio.

2. Trust Is Earned, Not Assumed: Introducing new tools in conflict zones means anticipating fear and resistance. The Sentinel Project trained 400 local ambassadors across the DRC and South Sudan and conducted deep community engagement to build acceptance. Success demands a preparatory phase with stakeholder mapping, risk planning, and culturally tailored outreach.

3. Verified Data Has Real-World Value: Beyond rumour control, WikiRumours reports have allowed communities to be active participants in their own safety. This has demonstrated to communities that their participation matters. Collecting timestamped, geolocated data—and sharing it securely with partners—can extend the platform’s impact well beyond verified information sharing and counter messaging.

[22] Initial funding to the Sentinel project were provided by the International Development Research Centre.

[23] Data taken from the Una Hakika and Hagiga Wahid platforms analytics. A rumour is resolved when it is not under investigation, impossible to verify or yet to be verified. When resolved the rumour is marked as: Confirmed True, Confirmed False, Probably True, Probably False.

“

An elderly woman was stopped by an armed group in the DRC. They accosted her, rifled through her bag, inspected her phone and even her food, then finally let her go, saying, “We’re not dealing with you—you’re just an old woman.” She went home and warned her children. The WikiRumours ambassadors immediately went to speak with the local commander to stress the urgency of the situation, which dispatched security personnel to the area. Without that direct intervention, entire villages could have been attacked...

Project Leader James Kataliko - DRC

”

4. Ambassadors Are More Than Messengers: Trained ambassadors became trusted figures. Their training included not just reporting, but also conflict resolution and safety protocols. Community members, properly equipped, can better step into the role as first responders and peacebuilders.

5. Fast Responses Build Community Trust: The Sentinel Project’s team members tried to respond to every high-risk report within 24 hours. This consistency encouraged participation and trust. Future programs should prioritize lean, decentralized verification systems that quickly close the loop with reporters using the same channel through which they submitted a given report.

6. Sustainability Requires Private-Sector Foresight: Developing partnerships with mobile network operators like MTN and Zain has proven to be slow and difficult. Future implementers must budget time, staff, and third-party support for private sector collaboration early, treating it as a core—not auxiliary—component of sustainability.

7. Gender Barriers Must Be Proactively Addressed: Despite a gender strategy, women’s participation sometimes lagged due to lower levels of phone access and literacy. Effective inclusion requires pairing digital tools with offline support—like device access, digital skills training, and engaging male gatekeepers in awareness campaigns.

8. Measuring Impact Takes More Than Surveys: Qualitative evaluations like focus groups provided valuable insights but required more resources than expected. Future projects should budget for M&E as a distinct workstream with flexible timelines, contingency funds, and local support built in from the start.

THE FUTURE

By mid-2022, as WikiRumours solidified its role in North Kivu and South Sudan, the digital and policy landscape around it began to shift. The 2024 MDH scoping study sounded the alarm: the rise of AI-generated content—deepfakes, fake audio, and algorithmically amplified disinformation—posed new threats, capable of spreading panic faster than any human moderator could react.

For WikiRumours, this meant evolving beyond reactive rumour verification. The challenge now is to pre-filter large volumes of synthetic content—especially in fast-moving WhatsApp or Facebook groups. However, the same AI tools that can automate falsehoods can also help to combat them. The Sentinel Project is exploring machine learning to detect urgent or novel threats early, thus freeing moderators to focus on local nuance and trust building.

Policy changes mirror these shifts. While early hopes centred on prioritizing lifesaving SMS channels, global discussions have moved towards content liability and censorship. Takedown laws meant to curb hate speech could unintentionally silence community voices. Elements of the WikiRumours model, such as anonymized submissions stored offshore, offer a template for balancing openness and user safety.

Two key technology fronts stand out. First, data privacy: storing submissions anonymously and have kept users safe so far but rising risks like state surveillance require stronger safeguards. Second, the early warning system—born from routine analytics—can give community safety committees the tools to act before crises escalate.

Support has come from many fronts. The Organisation Internationale de la Francophonie backed the platform during DRC's election period. Government actors and local media did not just broadcast information—they helped to verify it. One radio editor even joined the team.

The platform also reached a major technological milestone: development of a mobile app, which significantly expands its future reach—especially in smartphone-using regions.

These developments point to key imperatives for the MDH field:

- **Blend human trust with AI tools that scale.**

In conflict settings, trust lives with people (ambassadors, elders, clergy, local officials), while volume lives in the channels (SMS, WhatsApp, radio). AI can triage the flood; humans carry legitimacy. The 2024 MDH scoping study [24] cautions that AI alone won't fix MDH — social and behavioural dynamics still drive what people believe, and approaches must stay adaptable as tech changes.

- **Advocate for smart policies that protect community-led verification.**

Well-intended takedown regimes and blunt platform rules can silence the very channels communities use to keep themselves safe. The MDH scoping study[25] flags a sector tilt toward regulation while stressing local ownership and accountability to affected populations as non-negotiables.

- **Design systems not just for today's rumours, but tomorrow's threats.**

Threats are evolving: AI-generated audio/video, faster amplification, and offline intimidation. The MDH scoping study[26] urges proactive, preventative approaches and investment in MDH infrastructure, not just one-off responses.

As WikiRumours prepares for its next phase—deeper AI integration, expanded telecom partnerships, and scaled early warning tools—it offers a clear lesson: combating harmful information in crisis zones will not be solved by top-down or technological fixes alone. The future lies in adaptive, locally grounded networks that pair innovation with the wisdom and agency of the communities they serve.

[24] Anahi Ayala Iacucci, [Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict](#), Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

[25] Anahi Ayala Iacucci, [Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict](#), Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

[26] Anahi Ayala Iacucci, [Scoping Study – Misinformation, Disinformation and Hate Speech in Conflict](#), Grand Challenges Canada, Creating Hope in Conflict: a Humanitarian Grand Challenge, November 2024

CONCLUSION



What makes the Sentinel Project's model stand out is its fusion of grassroots ownership with lightweight technology. Rather than relying on costly, internet-heavy systems that exclude the most vulnerable, WikiRumours transforms everyday mobile phones into nodes in a decentralized verification network. It does not just stop rumours; it builds a lasting culture of community trust.

The Sentinel Project proves that innovation in MDH requires systems that centre people, trust, and adaptability. WikiRumours shows what is possible when communities lead and funders listen.

From this journey, several key recommendations emerge for future funders and implementers:

Recommendation	Actions for Funders	Actions for Implementers
Build local ownership from day one. Tech alone doesn't create trust—relationships do. Early co-design and compensation for community actors make systems relevant and resilient.	Resource co-design sessions with community, faith, women's and youth groups; fund local partner contracts and training.	Formalize MOUs with radio/media and civic actors; recruit and pay community ambassadors; hold regular listening sessions and iterate based on feedback.
Support long-term, flexible programming. Behaviour change and trust take years, not months. Flexibility protects impact when contexts shift.	Structure 2–5-year early disbursements or readiness grants, with phase gates, pivot clauses, and contingency funds for shocks (elections, outbreaks).	Set phased targets (reach → trust → behaviour); run quarterly adaptation reviews; use early-stage funding mechanisms or local partnerships to sustain pre-launch activities.
Treat private-sector engagement as strategy, not outreach. Mobile network operators and media access takes time, negotiation, and dedicated ownership. Treat it like a core workstream.	Budget a relationship manager; fund travel and compliance; tie milestones to signed MOUs and service level agreements.	Map decision-makers (technical + regulatory); start procurement/security reviews early; lock predictable airtime blocks with radio/TV partners.
Keep it human-led, AI-assisted. Trust lives with people; scale lives in the channels. Pair light machine learning (dedupe, keyword flags) with local moderators to move fast without losing legitimacy.	Finance simple ML pilots and data labelling; fund moderator stipends tied to a 24-hour SLA; cover basic data infrastructure.	Stand up a human-in-the-loop flow (intake → AI pre-filter → moderator verify → message back); define RACI; run weekly quality checks and maintain a 24-hour verification SOP.
Invest in anticipatory tools & playbooks. Moving from reaction to prevention saves lives and resources. Simple early-warning stacks and pre-approved messages speed response.	Fund lightweight toolsets (heat maps, spike alerts) and scenario exercises/tabletop drills.	Define trigger thresholds for action; pre-write SMS/radio "preventive postcards"; run quarterly drills and after-action reviews.
Break down barriers to participation. Inclusion isn't automatic; design must fit low-literacy, shared devices, and safety needs, especially for women and marginalized groups.	Make inclusion deliverables explicit (gender, language, disability); fund interactive voice response options and translation.	Train and compensate female ambassadors; design for low literacy and shared phones; localize languages; provide offline options (radio, town halls).
Design for resilience & security by default. Adversaries adapt; channels fail. Privacy-by-design and redundancy keep communities safe and services running.	Pay for privacy audits, threat modelling, and crisis comms reserves; fund core security tooling.	Rotate hotline numbers; anonymize data at rest/in transit; maintain channel redundancy (SMS/app/radio); maintain an OPSEC checklist and drills for staff/volunteers.



<https://humanitariangrandchallenge.org/>



info@humanitariangrandchallenge.org

CREATING
HOPE
IN CONFLICT:

A HUMANITARIAN
GRAND CHALLENGE