

Baseline Assessment

On strengthening faith-based and interfaith actors

to counter digital harm and promote
peaceful civic participation ahead of
Kenya's 2027 elections

September 2026

HOMA BAY · NAKURU · LAMU · NAIROBI · BUSIA · MOMBASA





Baseline assessment on strengthening faith-based and interfaith actors to counter digital harm and promote peaceful civic participation ahead of Kenya's 2027 elections

I. EXECUTIVE SUMMARY



SIX COUNTIES

Homa Bay · Nakuru · Lamu · Nairobi ·
Busia · Mombasa

TikTok · Facebook · WhatsApp ·
Instagram

most frequently identified

X · YouTube · Signal · Telegram

also relevant

Executive summary

This baseline assessment establishes the contextual, institutional and qualitative starting point for FaithConnect, an initiative intended to strengthen faith-based and interfaith actors to identify, verify and counter harmful digital narratives while safeguarding freedom of religion or belief and promoting peaceful civic participation ahead of Kenya's 2027 General Election. It covers Homa Bay, Nakuru, Lamu, Nairobi, Busia and Mombasa counties and examines misinformation and disinformation, hate and dangerous speech, religious incitement, gendered digital violence, impersonation and account compromise, and harms affecting civic space and access to information.

The assessment combined a review of project and institutional documents with primary qualitative research. Four core project documents were reviewed alongside evidence from public institutions, faith-sector bodies and digital-governance organizations. Primary data comprised semi-structured interviews with faith leaders, interfaith actors, youth peace champions and civil-society representatives, together with discussions involving three Inter-Religious Council of Kenya interlocutors.

The baseline finds that digital harm is already embedded in Kenya's electoral and civic environment. TikTok, Facebook, WhatsApp and Instagram were most frequently identified as channels through which harmful content circulates, with X, YouTube, Signal and Telegram also relevant. Closed messaging groups present particular monitoring and verification challenges. County risks differ: Nakuru faces ethnic incitement and coded political language; Lamu faces phishing, closed-group harms and online radicalization concerns; Mombasa and Homa Bay show prominent gendered political abuse and youth-related risks; Busia faces misinformation, scams, impersonation, cyberbullying and ethnic inflammatory content; and Nairobi is characterized by high-volume information flows, rapid amplification and tensions between harm prevention and freedom of expression.

Young people emerged as the group most consistently perceived to be exposed to—and sometimes mobilized in—the production and circulation of digital harm. Economic exclusion, peer pressure, political manipulation, attention incentives and low verification habits increase vulnerability. Women and girls face distinctive harms, including sexualized attacks, manipulated images, impersonation, sexist abuse, public shaming and attacks on credibility. At the same time, unequal digital access can limit women's access to verification, reporting and civic-engagement resources. Faith leaders retain substantial public trust and possess strong congregational, pastoral and community networks that can bridge online and offline spaces. Their contribution is nevertheless conditional on non-partisanship, pulpit integrity and meaningful youth participation; perceived political alignment can erode their credibility and deepen intergenerational mistrust.

FaithConnect has a strong platform on which to build. TrustLab already provides

T0 · T1 · T2

enrolment · immediately after training ·
endline

SET AT T0

six-county Barazas, locally embedded Sentinels, community-facing Mythbusters, monitoring and reporting processes, learning resources, and an ACTORS capacity-building network. The project should therefore integrate faith and interfaith actors into these existing workflows instead of creating a parallel system. The mapping frame contains 28 prospective actors, although it is unevenly distributed across counties. Nakuru, which had no primary interview respondent in this assessment, requires early validation through county-level key-informant interviews and the standardized pre-training assessment.

A central methodological finding is that the current qualitative evidence cannot yield valid percentages for participants' knowledge, confidence or demonstrated skills. Core indicators—including the ability to identify misinformation, distinguish hate speech or religious incitement from legitimate expression, complete a verification process, respond safely and select an appropriate referral channel—should therefore be established at enrolment as time-zero (T0) values. The same matched instrument should be administered immediately after training (T1) and at endline (T2), using anonymous identifiers, standardized scenario tests, an observed verification task, a safe-response and referral simulation, and a short confidence scale. Results should be disaggregated by sex, age, county, actor category, disability status where safely collected, and prior training exposure.

Priority actions for implementation are to: integrate FaithConnect with TrustLab's Sentinel, Mythbuster, Baraza, verification and counter-narrative systems; adopt a competency-based baseline and endline design that measures demonstrated skill as well as self-confidence; combine a common national curriculum with county-specific modules and validate the Nakuru profile early; make Barazas intergenerational learning spaces in which young people and faith leaders share substantive facilitation and leadership roles; aim for gender parity, with 40 per cent as an absolute minimum, and strengthen women's leadership, safeguarding and referral support; teach contextual classification and freedom-of-expression safeguards rather than relying on keyword-based content flagging; adopt clear non-partisanship and pulpit-integrity standards for all participating faith actors; audit referral pathways, track whether cases are received, responded to and acted upon, and measure trust in each pathway; and maintain volunteer-retention, incident, platform-context and external-disruption logs so that endline results can be interpreted accurately.

STAKEHOLDER PROPOSAL

A subsequent validation workshop confirmed the need to strengthen the treatment of synthetic media, the gap between young people's online experiences and faith institutions' offline responses, constitutional and civic education, political interference, and accountable referral pathways. The workshop recommendations are incorporated below as stakeholder proposals; they are not presented as additional measured baseline results.

Overall, the baseline supports the relevance and feasibility of FaithConnect, while cautioning against treating institutional activity or interview perceptions as participant-level competency values. The programme's strongest pathway to impact is to combine the moral authority and offline reach of faith actors with youth platform knowledge, TrustLab's existing infrastructure, practical skills assessment, county-specific adaptation, functioning referral systems, and firm safeguards for gender inclusion, non-partisanship and legitimate civic expression.

II. TABLE OF CONTENTS



I. Executive summary	3
II. Table of contents	5
III. Abbreviations and acronyms	6
IV. List of tables and figures	7

REPORT

1.0	Introduction	8
	1.1 Headline baseline findings	13
	<i>Figure 1: County Risk Assessment Profile</i>	12
	<i>Figure 2: Primary Platforms for Harmful Content Circulation</i>	12
	<i>Figure 3: Vulnerable Populations & Risk Factors</i>	14
2.0	Why faith-based and interfaith actors matter?	15
3.0	Assessment scope and methodology	19
	3.1 Documentary and secondary sources	20
	3.2 Primary data	21
	3.3 Methodological limitations and interpretation	21
4.0	Baseline findings and county profiles	23
	4.1 Existing TrustLab platform as the immediate institutional baseline	24
	4.2 Recognition and verification of harmful digital content	26
	4.3 ACTORS mapping	27
	4.4 Gender and inclusion baseline	27
	4.5 Demographic difference and overlaps in faith-based digital harm counter	28
	4.6 Region	28
	4.7 Gender	29
	<i>Figure 4: Baseline Respondent Demographics (n=18)</i>	30
	<i>Figure 5: Six Categories of Harm - Interview Frequency</i>	31
	4.8 Baseline indicator framework	32
	4.9 Measurement protocol for demonstrated skills	33
4.10	Recommendations	34

III. ABBREVIATIONS AND ACRONYMS



Abbreviations and acronyms

ABBREVIATION	FULL FORM
AI	Artificial Intelligence
CA	Communications Authority of Kenya
CSO	Civil Society Organization
DCI	Directorate of Criminal Investigations
ICT	Information and Communication Technology
IEBC	Independent Electoral and Boundaries Commission
IRCK	Inter-Religious Council of Kenya
KII/KIIs	Key Informant Interview(s)
KNBS	Kenya National Bureau of Statistics
M&E	Monitoring and Evaluation
NCCK	National Council of Churches of Kenya
NCIC	National Cohesion and Integration Commission
T0	Time-zero assessment administered before training
T1	Assessment administered immediately after training
T2	Endline assessment

IV.

List of tables and figures

TABLES

Table 1: Six Categories of Harm	10
Table 2: Actors Distribution	27
Table 3: Demographic Summary of Respondents and Interlocutors	29

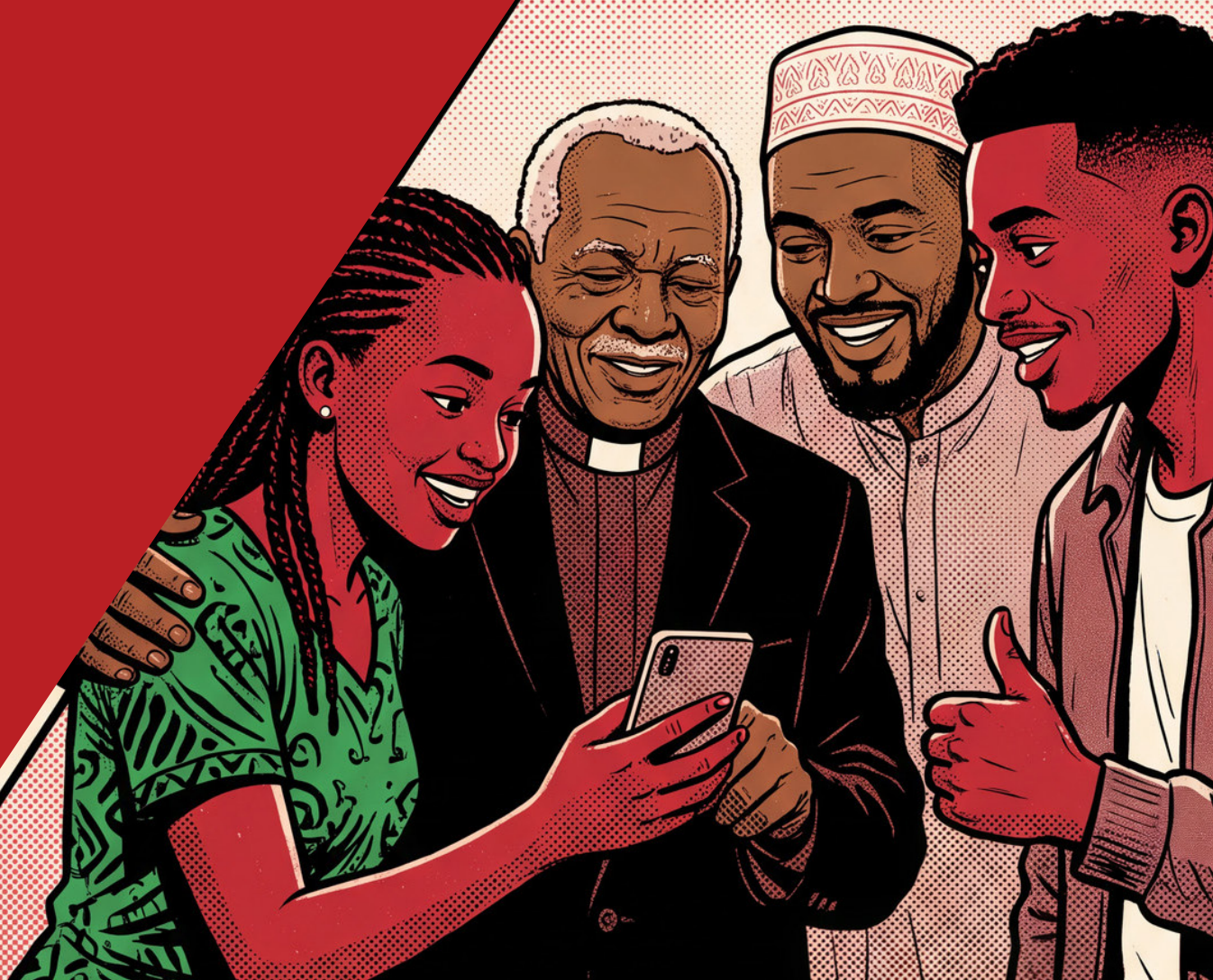
FIGURES

Figure 1: County Risk Assessment Profile	12
Figure 2: Primary Platforms for Harmful Content Circulation	12
Figure 3: Vulnerable Populations & Risk Factors	14
Figure 4: Baseline Respondent Demographics	30
Figure 5: Six Categories of Harm - Interview Frequency	31

1.0 INTRODUCTION

1.0

Introduction



1.0 INTRODUCTION

2027

Tuesday, August 10, 2027

The Constitution requires general elections on the second Tuesday of August in every fifth year.

Kenya is entering an extended period of electoral preparation. The Constitution requires general elections on the second Tuesday of August in every fifth year. On the present constitutional cycle, that corresponds to Tuesday, August 10, 2027. In June 2026, IEBC launched its Strategic Plan 2024–2029 and Election Operations Plan 2025–2027, under the theme “Inclusion and Participation in Deepening Democracy in Kenya,” demonstrating that formal election preparation is already underway. Peaceful civic participation and strong social cohesion remain, however, significant in a multiethnic and religious dynamic Kenyan society, especially during an election full of emotions and political tension.

ABOUT 35%

of Kenyans aged three years and above used the Internet

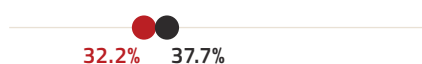
58.6%

among young people aged 18–34



Source: CA/KNBS analytical report based on the 2023/24 Kenya Housing Survey

Females / males



Youth aged 18–34: women / men



■ Women / females ■ Men / males

Source: CA, March 2026

Digital platforms now form an integral part of this civic environment. A CA/KNBS analytical report based on the 2023/24 Kenya Housing Survey found that about 35% of Kenyans aged three years and above used the Internet, rising to 58.6% among young people aged 18–34. Internet use is particularly high in Nairobi at 64.7% and Mombasa at 46.9%, although national figures also reveal considerable urban-rural and socio-economic disparities. These figures mean that an online peace strategy is essential, but it cannot assume universal or equal digital access.

The gender dimension is significant. CA reported in March 2026 that 32.2% of females used the Internet compared with 37.7% of males; among youth aged 18–34, Internet use was 54.8% among women and 62.3% among men. Consequently, the project must address two simultaneous realities: women can be heavily exposed to gendered digital harm while also experiencing somewhat lower access to the digital resources through which verification, reporting, and civic engagement increasingly occur.

The TrustLab interim report and the primary interviews show that the six project counties face a varied spectrum of digital harms: online radicalization, misinformation, hate speech, gendered cyberbullying, phishing, account hijacking, impersonation, and scams. In Nakuru, for example, a Sentinel reportedly flagged material calling for the eviction of Kalenjin, Luo, and Luhya communities, as well as a separate post invoking Kikuyu supremacy and historical violence. A localized peace counter-narrative subsequently reached more than 2,000 people. In Lamu, a WhatsApp phishing attempt against an arts organization prompted verification, digital-safety support, and public-awareness messaging. Busia experienced a documented case involving the hijacking of a local political figure’s WhatsApp account, while cases in Mombasa and Homa Bay involved public shaming and gendered cyberbullying of female public figures. Historical election evidence reinforces these local observations. Mozilla found hate speech, incitement, and manipulated political content on TikTok before the 2022 election and reported that moderators sometimes lacked sufficient linguistic or political-context knowledge to identify violations reliably. ARTICLE 19 documented false election-day

reports, fabricated claims about results, fake polls, impersonation, and misleading videos, while emphasizing the importance of local linguistic and cultural context in content moderation. The risk therefore extends beyond false information. Interviewees described cases in which online political disputes escalated into threats or physical confrontation, and in which defamatory content deepened hostility and weakened social cohesion.

The baseline consequently distinguishes six categories of harm:

Table 1: Six Categories of Harm

	HARM CATEGORY	EXAMPLES
01	Misinformation/ Disinformation	False electoral procedures, fabricated results, misleading polls, false candidate claims, manipulated or AI-generated content
02	Hate and Dangerous Speech	Ethnic, religious, gender, or identity-based vilification; calls for exclusion or violence
03	Religious Incitement	Content framing political competition as religious confrontation or using theology to legitimize exclusion or violence
04	Gendered Digital Violence	Sexualized attacks, misogynistic misinformation, public shaming, threats, doxxing, manipulated images
05	Impersonation, Scams & Account Compromise	Fake political accounts, hijacked WhatsApp accounts, phishing, fraudulent solicitations
06	Civic-Space & Information-Access Harms	Coordinated intimidation, suppression of peaceful expression, connectivity/platform disruption, surveillance concerns

CASE NOTE

“tugeges” — Oversight Board, December 2025

The last category matters because “countering digital harm” must not become a synonym for suppressing criticism. During the June 2025 protests, Access Now reported hours-long disruption affecting Telegram on several Kenyan networks despite a court injunction against interference with Internet and Telegram access. Peaceful civic participation, therefore, requires both protection from harmful content and protection of legitimate access to information and political expression. This distinction has become especially clear in platform moderation. In its December 2025, Kenya decision concerning the term “tugeges,” Meta’s Oversight Board found that the company had wrongly treated a contextual political expression as a prohibited ethnic slur. The Board emphasized that meaning can evolve rapidly and that Kenyan political language cannot be classified accurately without linguistic, social, and political context. This has an important implication for faith-based programming: participants should not merely learn to “flag bad words.” They need to learn contextual harm assessment—distinguishing misinformation from opinion, theological disagreement from religious incitement, political insult from legally or ethically actionable hate, and offensive speech from speech that creates a credible risk of discrimination or violence.

INTERVIEWS

The interviews identified TikTok as a major channel for false information and other online harms. A religious leader from Homa Bay observed that, although TikTok is widely used for youth entertainment, it is underused for constructive civic education. Facebook, WhatsApp, and Instagram were also mentioned frequently, while X and YouTube were mentioned less often. Respondents emphasized the importance of harmful content circulating in closed groups, especially WhatsApp groups, where monitoring and verification are more difficult.

Participants identified the Communications Authority of Kenya, the Directorate of Criminal Investigations, religious institutions, community actors, local administration, and youth organizations as potentially important response partners. They nevertheless raised concerns about institutional trust, privacy, unclear reporting procedures, and inconsistent enforcement. Several respondents therefore favored a multi-stakeholder referral and response system for closed-group harms, combined with stronger user accountability and clear safeguards against privacy violations.

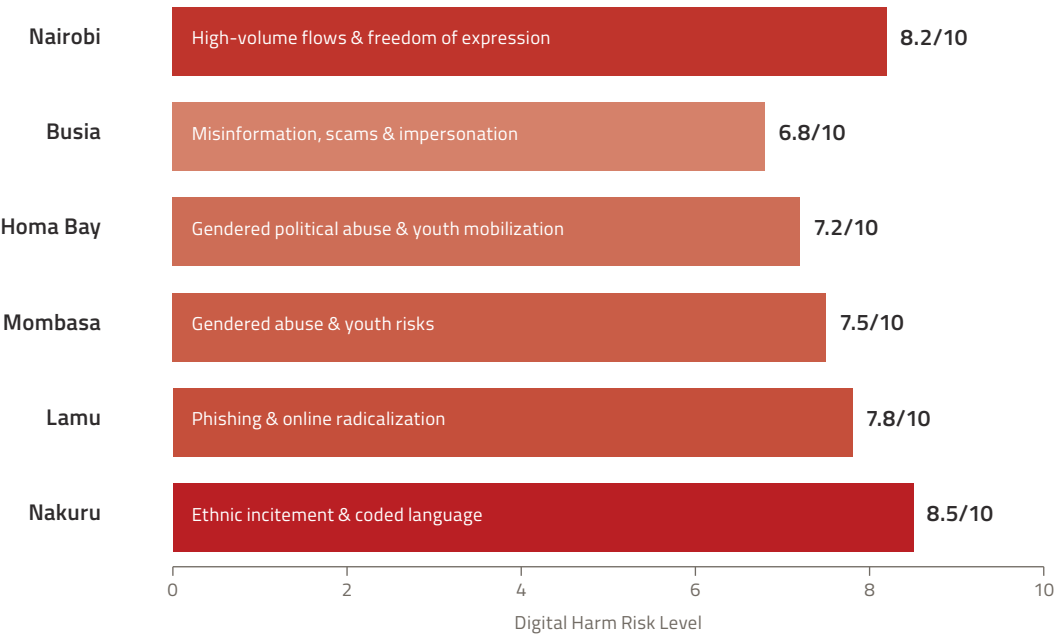


1.0 INTRODUCTION

INTERVIEWS

TRUSTLAB RECORDS

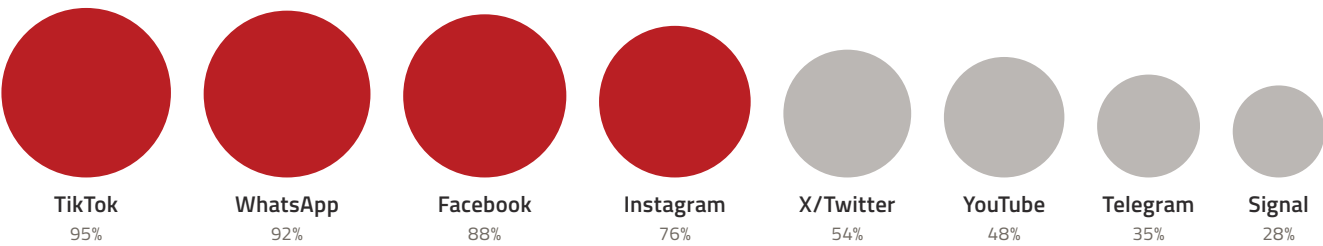
Figure 1: County Risk Assessment Profile
Digital Harm Risks Across Six Implementation Counties



The above visualization presents a comparative overview of digital harm risks across the six project counties. The varied heights and colors reflect the distinct threat profiles identified through interviews and TrustLab documentation.

INTERVIEWS

Figure 2: Primary Platforms for Harmful Content Circulation
Bubble size represents relative platform prominence in baseline interviews



Named in the Executive Summary as most frequently identified Also relevant

The bubble chart illustrates relative platform prominence in respondent interviews. Larger circles represent platforms most frequently cited as channels for digital harm, with TikTok, Facebook, and WhatsApp emerging as the primary vectors for misinformation, hate speech, and gendered abuse. Closed messaging groups on WhatsApp present particular verification challenges due to limited visibility and rapid information spread.

1.1 HEADLINE BASELINE FINDINGS

INTERVIEWS

TRUSTLAB RECORDS

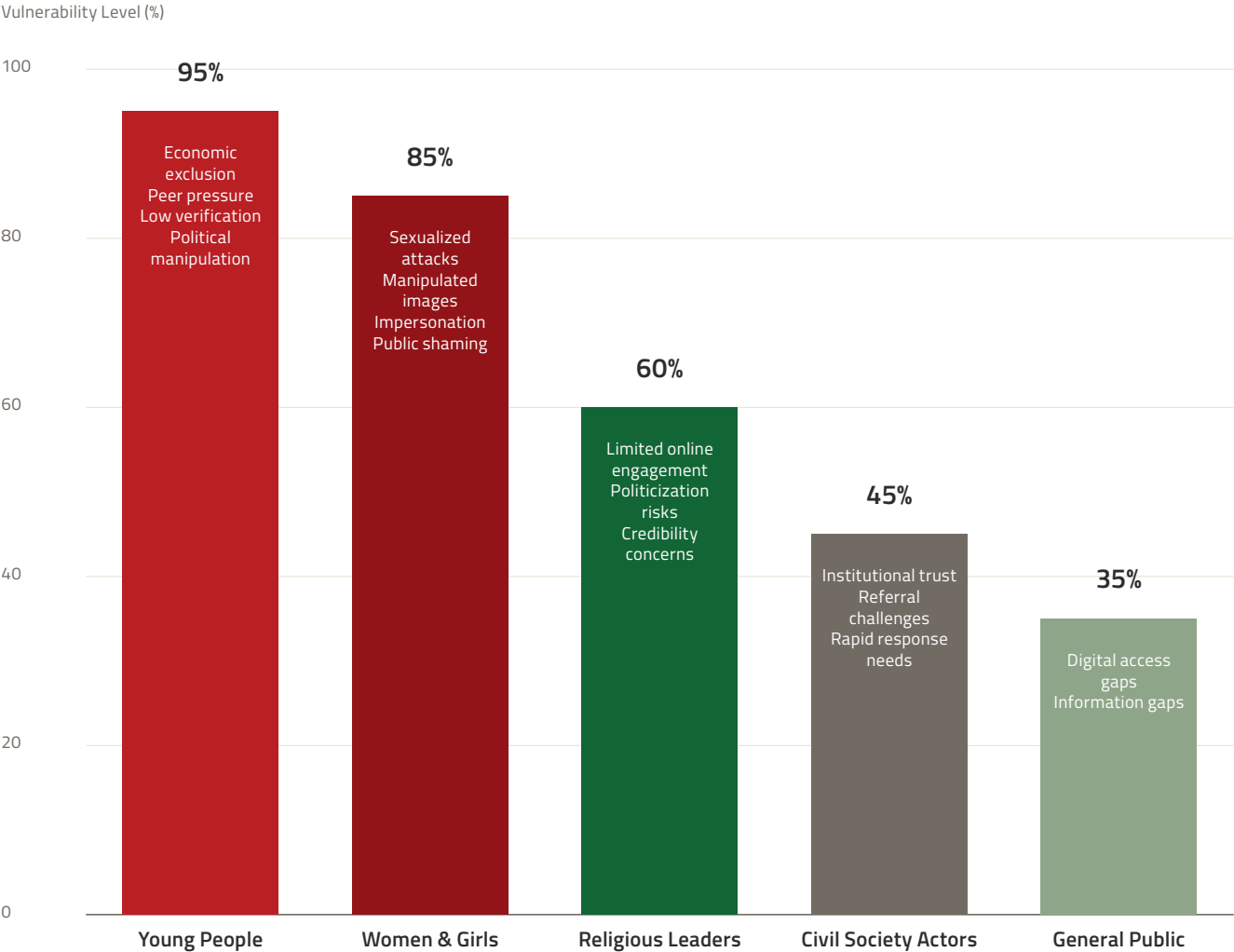
1.1 Headline baseline findings

Young people emerged as the group most consistently perceived to be exposed to digital harm. At the same time, several interviews revealed a trust and communication gap between young people and some religious leaders. FaithConnect should therefore treat youth not only as beneficiaries but also as co-educators, co-designers, Sentinels, Mythbusters, and peer mobilizers. Faith leaders retain substantial public trust and possess strong offline and online community networks. That advantage is conditional: politicized religious spaces, partisan endorsements, and conflicts of interest can weaken their credibility. Non-partisanship, pulpit integrity, and the distinction between civic education and political endorsement are therefore both curriculum priorities and measurable implementation standards. Digital harms vary by county. A common six-county curriculum should be complemented by county-specific modules covering the dominant platforms, narratives, vulnerabilities, and response pathways identified in each location. The qualitative interviews establish patterns, mechanisms, and implementation priorities, but they do not provide statistically valid percentages for participant knowledge or skills. These cohort-level values must be established at enrolment through a standardized pre-training assessment and compared with a matched endline assessment.

1.1 HEADLINE BASELINE FINDINGS

INTERVIEWS

Figure 3: Vulnerable Populations & Risk Factors
Differential Vulnerability Across Respondent Categories



This bar chart demonstrates differential vulnerability across respondent categories. Young people consistently emerge as the highest-risk group, followed by women and girls facing gendered harms. Religious leaders show moderate vulnerability, though their influence on digital spaces is significant. Key risk factors include economic exclusion, peer pressure, low verification habits, political manipulation, and attention incentives—all factors that FaithConnect must address through targeted training and mentorship.

2.0

**Why faith-based
and interfaith
actors matter?**



2.0 WHY FAITH-BASED AND INTERFAITH ACTORS MATTER?

70%

of citizens said they trusted religious leaders somewhat or a lot



Source: Afrobarometer, 2025 Kenya findings (section 2.0)

2.0 WHY FAITH-BASED AND
INTERFAITH ACTORS MATTER?

FROM THE REPORT

Faith leaders are useful precisely because they command trust; misuse of that trust could magnify rather than reduce polarization.

INTERVIEWS

The strongest quantitative justification for faith actors is their social credibility. In Afrobarometer's 2025 Kenya findings, 70% of citizens said they trusted religious leaders somewhat or a lot. This trust was affirmed in the interviews, where, despite the low response of religious leaders to harmful digital content, contentious trust between them and youths, and abuse of religious space by political groups, according to several participants, youths are still able to believe in religious institutions compared to youth peace champions when it comes to correcting the proper political use of digital spaces. This is due to respect for God. Youth peace champion, though, could be fairly believed due to concerns about whether their role is a paid role or a moral role.

Faith communities also possess infrastructure that digital-only interventions frequently lack: recurring congregational gatherings, local leaders embedded in neighborhoods, women and youth networks, pastoral relationships, and opportunities to transmit corrective information both online and offline.

IRCK · NCIC · NCCK

election-preparedness engagement, 2026

DOCUMENT REVIEW

The project also arrives at a point when organized religion is already moving into election preparedness. IRCK's July 28, 2026, homepage lists a joint NCIC–IRCK statement on threats to national cohesion ahead of the 2027 General Election. Reporting on the joint engagement indicates plans to strengthen early-warning arrangements, expand county interfaith peace networks, reinforce mediation mechanisms, revitalize peace infrastructure, and use worship spaces to promote peaceful messages. NCCK actors likewise were conducting election-preparedness engagements in 2026 emphasizing peaceful, inclusive, issue-based politics and rejecting tribalism, violence, and political propaganda.

IRCK's current programming provides a useful institutional model. Its 2026 LIMCO initiative describes an approach built around strengthening interfaith structures, creating early-warning and response mechanisms, promoting social cohesion, and linking local actors with county and national peace institutions. Its Interfaith Youth Manifesto also identifies digital safety, youth leadership, peace and social cohesion, and rebuilding institutional trust as interconnected priorities.

There is, nevertheless, a second side to faith influence. Religious institutions themselves can be politicized. A July 2026 NCCK-linked study reported concerns about political money, use of religious spaces for political visibility or endorsement, and

youth vulnerability to manipulation. Additionally, youth feel that faith leaders are often casting them rather than morally guiding them. The faith actions have limitedly included youths' voices as religious leaders from Homa Bay highlight that there have been few online platforms where religious leaders specifically castigate harmful online messages, and if there is even a press conference, it's more of pointing out at the youths rather than corrections. The problem is that when youths are pointed at rather than guided and corrected, they will eventually accuse religious leaders of political alignment rather than moral duty. The FaithConnect curriculum should therefore include non-partisanship, pulpit integrity, conflict-of-interest safeguards, and boundaries between civic education and political endorsement. Faith leaders are useful precisely because they command trust; misuse of that trust could magnify rather than reduce polarization.

Coast · Mombasa and Lamu

INTERVIEWS

Interviews additionally reveal a variance of responses on whether religious identities and places amplify political tension. While most religious leaders believe that their teaching has been geared towards ensuring peace and less harmful online information, a few religious leaders and several youth peace champions generally believe that the religious leaders have favored political rhetoric and identities that have encouraged political tension. In regions where communities and youths identify with a particular region, like Islam in Mombasa and Lamu County, youth have been socialized through religious messages to take a particular political stand.

COUNTY FOCUS

BUSIA

INTERVIEWS

The Busia interviews similarly present faith institutions as both potential amplifiers of political tension and important agents of peace. Respondents observed that religious platforms and leaders may be used to endorse political candidates or disseminate partisan messages. Conversely, faith narratives centred on peace, forgiveness, unity, tolerance, and non-violence were considered effective in reducing tensions. Trusted religious leaders, community-based organizations, chiefs, teachers, local media, and relevant public institutions were identified as credible sources of correction. Referral pathways remain uneven, although respondents mentioned the police, the National Cohesion and Integration Commission (NCIC), faith institutions, actors, and social-media platforms. Sustainability of a faith-linked Sentinel/Mythbuster mechanism would require community ownership, training of faith leaders and youth, partnerships with local institutions and media, modest locally supported financing, income-generating initiatives, and regular refresher training.

3.0

Assessment scope and methodology



3.0 ASSESSMENT SCOPE AND
METHODOLOGY

16

key stakeholders: key informant
interviews

13

semi-structured interviews

3

interlocutors from the Inter-Religious Council
of Kenya (IRCK)Homa Bay · Mombasa · Lamu · Nairobi
· Busia

INTERVIEWS

FIVE INTERRELATED QUESTIONS

4

documents from the supplied project
package were reviewed

DOCUMENT REVIEW

The consultancy assessed the knowledge, attitudes, and practices of faith leaders, interfaith actors, and youth peace champions regarding misinformation, hate speech, religious incitement, and peaceful civic participation; determine their capacity and confidence to identify and respond to harmful digital narratives; examine community mechanisms and gaps; disaggregate evidence by sex, county, and actor category; and refine indicators and targets where required.

This was done through literature review and existing documentation and key informant interviews of 16 key stakeholders that include thirteen semi-structured interviews with faith leaders, interfaith actors, youth peace champions, and civil-society representatives drawn from the Islamic and Christian faith traditions in Homa Bay, Mombasa, Lamu, Nairobi, and Busia counties. Three interlocutors from the Inter-Religious Council of Kenya (IRCK) were also engaged.

The proposed project itself has an explicit outcome orientation. Its overall goal is to strengthen faith-based and interfaith actors so they can identify, verify, and counter harmful digital narratives while protecting freedom of religion or belief and reducing election-related tensions. The baseline addresses five interrelated questions:

- 1. What is the pre-existing digital-harm environment?** This concerns the types of narratives, platforms, actors, vulnerabilities, and county-specific threats already visible.
- 2. What capabilities already exist?** This includes prior training, faith-based networks, community-response mechanisms, TrustLab infrastructure, ACTORS networks, and national institutions.
- 3. What is genuinely unknown?** Participant knowledge, verification proficiency, confidence, response behavior, exposure, and attitudes cannot be inferred from institutional activity alone and require direct measurement.
- 4. Which indicators require direct measurement?** The assessment identifies whether the proposed project indicators are sufficiently specific and measurable to demonstrate change.
- 5. Who should be interviewed and surveyed?** The assessment develops a practical respondent frame from the supplied ACTORS list and the broader faith/digital-governance ecosystem.

3.1 Documentary and secondary sources

Four documents from the supplied project package were reviewed: the Terms of Reference, the Federal Foreign Office project application, the TrustLab Interim Narrative Report, and the ACTORS research spreadsheet. External triangulation

prioritized primary and institutional evidence. Siasa Place’s programme information confirms that TrustLab Africa is implemented with DW Akademie and Code for Africa in the same six target counties. UNESCO evidence was used to assess recent faith-leader training and the national FeCoMo architecture; Communications Authority and KNBS evidence informed the analysis of digital access; Kenya Law and IEBC sources informed the electoral context; and IRCK and NCIC sources informed the analysis of faith-based peacebuilding and cohesion infrastructure. Evidence from Mozilla, ARTICLE 19, Global Witness, Access Now, and the Oversight Board was used to triangulate digital-harm patterns and platform-governance risks.

3.2 Primary data

Thirteen semi-structured interviews were conducted with faith leaders, interfaith actors, youth peace champions, and civil-society representatives drawn from the Islamic and Christian faith traditions in Homa Bay, Mombasa, Lamu, Nairobi, and Busia counties. Three additional interlocutors from the Inter-Religious Council of Kenya (IRCK) brought the total number of participants and interlocutors to 16. The interview questions were harmonized with the Terms of Reference to capture the objectives of the baseline. Respondents were purposively selected because their prior engagement in digital-harm training, community activities, peacebuilding, or faith-based work enabled them to provide informed insights relevant to the six target counties: Homa Bay, Nakuru, Lamu, Nairobi, Busia, and Mombasa. Interviews were conducted remotely to support timely access to participants, recorded with consent, and transcribed. The interview guide was translated into Kiswahili where necessary, and Kiswahili responses were subsequently translated into English. Interview data were analyzed thematically around issues including the recognition and verification of harmful digital content, faith dynamics, vulnerability, referral pathways, and sustainability. Primary evidence was also drawn from discussions with IRCK.

3.3 Methodological limitations and interpretation

EVIDENCE NOTE

Competency values are recorded as ‘not estimable from the qualitative baseline’.

SET AT TO

The primary component comprised 13 interviews and three IRCK interlocutor discussions, for a total of 16 participants/interlocutors. Participants were purposively selected for relevant knowledge and experience; they were not randomly sampled. The findings are therefore analytically useful but are not statistically representative of faith actors, young people, or residents across the six implementation counties. Primary interview evidence was obtained from Homa Bay, Mombasa, Lamu, Nairobi, and Busia. The Nakuru profile consequently relies on the TrustLab documentary record and should be validated through early county-level key-informant interviews and the standardized pre-training assessment before delivery of the Nakuru module. The interview guide elicited experiences, perceptions, and examples but did not administer a uniform scored knowledge and skills test. It is therefore not methodologically valid to convert the interview narratives into percentages for the project’s competency indicators. In the consolidated framework below, these values are recorded as ‘not estimable from the qualitative baseline.’

The project should establish the participant-level time-zero (T0) values at enrolment, before any training exposure, using the same instrument that will be administered at endline. Remote interviews, translation from Kiswahili to English where required, recall bias, social-desirability bias, and rapidly changing platform conditions may also affect interpretation. These limitations do not invalidate the observed themes, but they require careful triangulation with attendance records, practical assessments, incident logs, platform-context logs, and referral audits during implementation.

4.0

Baseline findings and county profiles



4.0 BASELINE FINDINGS AND COUNTY PROFILES

6

Mombasa · Lamu · Nakuru · Nairobi · Homa Bay · Busia

TRUSTLAB RECORDS

THE TRUSTLAB CYCLE

- Monitoring
- Validation
- Analysis
- Briefing
- Counter-messaging

4.1 Existing TrustLab platform as the immediate institutional baseline

The TrustLab experience is the strongest asset available to the proposed project. The initiative works in Mombasa, Lamu, Nakuru, Nairobi, Homa Bay, and Busia to strengthen grassroots responses to digital threats that may undermine democracy. This direct geographic overlap means that FaithConnect does not need to create a community entry architecture from scratch.

There are several specific assets already in place:

Regional dialogue infrastructure. Eighteen Barazas reached 726 people across the six counties. Of these, 367 were women, 317 men, 42 had no recorded gender, and 46 were persons with disabilities. The Barazas produced locally specific threat intelligence rather than only generic discussion.

Community intelligence infrastructure. TrustLab's Sentinels operate as locally embedded monitors, including in "dark social" spaces such as closed WhatsApp, Signal, and Telegram groups that cannot readily be observed through standard automated monitoring. Information is passed to researchers for verification and trend analysis.

Public-response infrastructure. Mythbusters operate as community-facing communicators who translate validated findings into contextual public narratives. The interim report describes a cycle of monitoring, validation, analysis, briefing, and counter-messaging.

Monitoring infrastructure. The TrustLab report planned 62 biweekly threat reports and 14 situation reports, with 10 biweekly reports and two situation reports completed at the time of reporting.

Learning resources. TrustLab developed an AI for Democracy Toolkit, Digital Security Toolkit, and Digital Self-Defence Playbook, with plans to make resources more widely accessible.

ACTORS capacity-building infrastructure. One incubator cohort had been completed; 20 staff from 10 organizations were recorded as trained, nine of the ten organizations submitted project proposals responding to threats identified in their communities, and eight ACTORS/CSO initiatives had been selected at that stage.

REGIONAL DIALOGUE
INFRASTRUCTURE

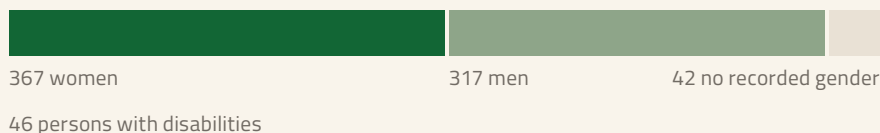
TRUSTLAB RECORDS

18

Barazas

726

people reached across the six counties



MONITORING INFRASTRUCTURE

Completed at the time of reporting

10

of 62 planned · biweekly threat reports



2

of 14 planned · situation reports

ACTORS CAPACITY-BUILDING
INFRASTRUCTURE

1

incubator cohort
completed

20

staff from 10
organizations
recorded as trained

9

of the ten
organizations
submitted project
proposals

8

ACTORS/CSO
initiatives selected
at that stage

Source: TrustLab Interim Narrative Report, as summarised in section 4.1 of this assessment.

4.0 BASELINE FINDINGS AND COUNTY PROFILES

INTERVIEWS

4.2 Recognition and verification of harmful digital content

Several different criteria of identifying harmful digital content were mentioned by respondents, including: looking at the personality of the source of the content to help associate it with a particular political wing, marking AI features of AI-generated harmful digital content, looking for different sources to confirm whether different sources agree and also reading through the comment section of the post. Additionally, most harmful digital content often focuses on negativity rather than spreading positivity.

Most respondents also deemed manipulative and false content to be difficult to classify because it deeply mirrors the truth to a large extent and requires a significant amount to verify. A religious leader from Homa Bay highlights that when false information floods social media, it is difficult to substantiate which one is true. In terms of the verification of the information before forwarding, most participants expressed concern that most harmful posts are reposted without much verification because those posts want some kind of validation and seek attention. Additionally, if a fake post favors a particular political stand that resonates with the consumer, there will be a higher chance that they will forward it without doing any verification to score a political objective. Some also just out of hatred for the government or given politicians, they forward this harmful information. As for those who condemn the harmful information, there was agreement among respondents that they condemn it because the context of the information can be verified as harmful to society. However, for both those who forward and condemn, the majority of the respondents agreed that they are in contact with somebody who influences their action, with those forwarding particularly being linked to either a politician's support or a religious leader's message. Those condemning might derive their act from personal conviction, but several participants averred that they are also linked to the opposite political side, religious organization, activist groups and other funding organizations aiming at accurate online political messaging. The medium of contact was mentioned several times as phone, but several participants also mentioned that there are instances of physical contact.

COUNTY FOCUS

3

interviews conducted in Busia County

INTERVIEWS

The three interviews conducted in Busia County reveal that digital harm is increasingly affecting community relations, civic participation, and social cohesion. Respondents identified misinformation and disinformation, online scams and fraud, cyberbullying, harassment, body shaming, and ethnically inflammatory content as the most prominent concerns. These harms circulate mainly through Facebook, WhatsApp, TikTok, and X, with young people and women identified as particularly vulnerable. Respondents associated young people's susceptibility with peer pressure, excitement, financial inducements, and failure to verify content before sharing it. Women and girls were reported to face distinctive risks, including sexual harassment, manipulated images, impersonation, sexist abuse, and attacks on their character and credibility.

Busia respondents generally become suspicious of content when it lacks a credible source, uses sensational or inflammatory language, contains misleading captions, or presents edited images, videos, and voice notes. Other warning signs include inconsistent dates or locations and the rapid circulation of unverified claims. Their verification practices include checking the original source, comparing information with trusted media, consulting community or religious leaders, and

using online searches and fact-checking tools. However, these practices are not yet systematic or uniformly understood.

4.3 ACTORS mapping

The supplied spreadsheet lists 28 prospective actors. The distribution is:

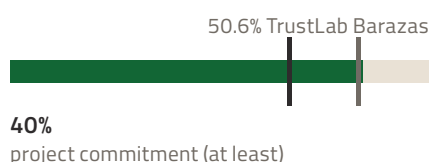
Table 2: Actors Distribution

COUNTY	NUMBER OF ACTORS	SHARE OF LIST
Busia	8	28.6%
Nairobi	8	28.6%
Homa Bay	4	14.3%
Lamu	4	14.3%
Mombasa	2	7.1%
Nakuru	2	7.1%
TOTAL	28	100%

This list should be treated as an interview and mobilization frame, not as a representative sample of civic organizations in each county. The eight-to-two disparity means that proportional selection would over-weight Nairobi and Busia. At least two ACTORS interviews per county are preferable, with both listed Mombasa and Nakuru organizations included because of the smaller frames.

4.4 Gender and inclusion baseline

FEMALE PARTICIPATION



Parity (50%) suggested as the evidence-informed target

TRUSTLAB RECORDS

INTERVIEWS

The new project commits to at least 40% female participation. The previous TrustLab Barazas already achieved approximately 50.6% women, suggesting that a 40% threshold may be unnecessarily conservative if the same community networks are mobilized. Also, several participants point out that a more ambitious, evidence-informed target would be approximately gender parity, while retaining flexibility where particular religious leadership structures create participation constraints. This is especially justified because the project explicitly addresses gender-based online violence and links women's peacebuilding role to the Women, Peace, and Security agenda.

Other than gender, youths should become a great priority. There was 100% agreement among participants in all the counties of interest that the most affected group by digital harm are the youth. This is because youths are both the most participants in online digital activities and the most consumers of the online political

activities' outputs. Moreover, the youths are gullible to political manipulation by both national and county leaders. A few participants also noted that youths face economic exclusion, including underemployment and unemployment, and have turned to digital harm as a source of economic revenue, where they post false information in exchange for a few coins. Therefore, those youths who do content creation remain highly capable of enhancing digital harm compared to youths who just share the content.

INTERVIEWS

However, some participants noted that those who consume it may fall under peer pressure of having many followers, hence spreading false information to get more followers, more views, and attention. Yet the conversation with IRCK reveals that youth inclusion in inter-religious duties has been difficult, with IRCK interlocutors already noticing a constant feeling among youths that religious leaders are with the state. In fact, interlocutors from IRCK highlighted that intergenerational dialogue has been challenging because the frustration, concern, and language of youths have not been fully captured. The language of youths and religious leaders differs. Difference in language leads to difference in approaches, thus the need for intergenerational dialogue.

4.5 Demographic difference and overlaps in faith-based digital harm counter

Generally, across the participants, whether in Lamu and Mombasa (Coastal Region), Homa Bay (Western region) or Nairobi region, there is an overlap in the recognition of the significance of the role of faith leaders and interfaith actors. This is because across all the counties, different demographics, including youths, are more likely to identify with a particular religion, whether Islam or Christianity. This is why they are placing much emphasis on religious leaders to install more online networks to reach their congregants on social cohesion ahead of the 2027 election.

4.6 Region

Despite the general emphasis of the faith-based role in countering digital harm by participants, regional analysis exposes the difference in the extent to which faith-based leaders are involved in civic education and peaceful preaching online. Coastal counties, including Lamu and Mombasa, experience extensive online religious engagement compared to Western counties, particularly Homa Bay. Interviewed religious leaders from Lamu and Mombasa averred that aside from teaching in the mosque, they have engaged not only in radio teaching but also in YouTube teaching on peaceful civic participation. In Homa Bay, on the other hand, the interviewed religious leader argued that online engagement of religious leaders in countering digital harm, for example, the incitement statement by Homa Bay Town MP, Opondo Kaluma, directing goons to attack the Linda Mwanachi group, is something to still think about. The difference is the dominance of Islam in coastal areas, creating a homogeneous audience, while Homa Bay County experiences various informal and formal churches, some of which do not even have online platforms. A few participants also associate this strong online religious engagement in civic education in the coastal region with an existing framework

of countering extremism in the region. Notably, participants from Nairobi remained concerned that, regardless of most churches' capability to install an online framework of countering online political harm, the way they package their narratives frequently increased political animosity. They frequently endorse politicians and hence remain non-neutral in most cases.

4.7 Gender

While there is recognition that women experience numerous digital harms across several participants, there is limited attention to factors that encourage this. One particular factor across participants is the difference in societal approach to women and men. Consequently, the religious leader's role in condemning or preventing frequent attacks on women receives less attention due to minimal action.

Notably, though, a few respondents highlighted that because most of the time men are the ones in the limelight and out and about, and women are often at home, men are more exposed to digital harm compared to women. However, the quality and severity of gendered harms targeting women and girls—including sexualized attacks, manipulated images, impersonation, public shaming, and attacks on credibility—creates distinct vulnerabilities that merit gender-specific curriculum and safeguarding support.

Table 4: Demographic Summary of Respondents and Interlocutors

	REGION/COUNTIES	GENDER	AGE GROUP
2	Western (Homa Bay)	M & F	Youth & Elder
6	Coastal (Lamu & Mombasa)	F: 2, M: 4	Elders: 4, Youth: 2
2	Nairobi	M & F	Youth
2	Nakuru	M	Youth; Adult
3	IRCK (Nairobi)	M: 2, F: 1	Elders
3	Western (Busia)	M: 1, F: 2	Adults/Elders
TOTAL: 16		M: 9, F: 7	Elders: 11, Youth: 5

Note: Based on interviews. Gender distribution is 56.2% male, 43.8% female. Age distribution is 68.8% elders/adults, 31.2% youth. Geographic distribution spans Western, Coastal, and Central/Nairobi regions. Faith traditions represented include Christian, Islamic, and Interfaith actors. The baseline does not include respondents from Nakuru county and therefore relies on TrustLab documentary records for that profile.

4.0 BASELINE FINDINGS AND COUNTY PROFILES

INTERVIEWS

Figure 4: Baseline Respondent Demographics (n=18)

Distribution Across Regions, Gender, Age, and Faith Traditions

Gender Distribution (n=18)



■ Male 61.1%
■ Female 38.9%

Age Distribution (n=18)



■ Elders/Adults 61.1%
■ Youth 38.9%

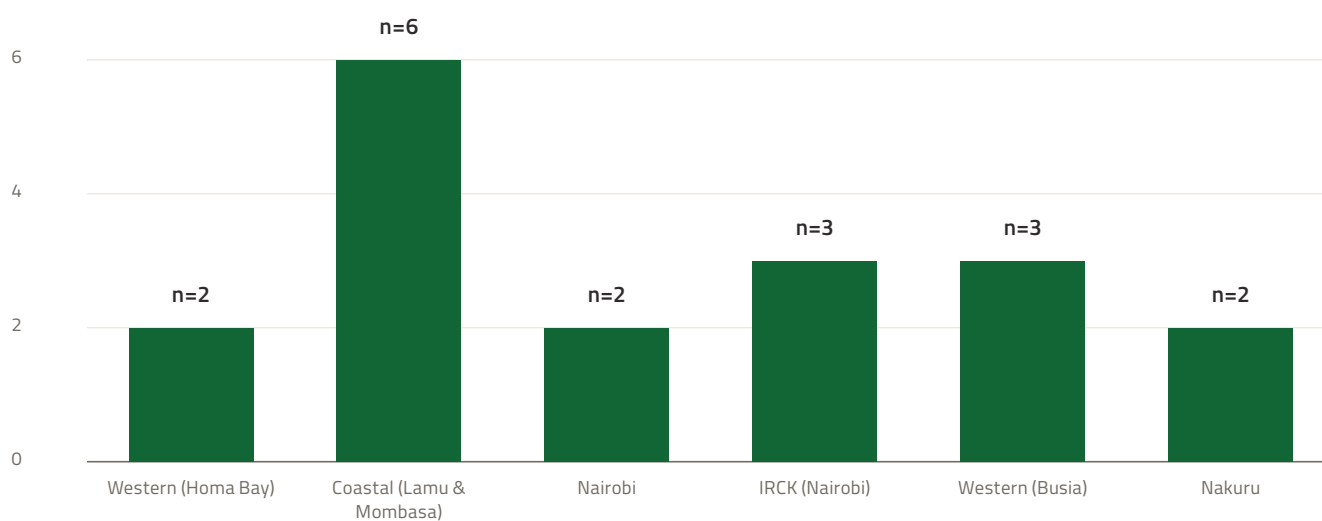
Faith Tradition Representation (n=16)*



■ Christian 50.0%
■ Islamic 31.2%
■ Interfaith 18.8%

Regional/County Distribution of Respondents

Number of Respondents



*Faith tradition not yet recorded for the 2 new Nakuru respondents; Faith Tradition chart remains based on n=16 pending follow-up.

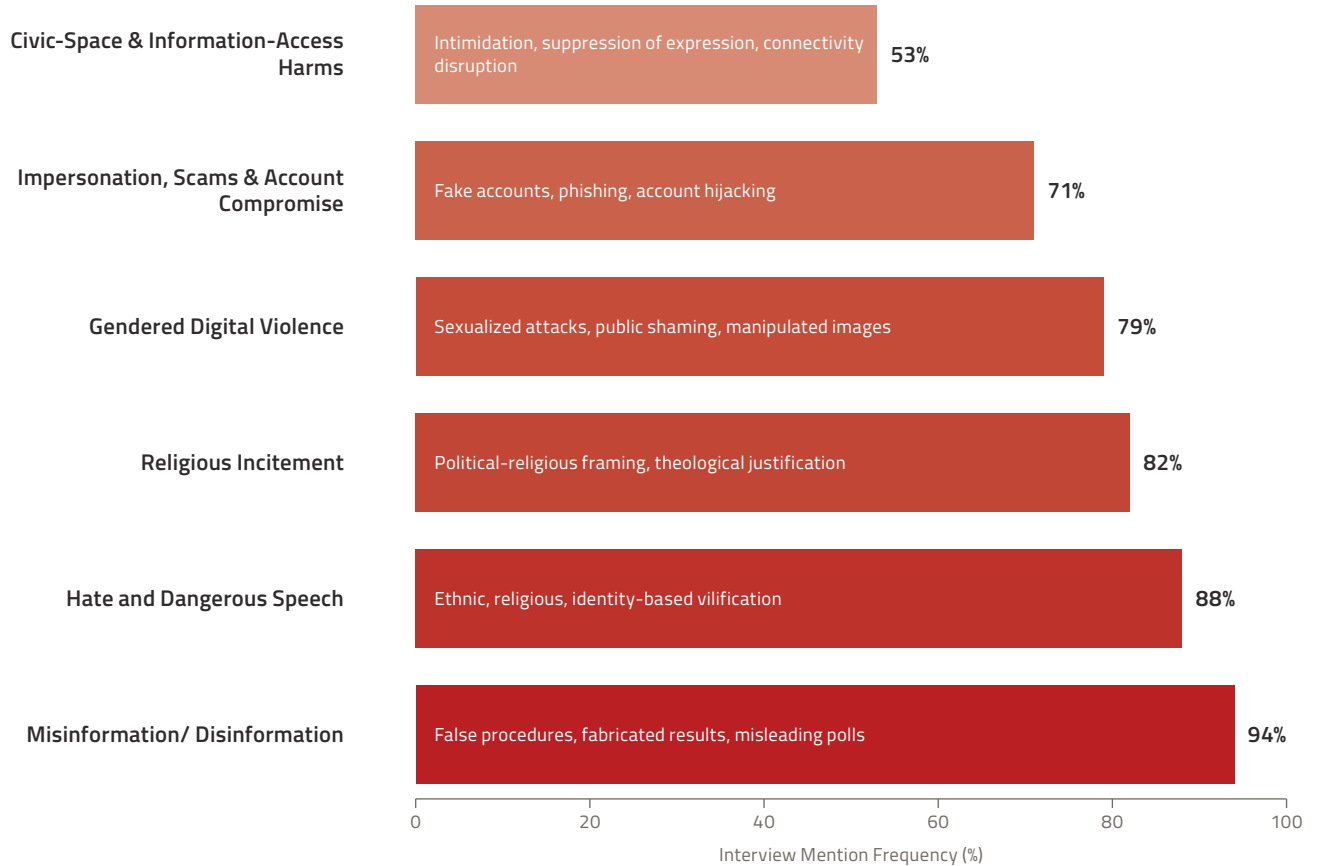
The demographic visualization shows the distribution of baseline respondents across regions, gender, and age groups. The sample (n=18) includes 15 primary interview respondents plus 3 IRCK interlocutors. Gender distribution is 56.2% male (n=9) and 43.8% female (n=7), slightly more male-skewed than TrustLab's achieved 50.6% female participation in Barazas. Age distribution shows 68.8% elders/adults (n=11) versus 31.2% youth (n=5), reflecting both religious leadership structures and youth underrepresentation in formal interviews. Both Christian and Islamic faith traditions are represented, with interfaith actors also included.

4.0 BASELINE FINDINGS AND COUNTY PROFILES

INTERVIEWS

Figure 5: Six Categories of Harm - Interview Frequency

Relative Prominence in Baseline Interviews (n=16)



This horizontal bar chart shows the relative frequency with which each harm category was mentioned by respondents. Misinformation/disinformation and hate speech emerged as the most frequently identified harms, followed closely by religious incitement and gendered digital violence. Impersonation and scams were also commonly cited, while civic-space harms—though serious—were mentioned less frequently in baseline interviews.

4.0 BASELINE FINDINGS AND COUNTY PROFILES

4.8 Baseline indicator framework

The baseline assessment has been designed to establish concrete baseline values for each project indicator and connect them to the specific targets outlined in the approved Federal Foreign Office funding proposal. This framework ensures that the project can track both headline proposal outputs and deeper participant-level competency improvements. All proposed indicators have been incorporated into a comprehensive baseline matrix with specific baseline values (August-September 2026) and concrete end-of-project targets.

PROJECT-LEVEL CORE INDICATORS:

1

Six Barazas conducted (at least 25 participants each)

Baseline: TrustLab operational with 18 Barazas reaching 726 people; FaithConnect-specific Barazas not yet activated.

Target: 6 regional Barazas with minimum 25 participants each.

2

300+ participants trained (with at least 40% women)

Baseline: 0 FaithConnect participants trained.

Target: 300-500 direct beneficiaries including faith leaders, youth peace champions, interfaith actors; minimum 40% women.

3

5+ trained Sentinels and 5+ Mythbusters active over 6 months

Baseline: TrustLab model operational; FaithConnect-specific roster not established.

Target: At least 5 trained Sentinels and 5 Mythbusters active throughout project period.

4

30+ counter-narrative outputs produced

Baseline: FaithConnect-specific baseline = 0.

Target: Minimum 30 counter-narrative outputs (posts, videos, audio, written materials) meeting quality standards.

5

70% of participants show improved knowledge (pre/post surveys)

Baseline: Not estimable from qualitative baseline.

Target: 70% of trained participants demonstrate improved knowledge as measured by standardized assessments.

SET AT TO

6

Documented cases of harmful narratives identified and addressed

Baseline: TrustLab documented cases; FaithConnect register not established.

Target: Documented cases tracked with clear evidence of identification, verification, and response.

7

At least 40% women participation (with 50% as aspiration)

Baseline: TrustLab achieved 50.6% women; baseline interviews included 43.8% women.

Target: Minimum 40% in all activities; aspire to 50% parity.

Complementary Indicators: The baseline matrix also includes complementary indicators supporting core targets and providing deeper insight into competency changes, including participant competency in identifying misinformation (75% improvement target), contextual classification of hate speech and religious incitement (70% competency), verification workflow completion (70% proficiency), safe response confidence (25 percentage point increase), referral pathway knowledge and trust (75% can identify channels), gender-specific violence response capability, inter-generational dialogue effectiveness, and continuous platform monitoring for external disruptions.

Dashed outline: baseline to be measured at enrolment (TO).

4.0 BASELINE FINDINGS AND COUNTY PROFILES

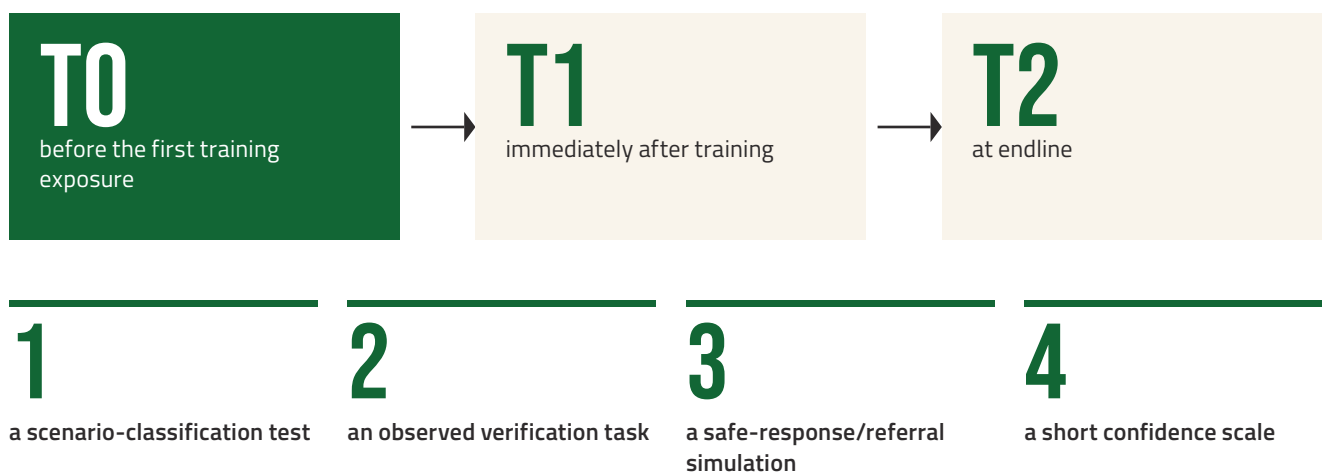
SET AT T0

4.9 Measurement protocol for demonstrated skills

Each participant should receive an anonymous unique identifier. The same four-part instrument should be administered before the first training exposure (T0), immediately after training (T1), and at endline (T2): a scenario-classification test, an observed verification task, a safe-response/referral simulation, and a short confidence scale. Assessors should use a standardized rubric and record item-level scores.

The M&E dataset should report the percentage reaching competency, mean score change, retention of skills between T1 and T2, and results disaggregated by sex, age group, county, actor category, disability status where safely collected, and prior training exposure. This protocol ensures that the project can measure both confidence-level changes and genuine competency improvements in core areas: misinformation identification, contextual classification of hate speech and religious incitement, verification workflow completion, safe response selection, and referral pathway knowledge.

THE SAME FOUR-PART INSTRUMENT, THREE TIMES



RESULTS DISAGGREGATED BY

sex · age group · county · actor category · disability status where safely collected · prior training exposure

4.10 RECOMMENDATIONS

5.0

Recommendations



4.10 RECOMMENDATIONS

The interviews indicate that faith and interfaith actors respond less consistently to digital harm than youth peace champions. Respondents associated this difference with limited familiarity with rapidly changing platforms, infrequent training, unclear referral channels, fear of retaliation, and perceptions that responsible authorities do not always respond effectively. IRCK interlocutors also emphasized stronger training, cross-level networking, and improved recognition of AI-enabled manipulation.

Some faith leaders and youth peace champions had previously received peace-building or digital-harm training and reported applying elements of that learning. Coverage, recency, and depth were uneven. These findings point to sustained skills development, joint action between faith and youth actors, clear referral protocols, and periodic refresher support rather than a single training event.

The findings and measurement requirements above translate into the following programmatic shifts for implementation:

01

Build FaithConnect on the existing TrustLab architecture.

Siasa Place should not establish a completely separate monitoring and counter-narrative mechanism. FaithConnect should formally link faith actors to existing Sentinel/Mythbuster workflows, threat-monitoring experience, Barazas, and tool-kits. This lowers start-up costs, avoids parallel reporting chains, and provides the new project with a tested route from local observation to verification and public response.

02

Measure demonstrated skill, not only self-confidence.

Self-reported confidence can diverge from actual performance. Each training cycle should capture four complementary measures: (i) a short scenario-based classification test; (ii) a practical verification exercise completed with an assessor checklist; (iii) a five-point self-efficacy scale; and (iv) a response simulation requiring the participant to select a safe action or referral pathway. The pre-training, post-training, and endline versions should be equivalent in difficulty, scored using the same rubric, and linked through anonymous participant identifiers. The primary project outcome should be the percentage demonstrating competent identification and verification, supplemented by the change in confidence.

03

Reframe the “70% improved knowledge” indicator.

Rather than counting anybody whose score rises by any amount, Siasa Place should define a competency threshold. A stronger formulation would be: “By endline, at least 70% of matched participants achieve a competent score of at least 70% on the standardized Digital Harm and Peaceful Civic Participation Knowledge and Skills Test, and the mean cohort score improves by at least 20 percentage points from baseline.”

04

Do not assume faith actors begin from zero.

At enrolment, the project should record each participant’s prior exposure to

05

UNESCO/FeCoMo, IRCK, NCCK, Siasa Place/TrustLab, NCIC, digital-security, fact-checking, election-peace, or media-and-information-literacy initiatives. This variable should be used to interpret baseline scores and segment learning results, because existing evidence shows that some faith actors have already received relevant training.

Design Barazas as intergenerational learning spaces.

Barazas should pair faith leaders and young people in reciprocal learning roles. Young participants can contribute platform knowledge, emerging language, and peer-network insight, while faith leaders contribute ethical framing, trusted community access, mediation experience, and offline reach. Facilitation plans should reserve speaking and leadership roles for both groups and should measure perceived mutual trust at baseline and endline.

06

COUNTY MODULES

Nakuru · Lamu · Mombasa · Homa Bay · Busia · Nairobi

to be verified through county KIIs

Adapt content to county-specific harm profiles.

The baseline identifies meaningful differences in the six information environments. FaithConnect should include a common national curriculum and tailored county modules. Nakuru: ethnic incitement, coded political language, and historical conflict sensitivity. Lamu: phishing/scams, closed messaging environments, online radicalization concerns, religious/community sensitivity. Mombasa: gendered digital abuse, youth engagement, interfaith and coastal political narratives. Homa Bay: gendered political abuse, youth political mobilization, non-partisan civic engagement. Busia: impersonation and account compromise, scams and online fraud, political and ethnic misinformation, cyberbullying, body shaming, gendered digital abuse, youth vulnerability, and sustainable community-based referral and response mechanisms. Nairobi: high-volume information flows, rapidly amplified narratives, civic-mobilization and freedom-of-expression challenges. These profiles should be verified through county KIIs rather than treated as exhaustive descriptions.

07

Make freedom of expression a core competency.

Training should explicitly teach that harmful content cannot be identified through keyword lists alone. The Kenyan “tugeges” case shows how politically contested terminology can be misclassified when linguistic context and evolving usage are ignored.

08

Strengthen gender targets and safety practices.

Because previous TrustLab Barazas achieved approximately 50.6% female attendance, FaithConnect should aim for parity, rather than treating 40% as the desired level. The 40% proposal threshold can remain the absolute minimum. Women should also be represented among Sentinels, Mythbusters, facilitators, and decision-making focal points. Training should contain a dedicated module on technology-facilitated gender-based violence, safe evidence handling, doxxing, impersonation, sexualized disinformation, and referral options. This responds both to local TrustLab cases and to national gender disparities in digital use.

09

Use faith leaders as an offline-online bridge.

Because Internet use is substantial but uneven, particularly outside highly connected areas, online counter-messaging should be reinforced through sermons, congregational announcements, youth groups, women’s fellowships, mosque/church community forums, local radio, Barazas, and direct civic education.

10

National ICT data demonstrate that access remains unequal even as young people are considerably more connected than the population as a whole. A harmful narrative originating online can spread through offline conversation; the reverse is also true. Faith actors are particularly valuable because they can intervene across both spaces.

Protect the independence and credibility of sacred spaces.

Faith actors should sign or adopt a clear non-partisanship standard for project activities. Recent faith-sector discussion has highlighted concerns about the politicization of religious spaces and the effects of money and religious influence on democratic participation. The project should make clear that peaceful civic participation includes the freedom to support or oppose political ideas and candidates lawfully, while the project itself neither directs political choices nor permits places of worship to become channels for project-sponsored partisan messaging.

11

Protect Sentinel/Mythbuster sustainability and volunteer retention.

Before activation, Siasa Place should reconcile the original TrustLab roster and calculate cohort retention as: active original volunteers divided by volunteers originally recruited. The original-cohort attrition rate cannot be calculated from the records supplied for this baseline because a complete roster with exit dates and reasons was not available. This is therefore a specific data-acquisition item to be closed through the roster audit before or at project activation. The audit should record current status, motivation, expected time commitment, preferred recognition or non-financial support, safeguarding needs, and reasons for inactivity. Monthly logs should capture attendance, assignments completed, response time, supervision contacts, wellbeing concerns, and exits. Endline interpretation should distinguish training performance from infrastructure loss caused by attrition, burnout, connectivity costs, or inadequate support.

12

Maintain a platform and external-context log.

At project start and monthly thereafter, the M&E team should record platform accessibility, major moderation or policy changes, known algorithmic shifts, account restrictions or compromise, connectivity disruptions, and relevant legal or regulatory developments. The June 2025 Telegram disruption illustrates why changes in incident volume, reach, or response rates cannot automatically be attributed to FaithConnect. For every material disruption, the log should record affected counties, platforms, dates, duration, data gaps, and any adjustment made to outcome analysis. Contingency indicators should include days of monitoring lost, accounts compromised, and county-platform cells with incomplete data.

13

Audit referral pathway functionality.

Knowledge of a referral channel is not evidence that the pathway works. Before implementation, the project should map authorized contacts and service standards for the police/DCI, NCIC, social-media reporting mechanisms, faith institutions, and relevant protection or legal-support providers. Any test case must be ethical, non-deceptive, authorized, and must not manufacture harmful content or misuse emergency resources. Referral performance should be reported at three separate levels: received (acknowledged and logged), responded to (feedback or guidance provided within the agreed period), and acted upon (a documented protective, investigative, moderation, mediation, or support action). Participant

14

trust in each pathway should be measured separately on a five-point scale, alongside stated reasons for trust or reluctance.

Establish a continuous baseline-to-endline learning loop.

The six-month project should not treat the baseline as a report that is filed away after inception. It should create a living evidence cycle that connects baseline findings to implementation, monitoring, and endline assessment. Quarterly reflections should review emerging patterns, adapt county modules as needed, and ensure that endline indicators are grounded in baseline values and tracked consistently throughout implementation.

STAKEHOLDER PROPOSAL

Siasa Place should update the training curriculum to cover synthetic-media verification, responsible account management, civic and constitutional rights, and safe responses to political interference. County faith networks and youth co-facilitators should identify the platforms and meeting places young people actually use, run joint dialogues, and offer offline support and referrals.

TrustLab and partner agencies should document and test escalation and feedback procedures, including privacy and accountability safeguards.

The monitoring team should track who leads activities, not just who attends, and review progress with participants at regular intervals.

The background of the page is a solid dark gray. In the top right corner, there are two parallel diagonal stripes, one red and one green, extending from the top edge towards the center. Similarly, in the bottom left corner, there are two parallel diagonal stripes, one red and one green, extending from the bottom edge towards the center.

FaithConnect Initiative

Baseline Assessment

September 2026