Citation Localisation in Generative AI Answers: Whose Evidence Is Visible in African Humanitarian Crises?

Research Article

Hezron Ochiel

Affiliation: Hezron Insights, Nairobi, Kenya
Corresponding author: Hezron Ochiel
Email: info@hezroninsights.com or hezronochiel2021@gmail.com

Abstract

Generative AI systems increasingly answer humanitarian questions by synthesising web sources, yet little is known about whose knowledge becomes visible. This study introduces citation localisation: the degree to which an AI-generated humanitarian answer visibly attributes evidence to producers based in the affected country. The first 100-question Batch 1 of the African Humanitarian AI Citation Index covered ten African countries, ten themes and ten question types. Ten questions were piloted and froze the protocol; 90 production questions were then submitted independently to Gemini, DeepSeek, ChatGPT and Perplexity. Source-level citation status was recoverable for 255 of 360 observations; 186 cited observations yielded 1,083 verifiable answer-linked citations. In this recoverable, citation-weighted corpus, country-local producers supplied 5.5 per cent, African-based producers 6.6 per cent, and UN/multilateral organisations 68.8 per cent; the top five producers captured 40.3 per cent. A complete 90-question ChatGPT sensitivity analysis found 7.2 per cent country-local citations (question-cluster bootstrap 95 per cent interval 4.2–10.7 per cent). Local sources appeared in 33 cited observations, and 84.8 per cent also included UN/multilateral evidence. The findings identify a citation hierarchy shaped by representation, concentration, intermediation and supplementation: local evidence is rarely visible and usually diversifies, rather than replaces, established international authority.

Keywords: Generative AI; humanitarian information; localisation; Africa; citation analysis; epistemic justice

1 Introduction

Humanitarian decisions depend on information: where people are displaced, which services have failed, who faces the greatest protection risk, whether food security is deteriorating and which actors can reach communities. Increasingly, some people seeking that information will not begin with a humanitarian dashboard, a country cluster page, or a search engine results page. They will ask a generative artificial intelligence system and receive a synthesised answer, often accompanied by citations.

That shift changes the politics of visibility. Traditional search distributes attention through ranked links that users can open and compare. Generative systems instead select a smaller set of sources and present a synthesised narrative. Citation therefore becomes more than a technical reference mechanism: it determines which institutions are visible as credible knowledge producers. Research on generative search increasingly treats this visibility as measurable, showing that AI citations can diverge from conventional rankings and concentrate attention on a relatively narrow set of domains (Aggarwal et al., 2024; Sünkler et al., 2026; Xu, Iqbal and Montgomery, 2026).

For African humanitarian crises, source visibility matters especially. Humanitarian reform has spent a decade debating localisation: shifting resources, leadership and recognition towards national and local responders. The 2016 Grand Bargain committed major actors to increase support for local and national responders and set a target of channelling at least 25 per cent of humanitarian funding to them as directly as possible (Inter-Agency Standing Committee, 2016). Progress remains limited: Pearson and Rieger (2026) estimate that below 10 per cent of international humanitarian funding reached local and national actors directly or indirectly in 2024, with direct funding at 3.8 per cent.

However, localisation debates have focused mainly on funding, partnerships, coordination and operational leadership. Generative AI introduces another form of authority: discoverability. A locally legitimate organisation may be central to a crisis response yet largely absent from the sources AI systems use to construct answers. At the same time, institutions with stronger publishing and syndication infrastructure may become disproportionately visible.

This study calls that problem citation localisation: the degree to which an AI-generated answer about a humanitarian situation visibly attributes supporting evidence to knowledge producers based in the affected country. The measure is deliberately narrower than humanitarian localisation as a whole. It concerns visible attribution when an AI system presents evidence to a user, not who funded, delivered, or collected the underlying information.

The paper reports the production analysis from the first 100-question randomised Batch 1 of the African Humanitarian AI Citation Index (AHACI), a structured study of source selection across ten African countries. The 100-question batch comprises a ten-question protocol pilot and a 90-question production sample; all citation percentages reported as primary findings derive from the production sample. It asks: what share of visible citations comes from country-local and African-based producers; how concentrated is citation attention; how does local visibility vary by platform, country, humanitarian theme and question type; when local evidence appears, does it replace or supplement international sources; and how does intermediary hosting affect attribution between URL and original producer?

The study does not claim to measure algorithmic bias against African sources. That would require a defensible denominator of relevant, available and retrievable sources for each query. AHACI instead audits what the systems visibly cited under a standardised protocol. The contribution is a mapped citation hierarchy rather than an inference about hidden model intent: an empirical account of representation, concentration, hosting intermediation and the position of local evidence within AI-generated source stacks.

2 Generative Search, Humanitarian Localisation and Epistemic Visibility

Generative answer systems combine information retrieval and language generation. They may search the live web, retrieve documents, rerank or filter them, allocate limited context and then generate a response citing only some material. Users see the end of that pipeline rather than its intermediate decisions. Venkit et al. (2024) show that source-cited answers can still contain hallucinations and citation errors, while Kakimov et al. (2026) argue for auditing retrieval and citation as observable processes rather than assuming that a cited answer is inherently transparent.

Early work on AI citation behaviour asked whether references existed and were accurate. Walters and Wilder (2023) found substantial fabrication and citation errors in bibliographic references generated by GPT-3.5 and GPT-4, although the newer model performed better. Mugaanyi et al. (2024) likewise found disciplinary variation in citation and DOI reliability. These studies establish an important distinction: citation validity and source selection are separate research problems. A citation can be accurate while the overall evidence ecosystem remains highly concentrated or geographically narrow.

Source selection has become more visible as generative search has expanded. Sünkler et al. (2026), auditing Google AI Overviews for health queries, identify substantial source concentration and limited overlap with organic search. Xu, Iqbal and Montgomery (2026) similarly find that many AI Overview citations are absent from co-displayed first-page results. Aggarwal et al. (2024) further show that content structure can affect visibility in generative answers. Together, these studies frame citation as a selective allocation of attention.

This selectivity matters in humanitarian settings because knowledge production already reflects institutional inequality. International organisations typically have dedicated information-management teams, standard formats, stable domains and global syndication networks. Local organisations may have richer contextual knowledge but publish less frequently, use less indexable formats, or distribute evidence through channels that generative systems retrieve less reliably. Localisation scholarship shows that proximity to a crisis does not automatically translate into institutional power (Barter and Sumlut, 2023; Robinson, 2026).

The financial dimension reveals a related hierarchy. The Grand Bargain sought to redirect humanitarian resources towards local and national actors (Inter-Agency Standing Committee, 2016), yet large gaps remain between commitment and practice (Pearson and Rieger, 2026). Earlier work in Somalia and South Sudan similarly documented persistent difficulties in tracking and increasing funding to local actors (Majid et al., 2018). Digital knowledge infrastructures may reproduce related authority patterns, although that mechanism must be measured independently.

Humanitarian information management is itself a form of institutional ordering. Situation reports, needs assessments, cluster products, and funding appeals do more than describe crises: they establish categories, baselines, and institutional accounts that circulate across donors, agencies, media, and policy actors. Organisations already embedded in global humanitarian information infrastructure therefore hold advantages in publication regularity, metadata, syndication and recognisable institutional authority. When generative AI draws mainly from that layer, it can reproduce the distributional advantages of actors that are already easiest to discover. Citation localisation consequently treats discoverability and attribution as part of localisation's power question, alongside resources, representation and decision-making (Barter and Sumlut, 2023; Robinson, 2026).

Critical AI scholarship provides a vocabulary for examining that possibility without reducing it to a technical defect. Mohamed, Png and Isaac (2020) call for decolonial scrutiny of AI systems and the historical power relations in which they operate. Birhane (2020, 2021) argues that technological dependence and algorithmic injustice are relational phenomena embedded in social structures. Couldry and Mejias (2019) similarly locate datafication within broader processes of extraction and control. These perspectives do not prove that any particular citation is unjust, but they make source visibility a legitimate empirical object.

AHACI therefore treats citation localisation as an observable layer of epistemic power. It relates to Fricker's (2007) account of epistemic injustice, especially unequal credibility among knowers, but operationalises it more modestly. The study asks who is visibly cited, not whether a model internally assigns credibility based on identity. This distinction avoids attributing unobservable motives while recognising that repeated attribution patterns shape which knowledge users encounter.

The analytical framework has four layers. The first is representation: the share of citations attributed to African-based and country-local producers. The second is concentration: how much citation attention a small set of organisations captures. The third is intermediation: whether the URL host is the same organisation that produced the underlying knowledge. The fourth is source-stack substitution: whether local evidence displaces dominant international sources or merely appears alongside them. These layers move the analysis beyond a binary local-versus-international count. A citation ecosystem can contain some local sources and

remain highly concentrated, mediated through international hosting infrastructure, or dependent on international organisations for most evidentiary weight.

3 Methods

AHACI is a structured observational audit of source selection in generative AI answers to African humanitarian questions. Its 1,000-prompt question bank crosses ten countries, ten humanitarian themes and ten question types. Countries are Burkina Faso, Chad, the Democratic Republic of the Congo (DRC), Ethiopia, Kenya, Mozambique, Nigeria, Somalia, South Sudan and Sudan. Themes cover conflict and protection; displacement; food insecurity; health emergencies; WASH; climate hazards; humanitarian access; protection of women and children; livelihoods, shelter and basic services; and humanitarian response, funding and localisation. Question types range from current situation and drivers to hotspots, unmet needs, response, barriers, evidence of change and outlook.

Batch 1 contained 100 independently randomised questions (Runs 001-100), selected from the 1,000-prompt bank with seed 20260823. The first ten questions served as a protocol pilot used to refine and freeze the collection procedure and were excluded from the production analysis. The production sample therefore comprises Runs 011-100: 90 questions collected under the final protocol. Because the factorial bank was randomly sampled rather than quota-balanced, country, theme and question-type counts vary; the analysis reports the realised sample rather than implying equal representation.

Each production question was submitted to the consumer-facing Gemini, DeepSeek, ChatGPT and Perplexity interfaces on 23 August 2026. The protocol required one question per fresh conversation, the first successful answer only and normal web-enabled behaviour where available. A frozen suffix instructed each system to answer independently, use current reliable information, cite relied-upon sources with direct links where possible, select sources naturally without favouring geography or organisation type, and treat each query as standalone. Product names, collection date, fresh-chat conditions and visible web-enabled behaviour were recorded. Consumer interfaces did not consistently expose exact underlying model builds, retrieval partners, ranking parameters, or backend versions, so we could not independently standardise them; platform labels in this paper therefore refer to the observed consumer products on the collection date rather than stable model families.

An early platform constraint required replacing Claude with DeepSeek. Runs 011-015 were rerun so the production sample used the same four-platform set. Batched attempts that handled multiple questions from shared retrieval context were preserved as protocol deviations but excluded, because a failed batched search is not equivalent to an independently generated zero-citation response.

The study distinguishes three units. A question is one humanitarian prompt. An observation is one platform's response to one question. A citation is one distinct answer-linked source endpoint. Batch 1 contains 100 questions overall: 10 pilot questions and 90 production questions. The analyses reported in this paper use the 90-question production sample, generating 360 platform-question observations. Citation-level coding was performed only for URLs visibly attributable to the original answer. Named organisations without URLs were not converted into citations, and missing links were never reconstructed after collection. Duplicate URLs within the same answer were deduplicated at the endpoint level after removing tracking parameters where the underlying resource was otherwise identical.

Perplexity presented a special attribution problem because its interface could surface sources that were not actually referenced in the narrative answer. AHACI separates answer-linked citations from surfaced-only sources. The primary analyses in this paper use answer-linked citations only. Twenty-five surfaced-only citations were retained in the audit workbook but excluded from source-composition calculations. This

choice prevents the interface's wider retrieval or display pool from being conflated with the sources that visibly supported the answer.

A second coding distinction separates URL host from original knowledge producer. Repositories, syndicated mirrors and asset hosts can expose a document without creating its evidence. For each citation, AHACI records host domain, identifiable publisher or producer, publisher type and geographic base, whether the producer is African-based or country-local, UN/multilateral status, and whether the host is direct or intermediary. Country-local means institutionally based in the country discussed; African-based also includes regional African organisations. UN country offices remain multilateral because classification concerns institutional governance rather than field-office location.

Publisher types include UN/multilateral organisations, universities and research institutes, international NGOs and media, donor or bilateral institutions, African national and regional media, national public agencies, local or national NGOs/CSOs, data platforms and other organisations. Classification confidence is stored separately from source-quality judgements. Of 1,083 answer-linked citations, 1,023 have high-confidence source classification, 59 medium and one low. Claim support and source-quality coding remain incomplete and are therefore outside this paper's scope.

Citation classification was conducted by a single analyst using the predefined codebook, with classification-confidence flags used to identify ambiguous cases. These confidence labels are quality-control indicators, not inter-coder reliability statistics; no independent second-coder assessment was performed in this wave. To make the classifications auditable, an anonymised reproducibility package accompanies the manuscript and contains the Batch 1 question list, frozen production prompt structure, coding framework, citation-ledger workbook and analysis workbook. Future validation should independently double-code a random subset and report agreement for country-local status, African-based status, publisher type, host role and original producer.

Source-level recovery is incomplete for 105 of 360 production observations. These are archival capture gaps, not evidence that the original responses lacked citations. The recoverable set contains 255 observations with resolved citation status: 186 with at least one answer-linked citation and 69 valid zero-URL answers, predominantly from DeepSeek. The 186 cited observations contain 1,083 answer-linked citations. Source-level capture is available for 19 Gemini observations, 80 DeepSeek observations (11 cited and 69 zero-URL), all 90 ChatGPT production observations and 66 Perplexity observations. Because missingness is platform-clustered rather than random, platform comparisons are descriptive; ChatGPT's complete 90-question production archive is used as a sensitivity analysis.

The pooled 1,083-citation corpus is therefore citation-weighted rather than platform-balanced: platforms that exposed more recoverable answer-linked URLs contribute more observations to citation-level percentages. ChatGPT and Perplexity together account for most recoverable citations. Pooled cross-platform percentages should consequently be read as the composition of the recoverable citation corpus, not as equal-weight estimates of the four products. The complete ChatGPT archive provides the principal platform-complete sensitivity check for the representation findings.

Citation concentration is measured with top-five, top-ten and top-twenty concentration ratios and the Herfindahl-Hirschman Index (HHI); inverse HHI is reported as an effective number of equally weighted publishers. Observation-level source-stack analysis compares cited answers containing at least one country-local source with those containing none, using total citation count, UN/multilateral count and within-answer UN share. These comparisons are descriptive: a lower UN share in local-citing answers is treated as a diversification or partial-substitution signal, not proof of causal replacement. As a post hoc sampling sensitivity for ChatGPT, the only platform with complete production source capture, 20,000 cluster-bootstrap

replicates resampled the 90 production question-level citation stacks with replacement using seed 20260823. This preserves within-answer citation clustering and quantifies prompt-sampling uncertainty without addressing model drift or unobserved retrieval.

Table 1. AHACI production sample and citation-recovery status

Component	N	Notes
Question bank	1,000 prompts	10 countries x 10 themes x 10 question types
Batch 1 randomised sample	100 questions	Runs 001-100; randomisation seed 20260823
Protocol pilot	10 questions	Runs 001-010; used to refine/freeze protocol; excluded from production analysis
Production analysis sample	90 questions	Runs 011-100; final frozen protocol
Platforms	4	Gemini, DeepSeek, ChatGPT, Perplexity
Platform-question observations	360	One first-successful answer per platform/question
Resolved citation-status observations	255	Source-level citation status recoverable
Cited observations	186	At least one answer-linked URL
Resolved zero-URL observations	69	Valid answers with no explicit URL
Source-capture gaps	105	Not treated as zero-citation observations
Answer-linked citations analysed	1,083	Primary citation-level corpus
Surfaced-only sources excluded	25	Retained separately for QA

4 Results

The recoverable, citation-weighted cross-platform corpus is strongly international and multilateral. Of 1,083 answer-linked citations, 745 (68.8 per cent) were attributed to UN or multilateral producers. Universities and research institutes contributed 108 (10.0 per cent), international NGOs 67 (6.2 per cent), international media 42 (3.9 per cent) and donor/bilateral organisations 32 (3.0 per cent). African national media, national public agencies and local/national NGOs or CSOs each supplied 21 citations (1.9 per cent). Because platforms contribute unequal numbers of recoverable citations, these percentages describe the pooled recoverable citation corpus rather than an equal-weight four-platform estimate.

Within that recoverable citation-weighted corpus, African-based organisations produced 71 citations (6.6 per cent), and 60 (5.5 per cent) were country-local to the crisis being discussed. The distinction matters because a source can be Africa-based without being local to the country in question. The 60 country-local citations spanned 35 publishers; the 71 African-based citations spanned 43 publishers. The pooled result therefore identifies a representation gap in visible citations, while the complete-platform sensitivity analysis below tests whether the same pattern persists without source-capture missingness.

Citation attention was also concentrated. The five most-cited producers supplied 40.3 per cent of answer-linked citations, the top ten 58.8 per cent and the top twenty 71.3 per cent. The HHI corresponds to roughly 22 equally weighted publishers despite 183 distinct producers. UNICEF led with 136 citations, followed by UNHCR (100), OCHA/Humanitarian Country Team products (73), OCHA (64) and IOM (63).

Local visibility was itself concentrated. The top five country-local publishers accounted for 40.0 per cent of local citations, the top ten 56.7 per cent and the top twenty 75.0 per cent; the effective number of local publishers was about 16.7. Kenya Red Cross Society led with eleven citations, followed by The Star Kenya with five. Nigeria's National Emergency Management Agency and the Ethiopian Human Rights Commission each had three; most local producers appeared only once.

Geographic variation was substantial in the recoverable cross-platform set. Kenya supplied 31 of 60 country-local citations and Nigeria nine; DRC and Sudan supplied none. Because source recovery is uneven, these differences are descriptive. The complete ChatGPT sensitivity analysis reproduces the main pattern: country-local sources were 41 of 567 citations (7.2 per cent), including 29.4 per cent for Kenya, 12.5 per cent for Nigeria and 8.9 per cent for Mozambique. ChatGPT cited no country-local source for DRC, Somalia or Sudan. In the complete ChatGPT sample, the cluster-bootstrap 95 per cent interval was 4.2-10.7 per cent for country-local citation share and 65.1-76.4 per cent for UN/multilateral share; 23.3 per cent of questions contained at least one local citation (95 per cent interval 14.4-32.2 per cent).

Theme differences were also pronounced. In the recoverable cross-platform set, WASH and humanitarian response/funding/localisation each produced thirteen country-local citations, while food insecurity and malnutrition produced none. The complete ChatGPT archive showed the same contrast: local sources were 23.9 per cent of WASH citations and 14.1 per cent of response/funding/localisation citations, but zero of food insecurity citations. Question framing also mattered in ChatGPT: geographic-hotspot questions had the highest local share (15.9 per cent), whereas unmet-needs questions had none; these question-type comparisons are descriptive because realised counts vary.

Platform citation behaviour differed sharply, although uneven source recovery prevents direct comparison as complete samples. Gemini's 71 recoverable citations were 78.9 per cent UN/multilateral and 1.4 per cent country-local. DeepSeek exposed 52 URLs (34.6 per cent UN/multilateral and 17.3 per cent local), but only 11 of 80 resolved observations contained any URL. ChatGPT produced 567 citations across all 90 production questions: 70.9 per cent UN/multilateral, 7.8 per cent African-based and 7.2 per cent local. Perplexity produced 393 answer-linked citations across 66 captured observations: 68.4 per cent UN/multilateral and 2.3 per cent local.

Separating host domain from producer reveals another layer. Overall, 231 citations (21.3 per cent) used an intermediary, republisher or asset host. ReliefWeb alone hosted 217 answer-linked citations: 20.0 per cent of the entire recoverable corpus and 93.9 per cent of intermediary-hosted citations. A domain-only audit would therefore make ReliefWeb appear to be the largest single source, even though OCHA, UNHCR, UNICEF, clusters, NGOs and others produced many hosted reports.

Observation-level analysis clarifies how local sources enter an answer—of the 186 observations with at least one recoverable answer-linked citation, 33 (17.7 per cent) included a country-local source. Twenty-eight of those 33 (84.8 per cent) also cited at least one UN or multilateral source. Only one observation relied exclusively on country-local sources. Eight combined local and UN sources without another international source; twenty combined local, UN and other international sources; four combined local sources with non-UN international sources. The dominant pattern is therefore supplementation.

Local-citing observations contained an average of 6.52 citations versus 5.67 in cited observations without a local source. They nevertheless contained fewer UN citations on average (3.18 versus 4.18) and had a lower mean within-answer UN share: 47.6 per cent versus 73.2 per cent. This larger overall source stack combined with lower UN weight is consistent with diversification plus partial substitution. In ChatGPT, the strongest complete-platform signal, 21 local-citing observations had a mean UN share of 51.5 per cent versus 76.3 per cent in 69 non-local observations. Perplexity showed a similar descriptive pattern (49.2 per cent versus 71.5 per cent).

Together, the results identify four connected properties of visible AI citation: scarce country-local representation, concentration among a small producer core, heavy intermediary hosting, and local evidence entering mainly as an additional but diversifying layer within internationally anchored source stacks.

Table 2. Publisher-type composition of the recoverable, citation-weighted cross-platform corpus

Publisher type	Citations	Share
UN/multilateral	745	68.8%
University/research institute	108	10.0%
International NGO	67	6.2%
International media	42	3.9%
Donor/bilateral	32	3.0%
African national media	21	1.9%
National government/public agency	21	1.9%
Local/national NGO or CSO	21	1.9%
Data platform	16	1.5%
African regional media	7	0.6%
Other	3	0.3%

Table 3. Platform citation behaviour in the recoverable archive (descriptive; uneven source recovery)

Platform	Resolved obs.	Capture gaps	Cited obs.	Zero-URL obs.	Citations	UN share	Local share	Top-5 share
Gemini	19	71	19	0	71	78.9%	1.4%	56.3%
DeepSeek	80	10	11	69	52	34.6%	17.3%	34.6%
ChatGPT	90	0	90	0	567	70.9%	7.2%	49.2%
Perplexity	66	24	66	0	393	68.4%	2.3%	46.8%

Note: Platform rows are not directly comparable as complete samples because source-capture missingness is concentrated in Gemini and Perplexity, and citation-level pooled estimates are not platform-balanced. DeepSeek zero-URL observations are valid first responses with no explicit source link. ChatGPT is the only platform with complete source-level capture across all 90 production questions and is therefore used for the principal complete-platform sensitivity analysis.

Table 4. Complete-platform sensitivity analysis: ChatGPT country-local citations

Country	Questions	Citations	Local citations	Local share
Burkina Faso	12	72	3	4.2%
Chad	5	31	1	3.2%
DRC	8	51	0	0.0%
Ethiopia	15	81	4	4.9%
Kenya	11	68	20	29.4%
Mozambique	7	45	4	8.9%
Nigeria	8	48	6	12.5%
Somalia	7	48	0	0.0%
South Sudan	13	93	3	3.2%
Sudan	4	30	0	0.0%

5 Discussion

Generative AI is not merely answering humanitarian questions; it is curating a visible hierarchy of humanitarian authority. In the recoverable, citation-weighted cross-platform corpus, country-local producers supplied 5.5 per cent of citations while UN and multilateral organisations supplied 68.8 per cent, and five producers captured 40.3 per cent of citation attention. These pooled figures are not equal-weight platform estimates, and they do not show that local sources were unavailable or prove bias. The complete ChatGPT sensitivity analysis, where all 90 production questions have source-level capture, nevertheless shows the same substantive pattern: 7.2 per cent country-local and 70.9 per cent UN/multilateral citations. Together, the analyses show that the evidence layer visible to users is strongly international and institutionally concentrated.

Citation localisation extends humanitarian localisation into this information layer. Traditional localisation asks who controls resources, programmes, partnerships and decisions. Citation localisation asks who becomes legible as a source when AI mediates a user's entry into a crisis. Generative systems compress many possible sources into a small evidentiary set; when that set lacks diversity, existing institutional hierarchies can become more visible even if each citation is sound.

This matters operationally because AI-generated summaries may increasingly become an early orientation layer for journalists, policymakers, donors, analysts, students and humanitarian staff who are not already embedded in specialist information systems. Visibility at that entry point can influence which organisations appear authoritative enough to investigate further, whose framing is repeated, and which evidence is absent from subsequent attention. Citation localisation therefore concerns recognition as well as retrieval: it measures whether local institutions are visible as named sources in the public-facing knowledge chain through which crises are interpreted.

The result should not be read as an argument that UN agencies ought to be cited less on principle. They aggregate country data, coordinate multi-sector reporting and publish standardised, frequently updated products that are genuinely useful in emergencies. The concern is structural dependence: repeated reliance on the same international layer can make locally produced observation and interpretation less visible exactly as users increasingly delegate source discovery to AI. The relevant question is whether credible local evidence has a viable route into the citation stack.

Concentration sharpens that concern. The top ten producers captured nearly 59 per cent of all citations, while only twenty local publishers captured three-quarters of local citations. The policy issue is therefore not simply 'cite more African sources'. It is whether a broad range of national and local organisations has the publishing, indexing, authority and distribution infrastructure required to become consistently retrievable and attributable.

Kenya illustrates what local visibility can look like when several institutional forms break through: the Kenya Red Cross Society, national media, government agencies, refugee authorities and the Kenya News Agency all appeared. Kenya's local share also remained high in the fully captured ChatGPT subset. By contrast, DRC had no country-local citation in either the recoverable cross-platform set or ChatGPT's complete archive. That absence is not evidence that credible Congolese sources do not exist; it prompts investigation into why they did not appear in this particular evidence chain.

Theme and question-type variation point in the same direction. Local visibility was highest in WASH and humanitarian response/funding/localisation, while food insecurity and malnutrition produced no country-local citation even in the complete ChatGPT archive. Geographic-hotspot questions also surfaced local sources more often than unmet-needs questions. One plausible explanation is that some humanitarian information systems are highly institutionalised around international classification and coordination products, while geographically specific questions create more openings for national reporting. This requires direct retrieval-level testing.

The host-versus-producer result is a methodological contribution in its own right. ReliefWeb hosted one-fifth of recoverable answer-linked URLs and almost 94 per cent of intermediary-hosted citations. A domain-only audit would mistake distribution infrastructure for knowledge production. In humanitarian research, where repositories are central to information circulation, citation studies should therefore code both the route through which evidence becomes discoverable and the organisation that produced it.

Source-stack analysis adds a further refinement. Local sources rarely replaced international authority outright: 84.8 per cent of local-citing observations still included a UN or multilateral source, and only one cited response was local-only. However, local-citing answers had a mean UN share 25.6 percentage points lower than non-local answers despite slightly larger citation stacks. The best description is supplementation with partial substitution: local evidence usually joins an international stack but is associated with a materially more plural distribution of visible authority.

In practice, evaluate citation localisation at the level of composition rather than as a quota. Adding one local link to an otherwise unchanged international stack would satisfy a superficial representation target without changing evidentiary weight. A stronger goal is for relevant local evidence to support substantive claims across countries and themes while reducing dependence on a narrow set of international authorities.

Platform interfaces also shape what auditors can assess. DeepSeek often produced substantive answers without explicit URLs; ChatGPT exposed many citations; Perplexity could surface sources beyond those linked to the narrative. Citation visibility is therefore a product feature, not a transparent window into hidden retrieval. Future comparative work should capture both retrieval and user-facing attribution where technically possible.

For humanitarian organisations, digital visibility is becoming part of knowledge power. Local actors may need support not only to produce evidence but to publish it through durable URLs, accessible HTML, clear dates and authorship, consistent institutional identities and trusted repositories. The aim is not to game AI systems; it is to make credible local evidence discoverable, attributable and reusable without losing provenance.

International organisations and repositories have a complementary responsibility. Syndication can expand discoverability, but metadata should preserve original authorship and institutional contribution when national or local partners collect data or produce analysis. Intermediary infrastructure can then function as a bridge to local authority rather than absorbing it into the host domain.

These findings also speak to decolonial AI debates. Birhane (2020) and Mohamed, Png and Isaac (2020) warn that technical infrastructures can reproduce global power relations. AHACI does not prove algorithmic colonisation. It identifies a narrower, measurable mechanism through which unequal authority can become visible: source attribution. That empirical claim is both more defensible and more actionable.

6 Limitations and Future Research

The study has eight principal limitations. First, source-level capture is missing for 105 of 360 production observations. Missingness is strongly platform-clustered, especially for Gemini and Perplexity, and therefore cannot be treated as random. The paper responds by separating observation-level coverage from citation-level composition, avoiding inferential platform claims and presenting a full 90-question production ChatGPT sensitivity analysis. Future collection waves should capture machine-readable source lists or full-page archives at the moment of response generation.

Second, AHACI observes visible citations rather than the complete retrieval process. A model may retrieve a source and use it without exposing a URL, or cite a page that only partially supports a claim. The current paper therefore does not equate absence of citation with absence from retrieval. Retrieval logs, where technically and ethically accessible, would allow future work to distinguish retrieval exclusion from citation exclusion.

Third, the study does not estimate the supply of eligible local sources for each question. Consequently, a 5.5 per cent local citation share cannot be interpreted as a formal under-citation rate or proof of bias. A stronger causal design would build query-specific candidate pools containing relevant local and international documents and measure selection conditional on availability, retrieval rank, freshness and content quality.

Fourth, localness is an institutional category rather than a complete measure of epistemic ownership. A UN report may incorporate data collected by national ministries; an international NGO report may be co-produced with local partners; a national media report may quote international agencies. The host-producer distinction reduces one source of misattribution but cannot reconstruct the full genealogy of every fact. Future research should code co-authorship, data provenance and partner acknowledgement.

Fifth, all prompts were asked in English. This is a structural limitation rather than a minor presentation choice. English-language querying and web retrieval can advantage organisations that publish consistently in English, maintain English-language pages or syndicate through international repositories, while disadvantaging locally authoritative sources in Francophone, Lusophone, Arabic-speaking and other multilingual information environments. The comparatively stronger local visibility observed in Kenya and Nigeria may therefore reflect language and publishing infrastructure alongside other factors. AHACI should be interpreted as measuring citation localisation within an English-language query environment, not general local-source visibility across languages. Multilingual replication is essential.

Sixth, one analyst produced the source classifications. The codebook and confidence flags increase transparency, but they do not substitute for independent inter-coder reliability. Although only one citation remained low-confidence after adjudication, future waves should independently double-code a random 10–20 per cent subset and report agreement statistics for the principal geographic and organisational variables.

Seventh, the audit is a one-day snapshot of rapidly changing consumer systems. The study records the products and collection conditions visible on 23 August 2026, but exact backend model builds, retrieval partners, and ranking systems were not consistently disclosed. Interface design and citation behaviour can change without notice. Longitudinal replication is necessary before interpreting differences as stable platform characteristics, and repeated queries and paraphrase variants would help quantify run-to-run variability.

Eighth, this paper does not evaluate factual accuracy, citation support or source quality across the full 1,083-citation corpus because those fields remain incompletely coded. This is deliberate. Representation and quality are separate questions, and increasing local-source visibility should never require lowering evidentiary standards. The next AHACI phase should jointly analyse provenance, claim support, timeliness and quality so that citation localisation is assessed alongside reliability rather than in place of it.

7 Conclusion

Generative AI is becoming an information gateway to humanitarian crises. The sources it cites are therefore part of the infrastructure through which it allocates authority. Within the first 100-question AHACI Batch 1, the 90-question production analysis produced a recoverable, citation-weighted corpus in which country-local knowledge producers accounted for 5.5 per cent of 1,083 answer-linked citations, African-based producers 6.6 per cent and UN or multilateral organisations 68.8 per cent. These pooled figures reflect unequal source recovery across platforms; importantly, the complete 90-question ChatGPT sensitivity analysis independently found 7.2 per cent country-local and 70.9 per cent UN/multilateral citations. The top five publishers captured 40.3 per cent of pooled citation attention. Local visibility was scarce and concentrated, while intermediary hosting—overwhelmingly ReliefWeb—shaped much of the visible URL layer.

The observation-level results add an important nuance. Local sources usually entered answers alongside international authority rather than replacing it: nearly 85 per cent of local-citing answers also cited UN or multilateral sources, and only one answer relied entirely on local sources. However, local-citing answers were materially less UN-dominated, suggesting that local inclusion can diversify the evidence stack rather than decorate it.

These findings support a new extension of the humanitarian localisation agenda. Localisation in an AI-mediated information environment concerns not only who receives funding or implements programmes, but also whose evidence is discoverable, attributable and visible when users ask machines to explain a crisis. Citation localisation offers a practical way to measure that layer. It directs attention towards the institutions that produce knowledge, the platforms that distribute it, the technical conditions that make it retrievable and the citation choices that make authority visible.

The policy goal should not be an arbitrary quota of local citations. Humanitarian information must remain accurate, current and relevant. The more defensible objective is a plural evidence ecosystem in which strong local sources can compete for visibility on evidentiary merit rather than disappear because they lack the publishing infrastructure, syndication pathways or machine-readable signals enjoyed by larger international institutions. Measuring that ecosystem is a first step towards changing it.

Declarations

Ethics: The study analysed publicly visible outputs from consumer AI systems and did not involve human participants, personal data or intervention with affected populations. No human-subject ethics approval was required.

Funding: This study received no external funding.

Competing interests: The author declares no competing interests.

Data availability: An anonymised reproducibility package accompanies the manuscript as supplementary material. It contains the Batch 1 question list, frozen production prompt structure, coding framework,

citation-ledger workbook and analysis workbook. The author intends to deposit the package in a DOI-bearing public repository at publication.

Use of generative AI: Generative AI was used as a research object in the data collection and as a language-editing and drafting aid during manuscript preparation. The author designed the study, set the protocol, reviewed the evidence, verified the analysis and accepts full responsibility for the final manuscript.

References

Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A. (2024). GEO: Generative engine optimisation. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 5–16). Association for Computing Machinery. https://doi.org/10.1145/3637528.3671900

Barter, D., & Sumlut, G. M. (2023). The "conflict paradox": Humanitarian access, localisation, and (dis)empowerment in Myanmar, Somalia, and Somaliland. Disasters, 47(4), 849–869. https://doi.org/10.1111/disa.12573

Birhane, A. (2020). Algorithmic colonisation of Africa. SCRIPTed, 17(2), 389–409. https://doi.org/10.2966/scrip.170220.389

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. Patterns, 2(2), Article 100205. https://doi.org/10.1016/j.patter.2021.100205

Couldry, N., & Mejias, U. A. (2019). The costs of connection: How data is colonising human life and appropriating it for capitalism. Stanford University Press. https://doi.org/10.1515/9781503609754

Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001

Inter-Agency Standing Committee. (2016). The Grand Bargain: A shared commitment to better serve people in need. https://interagencystandingcommittee.org/sites/default/files/migrated/2017-02/grand_bargain_final_22_may_final-2_0.pdf

Kakimov, R., Tan, X., Gillham, J., & Bejtic, N. (2026). Auditing citation behaviour in AI-generated search summaries: A framework and a case study of Google AI Overviews. In L. Bouzar-Benlabiod & C. Leung (Eds.), Proceedings of the 39th Canadian Conference on Artificial Intelligence (Vol. 318, pp. 224–235). PMLR. https://proceedings.mlr.press/v318/kakimov26a.html

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy & Technology, 33, 659–684. https://doi.org/10.1007/s13347-020-00405-8

Mugaanyi, J., Cai, L., Cheng, S., Lu, C., & Huang, J. (2024). Evaluation of large language model performance and reliability for citations and references in scholarly writing: Cross-disciplinary study. Journal of Medical Internet Research, 26, e52935. https://doi.org/10.2196/52935

Pearson, M., & Rieger, N. (2026). The state of international humanitarian funding to local and national actors. ODI Global. https://odi.org/en/publications/the-state-of-international-humanitarian-funding-to-local-and-national-actors/

Robinson, A. (2026). Localisation without transformation: Dynamics of "localisation" in South Sudan's humanitarian arena. Development in Practice. Advance online publication. https://doi.org/10.1080/09614524.2026.2662955

Sünkler, S., Lewandowski, D., Schultheiß, S., & Koop, O. (2026). Information diversity and authority shifts in Google's AI Overviews: An audit of health queries. In Companion Publication of the 2026 18th ACM

Web Science Conference (pp. 24–28). Association for Computing Machinery. https://doi.org/10.1145/3795513.3807429

Venkit, P. N., Laban, P., Zhou, Y., Mao, Y., & Wu, C.-S. (2024). Search engines in an AI era: The false promise of factual and verifiable source-cited responses [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2410.22349

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. Scientific Reports, 13, Article 14045. https://doi.org/10.1038/s41598-023-41032-5

Willitts-King, B., Majid, N., Ali, M., & Poole, L. (2018). Funding to local humanitarian actors: Evidence from Somalia and South Sudan. Overseas Development Institute. https://odi.org/en/publications/funding-to-local-humanitarian-actors-evidence-from-somalia-and-south-sudan/

Xu, H., Iqbal, U., & Montgomery, J. M. (2026). Measuring Google AI Overviews: Activation, source quality, claim fidelity, and publisher impact [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2605.14021