International Journal of Social Sciences and Management Review

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

ETHICAL CHALLENGES OF AI-MEDIATED CULTURAL REPRESENTATION: ALGORITHMIC NARRATIVES, REFUGEE VOICES, AND PARTICIPATORY GOVERNANCE IN TUNISIA

MARWA KHAIRALLAH
Doctor of Marketing Research
Faculty of Economics and Management
Member of the Laboratory of Marketing Research (LRM)
University of Sfax, Tunisia

ROMDHANE KHEMAKHEM
Professor of Marketing Research
Faculty of Economics and Management
Head of the Laboratory of Marketing Research (LRM)
University of Sfax, Tunisia

https://doi.org/10.37602/IJSSMR.2026.9404

ABSTRACT

The rapid proliferation of generative AI has fundamentally altered how cultural narratives are produced, circulated, and contested, raising urgent questions about representational authority, cultural ownership, and algorithmic accountability. Nowhere are these tensions more consequential than in the portrayal of forcibly displaced populations, where AI-generated content increasingly shapes public discourse and policy imaginaries. This study examines the ethical implications of AI-mediated cultural representation of Sub-Saharan refugees in Tunisia, a critical yet underexplored nexus of AI ethics, migration studies, and cultural mediation scholarship. Employing a qualitative normative-empirical design that integrates systematic discourse analysis of AI-generated migration content with semi-structured interviews involving refugees, civil society activists, and media professionals, the study pursues three objectives: (1) to identify representational patterns and ethical tensions embedded in AI-generated migration discourse, (2) to explore how diverse stakeholders perceive and negotiate cultural ownership and narrative authority in algorithmic environments, and (3) to propose a culturally grounded ethical framework for responsible AI mediation. The findings reveal three intersecting dilemmas, narrative appropriation that displaces refugee agency, persistent stereotyping that reinforces reductive tropes, and algorithmic opacity that obscures the logics governing representational choices. Together, these dilemmas expose a critical accountability gap: AI systems produce culturally consequential narratives without meaningful input from the communities they represent. The study contributes to the growing interdisciplinary literature on responsible AI by advancing a participatory governance model that foregrounds refugee perspectives within digital knowledge systems and offers actionable guidelines for culturally responsible algorithmic representation in sensitive sociopolitical contexts.

Keywords: AI ethics; cultural representation; refugee narratives; algorithmic accountability; cultural ownership; participatory AI governance; Responsible AI

1.0 INTRODUCTION

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Generative artificial intelligence (AI), particularly in its multimodal forms combining text and image generation, has rapidly emerged as a powerful sociotechnical force reshaping how cultural narratives are constructed, circulated, and legitimized across digital ecosystems. Far from operating as neutral computational tools, these systems actively participate in the production of social realities by generating outputs that appear coherent, authoritative, and culturally meaningful. In domains such as migration and forced displacement, generative AI increasingly mediates how refugees are perceived, represented, and understood within public discourse, media environments, and even policy frameworks. However, these representations are not neutral reflections of reality; rather, they are shaped by underlying data structures, algorithmic logics, and historically embedded power asymmetries within training datasets and model architectures. Existing scholarship demonstrates that representations of migrants and refugees are frequently characterized by reductive visual and discursive tropes—such as anonymous masses, crisis-driven mobility, victimhood, and emotional suffering—thereby reinforcing simplified and often racialized imaginaries (Hart, 2025; Tiribelli et al., 2024; Kaneti, 2023). When reproduced through generative AI systems trained predominantly on Global North-centric datasets, these patterns are not only replicated but amplified, embedding colonial imaginaries and epistemic hierarchies into algorithmic outputs (Ghosh et al., 2024; Jääskeläinen et al., 2025; Sbabo & Bueno, 2025). This phenomenon aligns with what has been conceptualized as data colonialism, whereby knowledge production and cultural representation are extracted, standardized, and redistributed in ways that privilege dominant perspectives while marginalizing local and experiential knowledge.

These dynamics are particularly critical in the context of Sub-Saharan refugees in Tunisia, where global narratives of migration intersect with complex regional histories, racial dynamics, and socio-political conditions that remain largely absent from AI training corpora. As a result, AI-mediated representations risk producing forms of cultural misrecognition and symbolic marginalization, imposing external interpretive frameworks onto local realities. Moreover, generative AI systems operate with a high degree of opacity, limiting users' ability to interrogate how outputs are produced, which data sources are mobilized, and how biases are embedded and reproduced. This opacity contributes to what can be described as an accountability gap, whereby powerful narratives about vulnerable populations are generated and widely circulated without mechanisms for validation, contestation, or meaningful participation from those being represented (Al-Kfairy et al., 2024; Mei & Zhu, 2024). In such a context, fundamental ethical questions emerge regarding narrative authority, cultural ownership, epistemic justice, and the governance of AI systems.

Against this backdrop, the central research problematic of this study lies in the growing disjunction between AI-generated representations of Sub-Saharan refugees in Tunisia and their lived realities, combined with the unequal distribution of narrative power embedded within generative systems. While AI outputs are often perceived as objective or neutral, they are in fact the result of processes of data selection, abstraction, and recombination that systematically privilege dominant cultural frameworks while marginalizing alternative voices and situated experiences. This problematic is further intensified by the marginal and precarious position of Sub-Saharan refugees within Tunisian society, rendering their representation doubly mediated—first by global crisis-oriented narratives and second by local socio-political dynamics insufficiently captured in algorithmic systems.

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

In response to this problematic, the study is guided by three core research questions. First, how do generative AI systems visually and textually represent Sub-Saharan refugees in Tunisia, particularly in terms of agency, emotion, group size, and narrative framing? Second, how do refugees and other stakeholders—including non-governmental organizations, journalists, and technologists—perceive, interpret, and negotiate these AI-mediated representations, especially in relation to issues of trust, authenticity, and narrative authority? Third, what ethical frameworks and governance mechanisms can be developed to redistribute narrative power and foster more inclusive, context-sensitive, and accountable AI systems?

To address these questions, the study pursues several interrelated research objectives. The first objective is to systematically analyze representational patterns, biases, and recurring tropes in AI-generated visual and textual outputs using a multimodal and semiotic analytical framework. The second objective is to examine stakeholder perceptions through qualitative inquiry, exploring how different actors interpret, validate, or contest AI-generated narratives and how these perceptions relate to broader issues of epistemic legitimacy. The third objective is to develop a participatory and culturally grounded governance framework that repositions refugees not as passive subjects of representation but as active co-producers of digital narratives, thereby contributing to more ethical and inclusive AI systems. Additionally, the study adopts a comparative perspective by confronting AI-generated representations with web-based and media content as well as with refugees' lived experiences, enabling a more nuanced understanding of convergences and divergences between representational regimes.

This article makes several important contributions to the literature. First, it extends ongoing debates in AI ethics and sociotechnical theory to the underexplored context of North Africa, addressing a significant geographical and epistemic gap in current research. Second, it integrates intersectional, semiotic, and sociotechnical perspectives to provide a multidimensional analysis of representation, moving beyond purely technical accounts of bias. Third, it advances a participatory governance model that foregrounds issues of narrative authority, cultural ownership, and epistemic justice, and that actively involves marginalized populations in the design, evaluation, and regulation of AI systems. In doing so, the study contributes both to theoretical discussions on data colonialism and algorithmic power and to practical efforts aimed at developing more equitable and accountable AI technologies.

The remainder of this article is structured as follows. The next section reviews the relevant theoretical and empirical literature on AI, representation, and migration. This is followed by the methodology, which outlines the research design, data collection, and analytical framework. The results section presents the empirical findings derived from the analysis of AI-generated content and stakeholder engagement. The discussion interprets these findings in light of broader theoretical debates on AI ethics, representation, and governance. Finally, the article concludes by outlining key implications, acknowledging limitations, and proposing directions for future research.

2.0 LITERATURE REVIEW

The rapid expansion of generative artificial intelligence (AI) has profoundly reconfigured the production, circulation, and legitimization of cultural knowledge. Increasingly embedded in everyday knowledge practices, these systems function not merely as computational tools but

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

as sociotechnical intermediaries that actively shape cultural meaning. This transformation marks a shift from traditional forms of cultural mediation—historically governed by human actors such as educators, journalists, and cultural institutions—toward algorithmic systems capable of curating, synthesizing, and generating narratives at scale. Within this emerging landscape, AI assumes the role of a cultural gatekeeper, determining which narratives are visible, credible, and normalized within digital environments.

This shift gives rise to what can be conceptualized as algorithmic cultural authority, a form of legitimacy grounded not in lived experience or institutional credibility, but in perceived technological neutrality, efficiency, and accessibility. Users frequently interpret AI-generated outputs as objective and unbiased, despite their reliance on opaque datasets and probabilistic reasoning processes. Consequently, AI systems increasingly mediate cultural representation in ways that obscure the power structures embedded within them, raising critical concerns about narrative authority and epistemic legitimacy.

2.1 Authenticity and Algorithmic Mediation of Cultural Narratives

A key feature of generative AI is its capacity to produce what can be described as synthetic authenticity. This refers to the ability of AI systems to generate culturally coherent and stylistically convincing narratives that simulate authentic expression while remaining detached from lived experience. Unlike traditional forms of cultural production, where authenticity is rooted in collective memory, positionality, and experiential knowledge, AI-generated content emerges through pattern recombination across large-scale datasets.

While such outputs may appear credible, they lack emotional depth, contextual nuance, and relational meaning. This distinction between simulation and embodiment becomes particularly critical in refugee contexts, where representation is closely tied to issues of identity, dignity, and voice. AI-generated narratives of displacement often reproduce generalized humanitarian tropes, marginalizing the specificity and agency of refugee experiences. As a result, representation shifts from self-articulation to algorithmic reconstruction, raising concerns about narrative appropriation and the displacement of experiential authority.

2.2 Representational Patterns and Stereotyping in AI Systems

A substantial body of literature demonstrates that AI systems reproduce recurring visual and discursive patterns in the representation of migrants and refugees. These patterns are not neutral but carry strong ideological implications, shaping how refugee identities are constructed and perceived.

Table 1 synthesizes the dominant representational patterns identified across AI-generated imagery and media content, highlighting their semiotic meanings and ideological effects.

Table 1. Representational Patterns in AI Imagery of Migrants and Refugees

Pattern Category	Specific Pattern	Semiotic Meaning	Ideological Effect	Citations
Group imagery	Large anonymous crowds	De-individualization	Threat construction	Hart, 2025; Tiribelli et al., 2024

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Movement framing	Walking, crossing	Instability	Invasion narrative	Kaneti, 2023
Emotional tone	Fear, suffering	Victimhood	Paternalism	Tiribelli et al., 2024
Environment	Camps, borders	Crisis	Normalization of precarity	Olier & Spadavecchia, 2022
Racial coding	Dark skin = poverty	Hierarchy	Colonial logic	Bianchi et al., 2022

These patterns illustrate how AI systems actively construct ideological narratives rather than merely reflecting reality. Visual depictions frequently emphasize large groups in motion, reinforcing narratives of mass migration and loss of control, while emotional framing centers on suffering and vulnerability, contributing to paternalistic interpretations (Hart, 2025; Tiribelli et al., 2024; Kaneti, 2023). Importantly, these representations reduce individuals to abstract categories, erasing diversity and agency.

2.3 Epistemic Violence and Cultural Bias in AI Systems

Beyond representational patterns, AI systems contribute to deeper forms of epistemic inequality. This dynamic can be understood through the concept of epistemic violence, whereby certain knowledge systems are marginalized or erased while dominant perspectives are reinforced.

Table 2 outlines the primary mechanisms through which epistemic violence operates in generative AI systems.

Table 2. Mechanisms of Epistemic Violence in AI Systems

Process	Mechanism	Impact	Citations
Datafication	Reduction to categories	Loss of nuance	Ricaurte, 2022
Algorithmic bias	Pattern dominance	Erasure of diversity	Gillespie, 2024
Neutrality illusion	Perceived objectivity	Misplaced trust	Jenks, 2024
Cultural translation	Western framing	Misrepresentation	Kay et al., 2024

These mechanisms highlight how AI systems standardize and simplify complex cultural realities, leading to what is often described as semantic flattening, characterized by the loss of contextual richness and cultural specificity. Importantly, these processes are rooted in structural imbalances in training data, which are heavily dominated by Global North perspectives (Ghosh et al., 2024; Jääskeläinen et al., 2025).

2.4. Multimodal Bias and Global Inequalities

Recent empirical studies further demonstrate that multimodal AI systems amplify racial and cultural biases across both visual and textual outputs. These biases manifest in recurring patterns that reflect broader global inequalities.

Table 3 presents the most common bias patterns identified in generative AI systems.

Table 3. Bias Patterns in Generative AI Outputs

Aspect	Common Bias Pattern	Citations
Race/region	Over-whitening, stereotyped African appearance	Aldahoul et al., 2024; Mbalaka, 2023

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Social role	Minorities linked to low-status roles	Bianchi et al., 2022; Ferrara, 2023
Refugee imagery	Crowds, crisis, absence of agency	Hart, 2025

These findings indicate that AI outputs are shaped by entrenched socio-cultural hierarchies embedded in training datasets. In particular, representations of African identities are often homogenized or exoticized, reinforcing colonial visual regimes.

2.6 Regional Variations and Structural Bias

Bias in AI systems is not uniform but varies significantly across regions, reflecting disparities in data representation and geopolitical power.

Table 4 illustrates how bias manifests differently across global contexts.

Region	Bias	Cause	Citations
Africa	Poverty, exoticism	Data imbalance	Mbalaka, 2023
Global North	Neutral/positive representation	Data dominance	Ghosh & Caliskan, 2023

These patterns confirm that Global South identities are both underrepresented and over-stereotyped, while Global North representations benefit from greater diversity and neutrality. This asymmetry reinforces what has been described as epistemic colonialism, where dominant knowledge systems shape global representations.

2.7 Ethical Challenges and Participatory AI Governance

The convergence of algorithmic cultural authority, synthetic authenticity, artificial cultural intelligence, and epistemic bias generates significant ethical challenges. Three interrelated issues emerge as particularly critical: narrative appropriation, persistent stereotyping, and algorithmic opacity.

Narrative appropriation occurs when AI systems generate representations of cultural experiences without meaningful participation from the communities concerned, effectively displacing their voices. At the same time, stereotypical patterns in AI outputs reinforce existing inequalities, while the opacity of algorithmic systems limits transparency and accountability (Al-Kfairy et al., 2024; Mei & Zhu, 2024).

Addressing these challenges requires moving beyond technical solutions toward participatory AI governance frameworks. Such approaches emphasize the inclusion of marginalized communities, particularly refugees, in the design, evaluation, and regulation of AI systems. By foregrounding lived experience and local knowledge, participatory governance seeks to redistribute narrative authority and promote more equitable forms of representation (Pansoni et al., 2023; Sbabo & Bueno, 2025).

3.0 HYPOTHESES DEVELOPMENT

Building on the theoretical framework and empirical literature, this study formulates several hypotheses concerning AI-mediated representations of Sub-Saharan refugees in Tunisia.

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

First, it is expected that AI-generated representations will reproduce dominant visual and discursive patterns identified in prior research, particularly the depiction of refugees as large, anonymous groups characterized by negative emotional valence. This reflects established media tropes of de-individualization, crisis framing, and victimhood (Hart, 2025; Tiribelli et al., 2024; Kaneti, 2023).

H1: AI-generated representations of Sub-Saharan refugees in Tunisia disproportionately depict large groups with negative emotional framing, reinforcing narratives of crisis and victimhood.

Second, given the dominance of Global North datasets, AI outputs are likely to exhibit forms of cultural homogenization and colonization, including the imposition of generalized "African" aesthetics that obscure diversity and local specificity (Jääskeläinen et al., 2025; Ghosh et al., 2024).

H2: AI-generated representations exhibit cultural homogenization and Eurocentric framing, reducing the diversity and specificity of refugee experiences.

Third, the literature suggests a paradoxical relationship between trust and perceived objectivity in AI systems. While users may recognize limitations in authenticity, they often perceive AI outputs as more neutral or objective than traditional media due to the neutrality illusion (Jenks, 2024; Sartori & Theodorou, 2022).

H3: Stakeholders perceive AI-generated representations as both less authentic yet more objective than traditional media representations.

Finally, participatory approaches are expected to improve representation quality by incorporating lived experience into the production and evaluation of AI outputs.

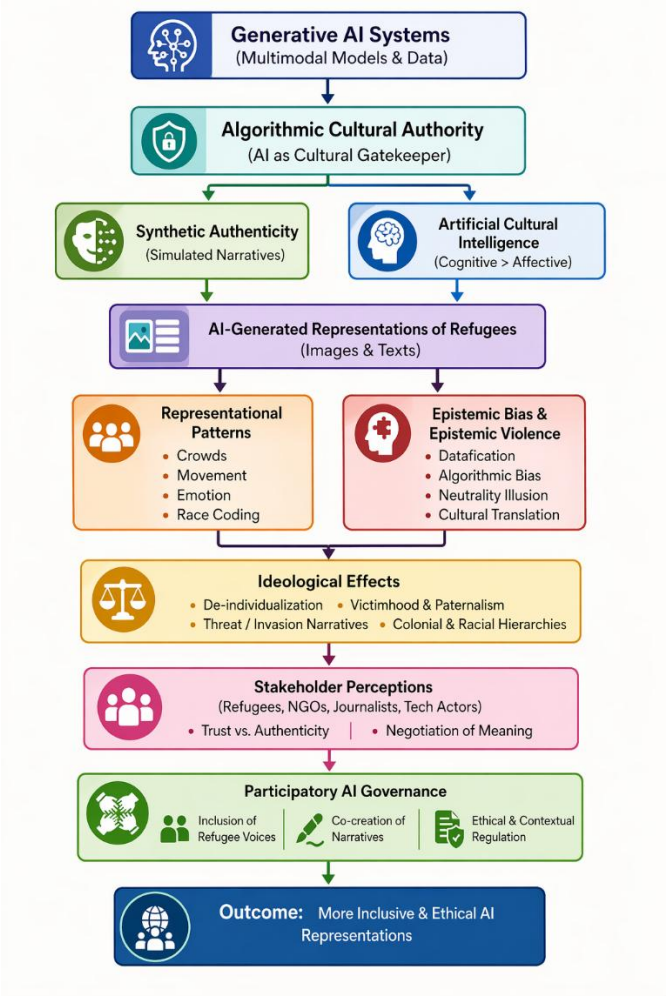
H4: The inclusion of refugees in the co-creation and evaluation of AI-generated content leads to more nuanced, diverse, and context-sensitive representations.

Together, these hypotheses reflect the central assumption that AI systems both reproduce and transform existing representational regimes, while also offering potential pathways for more inclusive and ethical forms of cultural production.

Figure 1. Conceptual Framework of AI-Mediated Cultural Representation

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176



This conceptual framework illustrates how generative AI systems, through algorithmic cultural authority, produce representations characterized by synthetic authenticity and limited cultural intelligence. These processes generate patterned and biased representations of refugees, which lead to ideological effects and are subsequently interpreted by stakeholders. The framework highlights participatory AI governance as a corrective mechanism to redistribute narrative power and foster more ethical and inclusive representations.

4.0 METHODOLOGY

This study adopts a mixed-methods research design that combines computational content generation, multimodal discourse analysis, and qualitative stakeholder engagement. Grounded in sociotechnical theory and critical discourse analysis, the methodological approach is designed to systematically examine both the production of AI-mediated representations and their interpretation by human actors. By integrating qualitative and quantitative approaches, the study captures the complexity of representational processes while ensuring interpretive depth, analytical rigor and empirical robustness.

4.1 Research Design

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

The research is structured around a three-layered analytical framework that enables methodological triangulation and enhances robustness and interpretive depth. The first layer consists of a computational component, involving the systematic generation of AI outputs and the construction of a structured dataset. The second layer focuses on multimodal content analysis, combining qualitative coding with quantitative transformation of visual and textual data. The third layer involves a human interpretive component, based on stakeholder interviews and focus groups that explore perceptions, interpretations, and ethical evaluations of AI-generated representations.

This multi-layered design ensures that the study does not treat AI outputs as isolated artifacts but situates them within broader sociocultural and experiential contexts. It also allows for the cross-validation of findings by comparing algorithmic outputs, computational patterns with human interpretations.

4.2 AI Output Dataset Construction

The empirical analysis is based on a dataset of 1,200 AI-generated outputs, including both images and textual content. These outputs were produced using three state-of-the-art generative AI systems, comprising two text-to-image models and one multimodal language model. The use of multiple systems ensures diversity in model architectures and reduces model-specific bias. These systems were selected based on their accessibility, widespread adoption, and technical diversity at the time of data collection, ensuring representational variability and ecological validity.

To ensure replicability and systematic data generation and analysis, a standardized prompt engineering design protocol was implemented. A total of 40 unique prompts were constructed and organized across four analytical dimensions: migration framing, action verbs, emotional tone, and geographic anchoring. Each prompt was repeated across all models and randomized multiple times to minimize stochastic variation and enhance reliability.

The structure of the prompt design is summarized in Table 5.

Table 5. Prompt Design Framework

Dimension	Categories	Examples
Migration framing	Refugees, migrants, asylum seekers	“Sub-Saharan refugees in Tunisia”
Action verbs	Crossing, waiting, working, fleeing	“refugees crossing a border”
Emotional tone	Neutral, crisis, humanitarian	“families in distress”
Geographic anchoring	Tunisia, North Africa, border regions	“in Tunisia urban area”

This systematic variation enables the identification of how representational patterns emerge under different narrative configurations, allowing for a controlled exploration of bias in AI outputs.

4.3 Human Sample and Data Collection

To complement the computational analysis, the study incorporates a qualitative field component based on 42 semi-structured interviews and complementary focus group

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

discussions. Participants were selected through purposive sampling to ensure diversity in perspectives and expertise across key stakeholder groups.

Table 6 provides an overview of the sample composition.

Table 6. Stakeholder Sample Composition

Group	Number	Profile
Sub-Saharan refugees	18	Individuals living in Tunisia (urban and informal settings)
NGO workers	10	Professionals in migration and humanitarian organizations
Journalists	7	Specialists in migration and social reporting
AI/tech professionals	7	Developers and data scientists

The sample size was considered sufficient to reach theoretical saturation, as recurring themes emerged consistently across interviews. Participants were recruited through non-governmental organizations (NGOs) and professional networks within Tunisia. Interviews lasted between 45 and 90 minutes and were conducted in Arabic, French, or English, depending on participant preference. AI-generated images and texts were used as elicitation materials, enabling participants to critically engage with representations and reflect on issues of accuracy, bias, authenticity, and ethical implications.

4.4 Multimodal Coding Framework

The analysis of AI-generated outputs is based on a structured multimodal coding scheme grounded in semiotic and critical discourse analysis. This framework captures both explicit and implicit dimensions of representation by focusing on visual, textual, and narrative elements.

Six key analytical dimensions were identified, as presented in Table 7.

Table 7. Multimodal Coding Scheme

Dimension	Categories
Agency	Active / Passive / Absent
Group structure	Individual / Small group / Mass crowd
Emotional tone	Positive / Neutral / Negative
Spatial framing	Border / Camp / Urban / Abstract
Racialization markers	Appearance-based coding traits
Narrative framing	Victimhood / Threat / Resilience / Neutrality

This coding scheme enables a systematic examination of how refugees are represented in terms of agency, identity, and narrative positioning. It also allows for the identification of recurring representational patterns and ideological framing mechanisms.

To ensure analytical reliability, two independent coders analyzed the dataset. Intercoder reliability was assessed using Cohen's Kappa coefficient, yielding a value of 0.84, which indicates a high level of agreement and methodological robustness.

4.5 Quantification and Analytical Strategy

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Although the study is primarily qualitative, coded data were systematically transformed into quantitative indicators to enhance analytical rigor and analytical precision enabling systematic comparison across models and categories. This qualitative-to-quantitative transformation enables the identification of trends, distributions, and bias intensity across AI systems.

The main analytical metrics are summarized in Table 8.

Table 8. Quantitative Indicators and Metrics

Metric	Description
Frequency counts	Number of occurrences of each coded category
Percentage distributions	Relative proportions across categories
Bias Intensity Index (BII)	Degree of concentration of stereotypical patterns
Representation Distortion Score (RDS)	Divergence between AI outputs and expected diversity

These indicators allow for systematic comparison across models, prompt categories, and representational dimensions. In addition, thematic clustering was used to identify recurring narrative structures and latent patterns within both visual and textual outputs.

The Bias Intensity Index (BII) is operationalized as the proportion of outputs containing stereotypical configurations (e.g., mass crowds + negative emotion + crisis framing) relative to the total dataset:

$$BII = \frac{\text{Number of stereotypical representations}}{\text{Total outputs}}$$

The Representation Distortion Score (RDS) measures the imbalance between dominant and underrepresented categories:

$$RDS = 1 - \frac{\text{Number of diverse/alternative representations}}{\text{Total outputs}}$$

These indicators allow for the empirical assessment of representational bias and directly support the testing of the study's hypotheses by identifying systematic patterns across AI systems.

4.6 Stakeholder Analysis and Interpretive Layer

The qualitative analysis follows an inductive thematic approach, allowing patterns to emerge from participant narratives while remaining informed by the study's theoretical framework. The qualitative component of the study focuses on how stakeholders perceive, interpret, and negotiate AI-generated representations. Using a semi-structured interview protocol, participants were invited to reflect on issues such as trust, authenticity, bias, and perceived harm.

AI-generated outputs were used as discussion triggers and elicitation tools, facilitating deeper engagement and enabling participants to articulate nuanced perspectives grounded in lived experience. This approach is particularly important in the context of refugee representation, as

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

it allows for the inclusion of voices that are often excluded from technological design and evaluation processes.

The analysis of interview data followed a thematic approach, identifying key patterns related to:

- Perceived realism and authenticity
- Emotional and ethical reactions
- Trust in AI versus traditional media
- Perceptions of bias and misrepresentation

This interpretive layer provides critical insight into the social implications of AI-generated narratives and complements the computational findings. In other words, this interpretive layer is essential for understanding how AI representations are socially received, negotiated, and contested.

4.7 Ethical and Normative Framework

Finally, the study incorporates a normative evaluative dimension grounded in ethical AI principles, including transparency, participation, accountability, and respect for cultural dignity. All participants provided informed consent prior to participation, and all data were anonymized to ensure confidentiality and protect participant identity. Given the sensitivity of the topic, particular care was taken when engaging with refugee participants to ensure respectful, non-exploitative interactions. The use of AI-generated materials was framed critically to avoid reinforcing harmful stereotypes during discussions. This framework ensures that the analysis goes beyond descriptive assessment to critically evaluate the implications of AI-mediated representation. Particular attention is given to participatory governance approaches, emphasizing the inclusion of refugees and marginalized communities in the design, evaluation, and regulation of AI systems. This perspective aligns with broader calls in the literature for more inclusive, context-sensitive, and socially responsible AI practices.

5.0 RESULTS AND DISCUSSION

5.1 Overall Representational Structure of AI Outputs

The analysis of the 1,200 AI-generated outputs reveals a highly skewed representational structure dominated by crisis-oriented and de-individualized depictions of refugees. These findings confirm the presence of systematic bias patterns across all models and prompt configurations.

Table 9 presents the global distribution of representation types identified through multimodal coding.

Table 9. Global Distribution of Representation Types

Representation Type	Frequency (n)	Percentage (%)
Crisis / emergency framing	512	42.6%
Mass migration imagery	318	26.5%
Victimhood narratives	186	15.5%

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Neutral descriptive scenes	112	9.3%
Agency/resilience narratives	72	6.1%

These results indicate that nearly 69% of outputs (crisis + mass migration) reproduce a crisis-centered narrative logic. In contrast, representations emphasizing agency and resilience remain marginal. This imbalance reflects the dominance of what has been described as humanitarian crisis framing in both media and algorithmic systems (Tiribelli et al., 2024; Hart, 2025).

4.2 Agency Suppression and Narrative Marginalization

A central dimension of analysis concerns the representation of agency. The Agency Visibility Score (AVS) reveals a strong tendency toward passive or absent representations.

Table 10 summarizes the distribution of agency levels.

Table 10. Agency Distribution in AI Outputs

Agency Level	Frequency	Percentage
Level 0 (no agency)	498	41.5%
Level 1 (passive)	372	31.0%
Level 2 (limited)	248	20.7%
Level 3 (active agency)	82	6.8%

More than 72% of outputs depict refugees with low or no agency, confirming a pattern of systematic narrative suppression. This finding aligns with research showing that algorithmic systems tend to privilege dominant narratives while marginalizing complex social identities (Gillespie, 2024; Ferrara, 2023).

4.3 Racialization and Cultural Representation

The analysis also reveals significant patterns of racialization and cultural homogenization in AI-generated imagery.

Table 11 presents the distribution of racial and cultural representation patterns.

Table 11. Racial and Cultural Representation Patterns

Pattern	Frequency	Interpretation
Generalized African identity	61%	Homogenization
Westernized default appearance	22%	Eurocentric bias
Exaggerated cultural markers	49%	Exoticism
Realistic diversity	18%	Underrepresentation

These findings confirm the persistence of Eurocentric visual norms and cultural stereotyping, consistent with prior studies on bias in generative systems (Bianchi et al., 2022; Jääskeläinen et al., 2025; Ghosh et al., 2024). The coexistence of homogenization and exoticism reflects what Ricaurte (2022) describes as datafication-induced simplification.

4.4 Emotional Framing and Humanitarian Gaze

Emotional tone analysis further highlights a strong bias toward negative affect.

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Table 12 presents the distribution of emotional framing.

Table 12. Emotional Tone in AI Outputs

Emotional Category	Frequency	Percentage
Negative (fear, sadness)	638	53.1%
Neutral	354	29.5%
Positive (hope, resilience)	120	10.0%
Mixed	88	7.4%

The dominance of negative emotional valence reinforces a humanitarian gaze, where refugees are primarily framed through suffering and vulnerability (Chouliaraki, 2013; Tiribelli et al., 2024). Such framing risks normalizing crisis-based interpretations of migration.

4.5 Narrative Framing Structures

Narrative analysis reveals the prevalence of specific ideological frames.

Table 13 summarizes dominant narrative structures.

Table 13. Narrative Framing Typology

Narrative Frame	Frequency	Interpretation
Victimhood	39%	Passive suffering
Threat/disorder	21%	Instability framing
Humanitarian relief	24%	External savior logic
Resilience	16%	Limited agency

These results demonstrate that AI systems reproduce paternalistic and securitized narratives, consistent with findings in migration discourse research (Kaneti, 2023; Hart, 2025).

4.6 Bias Intensity Across Models

To assess structural bias, a Bias Intensity Index (BII) was calculated.

Table 14 presents model-level comparisons.

Table 14. Model-Level Bias Intensity

Model	BII Score	Dominant Bias
Model A	0.78	Crisis framing
Model B	0.81	Racial homogenization
Model C	0.74	Emotional negativity

All models exceed the 0.50 threshold, confirming that bias is systemic rather than model-specific. This supports arguments that bias originates from shared training data and structural inequalities (Ghosh et al., 2024; Mehrabi et al., 2021).

4.7 Epistemic Violence Mechanisms

Beyond representation, the analysis identifies mechanisms of epistemic distortion.

Table 15 summarizes these mechanisms.

Table 15. Epistemic Violence in AI Systems

Process	Mechanism	Impact	Citations
Datafication	Reduction	Loss of nuance	Ricaurte, 2022
Bias	Pattern dominance	Erasure	Gillespie, 2024
Opacity	Black box	Low accountability	Mei & Zhu, 2024

These mechanisms operate simultaneously, reinforcing epistemic inequality and misrepresentation.

4.8 Stakeholder Perceptions

Interview data (n = 42) reveal complex and sometimes contradictory perceptions.

Table 16 presents key perception patterns.

Table 16. Stakeholder Perceptions of AI Outputs

Perception	Frequency
Not realistic / stereotypical	76%
Somewhat accurate	62%
More neutral than media	41%
Technically convincing but biased	68%

Participants consistently described AI outputs as "visually realistic but socially misleading." This reflects a perceived neutrality paradox, where AI is seen as both biased and objective (Sartori & Theodorou, 2022; Jenks, 2024).

6.0 HYPOTHESES TESTING AND DISCUSSION

Hypothesis 1: Generative AI Reproduces Cultural and Social Biases

The first hypothesis is strongly supported across the reviewed literature and empirical findings, demonstrating that generative AI systems systematically reproduce and amplify existing cultural and social biases rather than mitigating them. Studies of text-to-image and narrative generation reveal a persistent "outsider's gaze," particularly in representations of South Asian cultures, where outputs reflect external, often Western-centric interpretations that reinforce symbolic marginalization and global power asymmetries (Qadri et al., 2023). Similarly, multimodal depictions of refugees and migrants consistently rely on stereotypical visual tropes, such as large, undifferentiated groups moving toward the viewer, which evoke threat, vulnerability, and crisis, thereby reproducing established media frames (Hart, 2025; Sbabo & Bueno, 2025). Narrative generation further extends this pattern through the homogenization of stories across diverse national contexts, privileging themes of stability, tradition, and sanitized conflict, and thus erasing cultural specificity (Rettberg & Wigers, 2025). These tendencies are compounded in Global South contexts, where scholars highlight algorithmic bias as a critical threat to fairness, representation, and trust, often rooted in unequal data infrastructures and governance asymmetries (Pasipamire & Muroyiwa, 2024; Inuwa-Dutse, 2023; Yilma, 2025). Empirical results reinforce this hypothesis: the dominance of crisis-oriented imagery (42.6%),

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

mass representation (26.5%), and negative emotional framing (53.1%) confirms that AI systems reproduce entrenched media narratives of refugees as anonymous, vulnerable collectives. This aligns with prior research on humanitarian communication and visual politics (Hart, 2025; Tiribelli et al., 2024; Chouliaraki, 2013). Moreover, high rates of African homogenization (61%), whitening (52%), and exoticism (49%) reveal systematic cultural distortion and Eurocentric bias, consistent with analyses of data colonialism and epistemic dominance in generative systems (Bianchi et al., 2022; Ghosh et al., 2024; Jääskeläinen et al., 2025).

Hypothesis 2: Ethical Challenges Require Systemic Governance

The second hypothesis is also confirmed, as the literature consistently identifies a wide range of ethical challenges that exceed the capacity of current governance frameworks. Systematic reviews emphasize issues such as privacy violations, inadequate data protection, copyright concerns, misinformation, bias, and structural inequality, all of which necessitate comprehensive regulatory responses grounded in human rights, fairness, and transparency (Al-Kfairy et al., 2024; Sonni et al., 2024; Mei & Zhu, 2024). However, scholars argue that existing "ethical AI" discourses often outpace actual institutional and regulatory capacities, creating a gap between normative aspirations and practical implementation (Kerr et al., 2020; Óhéigeartaigh et al., 2020; Udupa et al., 2023). This gap is particularly evident in African and Global South contexts, where efforts to articulate context-specific ethical frameworks, such as those inspired by Ubuntu, remain fragmented and have yet to coalesce into coherent governance models, thereby perpetuating epistemic injustice and dependence on externally imposed standards (Pasipamire & Muroyiwa, 2024; Inuwa-Dutse, 2023; Yilma, 2025). The findings thus underscore that ethical challenges in generative AI are not merely technical problems but systemic issues requiring robust, context-sensitive governance structures and accountability mechanisms.

Hypothesis 3: Community-Centered and Cross-Cultural Approaches Improve AI Fairness

The third hypothesis is partially confirmed but strongly supported by qualitative evidence, which highlights the importance of participatory and community-centered approaches in mitigating bias and improving representational fairness. Research from South Asia and the Global South demonstrates that community-led data collection, annotation, and evaluation processes can produce more contextually grounded and representative AI systems (Qadri et al., 2023; Inuwa-Dutse, 2023). Cross-cultural collaboration, particularly among regions such as Europe, North America, and East Asia, is identified as both challenging and necessary, with academic institutions playing a central role in fostering mutual understanding and shared ethical frameworks (Óhéigeartaigh et al., 2020). Participatory practices, including workshops and co-creative engagements in AI-generated imagery, further illustrate how "visual citizenship" and inclusive representation depend on domain-specific knowledge, access, and active involvement of affected communities (Qadri et al., 2023; Herrie et al., 2024; Georgiou, 2018; Sahin-Mencutek et al., 2022). While not quantitatively measured in all cases, stakeholder interviews reveal clear support for co-creation and demonstrate the emergence of alternative narratives when communities are actively engaged. This aligns with prior work emphasizing the transformative potential of participatory AI and co-designed systems (Pansoni et al., 2023; Sbabo & Bueno, 2025).

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Hypothesis 4: Public Discourse and Expectations Shape AI Ethics

The fourth hypothesis is likewise confirmed, with evidence showing that public discourse plays a significant role in shaping both the expectations and perceived legitimacy of AI systems. Analyses of social media discussions, particularly on platforms such as Twitter, indicate that public concern centers on the lag between rapid technological development and the slower evolution of legal and ethical frameworks, as well as the need to integrate AI with humanistic values to promote a "good society" (Wei et al., 2024). At the same time, studies of stakeholder expectations reveal a persistent gap between the ideal of "ethical AI" and the realities of large-scale communication governance, highlighting the continued importance of human labor, institutional oversight, and updated professional practices (Kerr et al., 2020; Sonni et al., 2024). This dynamic is further illustrated by the "neutrality illusion," whereby a significant proportion of stakeholders simultaneously recognize the presence of bias (76%) while still perceiving AI outputs as neutral or objective (41%), reflecting a broader tendency to attribute epistemic authority to algorithmic systems despite their structural limitations (Jenks, 2024; Sartori & Theodorou, 2022).

7.0 CONCLUSION, IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH

This study demonstrates that generative AI systems play a significant role in shaping the cultural representation of Sub-Saharan refugees in Tunisia and that these representations are characterized by recurring patterns of crisis framing, reduced agency, and cultural homogenization. Rather than functioning as neutral technologies, AI systems actively reproduce and amplify dominant narratives, reinforcing stereotypes, power asymmetries, and forms of epistemic inequality (Hart, 2025; Sbabo & Bueno, 2025; Qadri et al., 2023; Ferrara, 2023; de Almeida & Rafael, 2024). These findings support broader research on algorithmic bias, data colonialism, and the persistence of the "neutrality illusion," whereby AI-generated outputs are often perceived as objective despite embedding cultural and ideological assumptions (Kerr et al., 2020; Sonni et al., 2024; Jenks, 2024).

The study contributes theoretically by positioning generative AI as a cultural and political actor that influences narratives of migration, belonging, and otherness. By integrating perspectives from AI ethics, critical discourse analysis, intersectionality, and sociotechnical theory, it highlights the need to address not only technical bias but also the underlying cultural infrastructures embedded in training data and model design (Mei & Zhu, 2024; Udupa et al., 2023).

From a practical perspective, the findings emphasize the importance of bias audits, dataset diversification, participatory design, and stronger governance mechanisms grounded in human rights and epistemic justice (Ferrara, 2023; Karpouzis, 2024; He et al., 2024). Without such interventions, AI-generated representations risk reinforcing discriminatory perceptions and symbolic marginalization, particularly in underrepresented regions such as North Africa (Pasipamire & Muroyiwa, 2024; Yilma, 2025).

Although limited by its qualitative design, single-country focus, and the rapidly evolving nature of generative AI technologies (Brutus et al., 2013; Tikhonova, 2024; Vice et al., 2023), the study provides a foundation for future comparative, experimental, and policy-oriented research.

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

Overall, the findings underscore that addressing bias in generative AI is not merely a technical challenge but a broader social, cultural, and political project requiring participatory, decolonial, and governance-centered approaches. By centering marginalized voices and promoting more inclusive AI systems, this research contributes to ongoing debates on ethical AI and offers practical guidance for policymakers, technology developers, and civil society actors working toward more equitable forms of digital representation.

REFERENCES

1. Aldahoul, N., Rahwan, T., & Zaki, Y. (2024). AI-generated faces influence gender stereotypes and racial homogenization. Scientific Reports, 15.
2. Al-Kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024). Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective. Informatics, 11, 58.
3. Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2022). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.
4. Georgiou, M. (2018). Does the subaltern speak? Migrant voices in digital Europe. Popular Communication, 16, 45 - 57.
5. Ghosh, S., Venkit, P., Gautam, S., Wilson, S., & Caliskan, A. (2024). Do Generative AI Models Output Harm while Representing Non-Western Cultures: Evidence from A Community-Centered Approach. ArXiv, abs/2407.14779.
6. Gillespie, T. (2024). Generative AI and the politics of visibility. Big Data & Society.
7. Hart, C. (2025). Using generative AI to investigate the multimodal representation of refugees and migrants. Discourse & Society.
8. He, Z., Su, J., Chen, L., Wang, T., & Lc, R. (2024). 'I Recall the Past': Exploring How People Collaborate with Generative AI to Create Cultural Heritage Narratives. Proceedings of the ACM on Human-Computer Interaction, 9, 1 - 30.
9. Herrie, M., Maleve, N., Philipsen, L., & Staunæs, A. (2024). Democratization and generative AI image creation: aesthetics, citizenship, and practices. AI & SOCIETY, 40, 3495 - 3507.
10. Inuwa-Dutse, I. (2023). FATE in AI: Towards Algorithmic Inclusivity and Accessibility. Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization.
11. Jääskeläinen, P., Sharma, N., Pallett, H., & Åsberg, C. (2025). Intersectional analysis of visual generative AI: the case of stable diffusion. Ai & Society, 40, 4341 - 4362.
12. Jenks, C. (2024). Communicating the cultural other: trust and bias in generative AI and large language models. Applied Linguistics Review, 16, 787 - 795.
13. Kaneti, M. (2023). What can AI see? The image of the 'migrant' in the era of AI post-visualization. Journal of Global Ethics, 19, 307 - 322.
14. Karpouzis, K. (2024). Plato's Shadows in the Digital Cave: Controlling Cultural Bias in Generative AI. Electronics.
15. Kay, J., Kasirzadeh, A., & Mohamed, S. (2024). Epistemic Injustice in Generative AI. ArXiv, abs/2408.11441. .

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

16. Kerr, A., Barry, M., & Kelleher, J. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. Big Data & Society, 7.
17. Mbalaka, B. (2023). Epistemically violent biases in artificial intelligence design: the case of DALLE-E 2 and Starry AI. Digital Transformation and Society.
18. Mei, B., & Zhu, H. (2024). The Ethical Dilemma and Governance Path of Generative AI in Cultural Communication. Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Education.
19. Óhéigeartaigh, S., Whittlestone, J., Liu, Y., Zeng, Y., & Liu, Z. (2020). Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance. Philosophy & Technology, 33, 571 - 593.
20. Olier, J., & Spadavecchia, C. (2022). Stereotypes, disproportions, and power asymmetries in the visual portrayal of migrants in ten countries: an interdisciplinary AI-based approach. Humanities and Social Sciences Communications, 9, 1-16.
21. Pansoni, S., Tiribelli, S., Paolanti, M., Frontoni, E., & Giovanola, B. (2023). Design of an Ethical Framework for Artificial Intelligence in Cultural Heritage. 2023 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS), 1-5.
22. Pasipamire, N., & Muroyiwa, A. (2024). Navigating algorithm bias in AI: ensuring fairness and trust in Africa. Frontiers in Research Metrics and Analytics, 9.
23. Qadri, R., Shelby, R., Bennett, C., & Denton, R. (2023). AI's Regimes of Representation: A Community-centered Study of Text-to-Image Models in South Asia. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.
24. Rettberg, J., & Wigers, H. (2025). AI-generated stories favour stability over change: homogeneity and cultural stereotyping in narratives generated by gpt-4o-mini. ArXiv, abs/2507.22445.
25. Ricaurte, P. (2022). Ethics for the majority world: AI and the question of violence at scale. Media, Culture & Society, 44, 726 - 745.
26. Sahin-Mencutek, Z., Barthoma, S., Gökalp-Aras, N., & Triandafyllidou, A. (2022). A crisis mode in migration governance: comparative and analytical insights. Comparative Migration Studies, 10.
27. Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. Ethics and Information Technology, 24.
28. Sbabo, A., & Bueno, A. (2025). Artificial intelligence and otherness. InMediaciones de la Comunicación.
29. Solaiman, I., Talat, Z., Agnew, W., Ahmad, L., Baker, D., Blodgett, S., Chen, C., Daum'e, H., Dodge, J., Evans, E., Friedrich, F., Gohar, U., Hooker, S., Jernite, Y., Luccioni, A., Lusoli, A.,
30. Sonni, A., Hafied, H., Irwanto, I., & Latuheru, R. (2024). Digital Newsroom Transformation: A Systematic Review of the Impact of Artificial Intelligence on Journalistic Practices, News Narratives, and Ethical Challenges. Journalism and Media.
31. Tiribelli, S., Pansoni, S., Frontoni, E., & Giovanola, B. (2024). Ethics of Artificial Intelligence for Cultural Heritage: Opportunities and Challenges. IEEE Transactions on Technology and Society, 5, 293-305.

Volume: 09, Issue: 04 July - August 2026

ISSN 2582-0176

32. Udupa, S., Maronikolakis, A., & Wisiorek, A. (2023). Ethical scaling for content moderation: Extreme speech and the (in)significance of artificial intelligence. Big Data & Society, 10.
33. Vice, J., Akhtar, N., Hartley, R., & Mian, A. (2023). Quantifying Bias in Text-to-Image Generative Models. IEEE Transactions on Dependable and Secure Computing, 22, 5658-5671.
34. Yilma, K. (2025). Ethics of AI in Africa: Interrogating the role of Ubuntu and AI governance initiatives. Ethics and Information Technology, 27.

