



ENCODING MARGINALIZATION: A DOCUMENT ANALYSIS OF HOW STATE AND CORPORATE AI ARCHITECTURES IMPACT RELIGIOUS MINORITIES IN THE GLOBAL NORTH AND GLOBAL SOUTH

Muhammad Kashif Mirza ¹, Muhammad Ali Bilgrami ²

DOI: <https://doi.org/10.63544/ijss.v5i5.334>

Affiliations:

¹ Independent Researchers,
Karachi, Pakistan
Email: qashif.mirza@gmail.com

² Independent Researchers,
Karachi, Pakistan

Corresponding Author's Email:

¹ qashif.mirza@gmail.com

Copyright:

Author/s

License:



Article History:

Received: 30.07.2026

Accepted: 25.08.2026

Published: 08.09.2026

Abstract

Background: The rapid global proliferation of artificial intelligence architectures is frequently framed as technologically neutral, obscuring how these systems inherently encode and operationalize structural biases against religious minorities. While the marginalization of these communities through digital surveillance and automated moderation is a universal phenomenon, the specific mechanisms driving this oppression diverge significantly across geographical contexts.

Objectives: This study aims to systematically extract, classify, and compare the regulatory language and structural frameworks governing AI deployment across the Global North and the Global South, identifying how state and corporate architectures disparately impact religious minority rights.

Methods: This study employed a qualitative documentary analysis, systematically reviewing state legislative texts, corporate AI policy guidelines, and triangulated human rights baseline reports spanning 2015 to 2026.

Results: The analysis reveals a fundamental hemispheric bifurcation in digital marginalization. In the Global North, oppression is predominantly covert, mediated by corporate algorithmic bias, opaque content moderation, and ostensibly neutral predictive policing architectures that inadvertently criminalize minority identities. Conversely, in the Global South, marginalization is heavily characterized by the overt state weaponization of imported biometric and surveillance tools, facilitated by deliberate domestic regulatory voids and broad national security exemptions. Despite these divergent mechanisms, both hemispheres share a structural absence of religious-rights-centered AI governance frameworks.

Conclusion: The study concludes that the systemic marginalization of religious minorities is a structural feature, not an anomaly, of current global AI deployment. Countering localized digital repression requires shifting the focus from generalized technological ethics to robust, context-specific regulatory frameworks that explicitly enforce the freedom of religion or belief against both corporate algorithmic negligence and authoritarian state mandates.

Keywords: Artificial Intelligence, Religious Minorities, Predictive Policing, Algorithmic Bias, Digital Repression

1. Introduction

1.1 Background

The rapid global proliferation of artificial intelligence (AI) and automated decision-making architectures has fundamentally restructured state security, public administration, and digital communication. These technologies are frequently framed within prevailing discourse as mathematically objective and politically neutral tools for optimizing governance and moderating vast digital ecosystems. However, this assumption of technological neutrality obscures embedded sociotechnical biases. AI architectures, ranging from predictive policing algorithms to automated content moderation systems, inherently encode the historical, structural, and political prejudices of both their training data and their developers (Council of Europe, 2018). The United Nations High Commissioner for Human Rights (2021) has issued an authoritative international warning that biometric technologies and AI systems carry profound, inherent risks of



exacerbating systemic discrimination against marginalized populations. Within this landscape, religious minorities represent a uniquely vulnerable demographic, frequently targeted through both algorithmic misclassification and deliberate digital surveillance under the guise of public safety (Shaheed, 2022). As advanced surveillance and moderation tools scale globally, the intersection of AI deployment and freedom of religion or belief (FoRB) has emerged as a critical frontier in human rights.

1.2 Statement of the Problem

Religious minorities are systematically marginalized by the deployment of AI architectures. Yet the structural mechanisms driving this marginalization remain fundamentally distinct across geopolitical contexts, and are rarely subjected to comparative evaluation. In the Global North, the marginalization of religious minorities is predominantly mediated by corporate algorithmic bias and ostensibly neutral predictive architectures. Northern state and corporate entities deploy systems, such as automated content moderation and risk-assessment algorithms, that frequently rely on structurally flawed datasets, unintentionally coding minority religious practices as extremist or toxic (Selbst, 2017; Jigsaw, 2022). Conversely, in the Global South, marginalization is heavily characterized by the overt state weaponization of imported AI and biometric surveillance tools. Authoritarian and hybrid regimes actively codify digital repression through sweeping cybersecurity legislation and mandatory biometric tracking, expressly targeting religious minority communities to enforce ideological compliance and monitor dissent (Digital Rights Foundation, 2022; Human Rights Watch, 2019). The central problem, however, is that despite this profound divergence, covert corporate bias in the North versus overt state-directed repression in the South, both hemispheres suffer from a critical structural absence of FoRB-centered AI governance, allowing digital oppression to operate largely unchecked.

1.3 Significance of the Topic

Understanding these divergent regulatory and operational architectures is critical for human rights advocates, policymakers, and civil society practitioners, particularly within South Asia. The United States Commission on International Religious Freedom (2024) notes an expanding global trend of digital authoritarianism, wherein surveillance architectures developed in the Global North or by authoritarian superpowers are exported to states with weak democratic safeguards. For advocates in Pakistan and the broader Global South, dissecting how these architectures function provides a crucial, evidence-based framework to anticipate and counter the localized importation of digital repression tactics. By demonstrating how AI reinforces existing political economies of marginalization, this study equips civil society with the theoretical grounding necessary to demand equitable, rights-based digital governance and to challenge the weaponization of domestic regulatory voids. Furthermore, mapping the distinct mechanisms of harm clarifies the accountability gap, illustrating why Northern corporate self-regulation is wholly insufficient to protect vulnerable groups from Southern state coercion.

1.4 Research Objective

The primary research objective of this study is to systematically extract, classify, and compare the regulatory language governing AI surveillance and moderation architectures across selected Global North (e.g., North America, Western Europe) and Global South (e.g., South Asia, East Asia) jurisdictions. Additionally, the study seeks to map the documented impacts of these divergent AI architectures on the freedom, privacy, and security of religious minorities using triangulated human rights conditions reports. Finally, the study aims to identify the specific structural mechanisms that facilitate the systemic marginalization of these communities in the absence of robust ethical safeguards.

1.5 Aim of the Study

The overarching aim of this study is to map the structural mechanisms, distinguishing between covert corporate-encoded biases and overt state-directed digital repression, that facilitate digital oppression globally. By evaluating the documented impacts of these systems on religious minorities, the study seeks to shift the global AI governance discourse beyond generalized data privacy and racial bias, anchoring it firmly in the protection of FoRB and the prevention of technologically enabled religious persecution.

1.6 Contribution and Scope of the Study

This study contributes a novel comparative policy evaluation to the fields of human rights and technology governance by directly auditing state regulatory frameworks and corporate policies side-by-side.



Its scope focuses strictly on the intersection of AI architectures and religious minority rights. The analysis deliberately excludes the broader impacts of AI on racial or socio-economic majorities, as well as pure computer science audits of machine learning code, in favor of a structural, policy-level evaluation. Geographically, the scope encompasses a cross-hemispheric comparison, utilizing public policies, corporate guidelines, and NGO documentation active between 2015 and 2025 to trace how global architectures disparately impact marginalized faith communities.

1.7 Organisation of the Study

The remainder of this study is organized as follows. Section 2 reviews the theoretical and empirical literature surrounding digital colonialism, surveillance capitalism, and AI bias, establishing the conceptual framework and articulating the core propositions. Section 3 details the research methodology, outlining the systematic qualitative document review protocol and the thematic coding framework utilized. Section 4 presents the results and discussion, featuring a descriptive analysis of the differing policy landscapes, an inferential analysis of structural marginalization patterns, an evaluation of the propositions, and an empirical estimation of the documented impact footprint. Finally, Section 5 synthesizes the major findings and provides actionable recommendations for policymakers and corporate entities to embed robust FoRB safeguards into global AI governance.

2. Literature Review

2.1 Theoretical Literature

The deployment of artificial intelligence architectures across diverse geopolitical landscapes is increasingly analysed through the lenses of digital colonialism and decoloniality. Foundational scholarship argues that the extraction of data and the subsequent imposition of Western-designed AI systems upon the Global South constitute a contemporary form of algorithmic colonization (Birhane, 2020). This structural dynamic ensures that sociotechnical systems inherently reinforce historical power imbalances, marginalizing vulnerable populations by imposing external, often majoritarian, biases disguised as objective technological solutions (Mohamed, Png, & Isaac, 2020). Extending this structural critique specifically to religious identity, the framework of “algorithmic Orientalism” posits that AI architectures, including language models and image classifiers, actively embed and amplify colonial stereotypes against minority groups, notably Muslims (Dahman & Saou, 2025).

Complementing these critiques of corporate AI development, theories of state technological control provide a framework for understanding authoritarian deployments. The concept of “techno-docility” illustrates how repressive states seamlessly integrate traditional physical coercion with pervasive AI and biometric surveillance, engineering a unitary society by rendering minority religious practices functionally inaccessible (Topal, 2023). Across these theoretical paradigms, a critical evolution has been the growing recognition that the impact of artificial intelligence must be evaluated not merely as a privacy or free expression concern, but directly through the lens of freedom of religion or belief (FoRB) (Ashraf, 2022). The prevailing theoretical consensus establishes that AI systems are not neutral arbiters; they are sociotechnical structures that systematically operationalize the marginalization of distinct religious identities unless actively countered by rights-based ethical frameworks.

2.2 Empirical Literature

Empirical investigations into AI deployment consistently document the disparate impacts these architectures exert on marginalized communities, though the mechanisms vary markedly between hemispheric contexts. In the Global North, the literature predominantly highlights the role of corporate algorithmic bias and the ostensibly neutral application of predictive models. Studies evaluating big data policing indicate that predictive architectures operationalize historical prejudices, relying on structurally flawed datasets that encode systemic discrimination under a veneer of objective, data-driven security (Selbst, 2017; Richardson, Schultz, & Crawford, 2019). Commercial facial analysis algorithms exhibit severe intersectional accuracy disparities, systematically misclassifying darker-skinned and female faces at substantially higher error rates than lighter-skinned male faces (Buolamwini & Gebru, 2018). A separate line of inquiry finds that commercial large language models persistently associate specific religious identities, such as Muslims, with violence and terrorism in generated text (Abid, Farooqi, & Zou, 2021). When Northern state actors adopt these unregulated



corporate tools for law enforcement and border security, the resulting surveillance architectures disproportionately target religious and immigrant minorities without adequate due process (Council on American-Islamic Relations, 2023; Patel & Melendi, 2024). Furthermore, judicial deference to national security interests frequently shields these racialized security architectures from robust democratic accountability (Sinnar, 2025). Parallel empirical findings raise similar concerns about corporate content moderation systems, warning that automated architectures which fail to grasp linguistic or cultural context risk over-censoring minority speech (UN General Assembly, 2018). A separate content analysis documents extensive hashtagged trolling and emoji-based hate speech directed at Muslim social media users, illustrating these same platforms' parallel failure to curb human-authored religious hate content (Al-Rawi, 2022).

Conversely, empirical literature centered on the Global South documents the explicit state weaponization of AI. Extensive human rights monitoring has reverse-engineered mass surveillance architectures, such as China's Integrated Joint Operations Platform (IJOP), revealing how algorithms explicitly code minority religious behavior as suspicious to facilitate the mass internment of Uyghur Muslims (Aras & Günar, 2026; Human Rights Watch, 2018, 2019). Similar algorithmic state violence and biometric dispossession have been documented against Palestinian populations under military occupation (Amnesty International, 2023; Goodfriend, 2023). In South Asia, empirical baseline reports demonstrate how sweeping cybersecurity legislation and digital identification frameworks are deliberately weaponized by state authorities to track, censor, and prosecute religious minorities under the guise of public order and anti-blasphemy enforcement (Digital Rights Foundation, 2022; Ranaware & Mishra, 2026; South Asia Collective, 2022). Crucially, corporate AI architectures frequently intersect with and facilitate this state-directed repression; researchers have documented a connection between corporate algorithmic engagement models and the amplification of hate speech linked to violence against the Rohingya in Myanmar (UN Human Rights Council, 2018), though Schissler (2024) cautions that framing this primarily as algorithmic amplification of 'hate speech' may understate the role of in-group social dynamics in driving the violence. The literature also confirms the transnational dimension of this threat, noting the unregulated export of biometric surveillance tools from the Global North to authoritarian regimes in the Global South, effectively equipping repressive states with the infrastructure necessary to oppress marginalized religious demographics (Byler & Ketter, 2024; Privacy International, 2022).

2.3 Research Gap

Despite robust documentation of AI's detrimental impact on human rights, the existing literature exhibits significant structural gaps. First, the majority of AI bias scholarship centers on generalized racial or socio-economic discrimination; dedicated analyses focusing exclusively on religious minorities as distinct targets of AI architectures remain severely under-theorized. Second, few systematic, comparative qualitative document analyses directly audit state regulatory frameworks and corporate policies side-by-side across both the Global North and the Global South. Finally, while empirical reports note the transnational export of surveillance tools, few studies closely examine how deliberate domestic regulatory voids, particularly in South Asia, function structurally to bypass FoRB protections and enable the state weaponization of these technologies.

2.4 Conceptual Framework

This study employs a two-axis conceptual framework mapping the intersection of artificial intelligence governance and religious minority rights (see Figure 1) to address these gaps. The framework categorizes deployment architecture (Corporate/Implicit vs. State/Explicit) based independent variables and moderates these through geopolitical context of Global North in comparison with Global South environments. This matrix posits that while systemic marginalization of religious minorities serves as universal dependent outcome, structural mechanisms driving it diverge fundamentally by hemisphere. It is illustrated through the framework that the structural absence of FoRB-centered governance across both axes acts as the primary enabling condition for digital repression, rendering minority communities uniquely vulnerable to both corporate algorithmic negligence and authoritarian state mandates.

2.5 Propositions/Hypotheses Development

Derived from the theoretical consensus and the identified empirical gaps, this study advances three



qualitative propositions to guide the documentary analysis.

- **Proposition 1 (P1):** In the Global North, the marginalization of religious minorities through artificial intelligence is predominantly mediated by corporate algorithmic bias, opaque content moderation, and ostensibly neutral predictive architectures, rather than explicit, identity-based state directives.
- **Proposition 2 (P2):** In the Global South, the marginalization of religious minorities is heavily characterized by the overt state weaponization of imported AI and biometric surveillance tools, a mechanism facilitated by deliberate domestic regulatory voids and broad national security exemptions.
- **Proposition 3 (P3):** Across both hemispheres, the structural absence of religious-rights-centered AI governance frameworks functions as a binding constraint on minority protection, functioning as the primary structural enabler of digital oppression.

3. Research Methodology

3.1 Study Area

The study's geographical scope was defined as a cross-hemispheric comparative analysis, contrasting selected jurisdictions in the Global North and the Global South. The Global North was represented primarily by North America (the United States and Canada) and Western Europe (the European Union and the United Kingdom), reflecting regions dominated by major corporate AI developers and ostensibly rights-based regulatory frameworks. The Global South was represented primarily by South Asia (Pakistan, India and Sri Lanka), alongside relevant comparators in Southeast Asia (Myanmar, Indonesia) and East Asia (China). This selection provided the geopolitical contrast necessary to evaluate how different state and corporate ecosystems govern the deployment of artificial intelligence.

3.2 Data Collection Techniques

The study employed a systematic qualitative document review through a purposive sampling strategy. The evidentiary dataset comprised 57 secondary documents extracted across three primary classifications: official state legislative and regulatory texts (e.g., the European Union AI Act, Pakistan's Prevention of Electronic Crimes Act); corporate transparency reports, system cards, and acceptable use policies (e.g., documentation from OpenAI, Google, Meta, and Palantir); and triangulated international human rights monitor reports and judicial rulings (e.g., from Amnesty International, Human Rights Watch, and the United Nations). Data retrieval used targeted Boolean search strings applied across academic databases and institutional portals, combining technological terms such as "artificial intelligence," "algorithmic bias," and "biometric surveillance" with demographic and rights-based identifiers including "religious minorities," "freedom of religion," and "digital repression."

3.3 Duration of the Study

The temporal scope of the documentary evidence reviewed spanned from 2015 to 2026. This window was deliberately selected to capture the exponential global acceleration of commercial AI deployment, the drafting and implementation of foundational Northern regulatory frameworks, and the simultaneous, aggressive expansion of Southern digital control laws and biometric identification architectures. The systematic document retrieval, coding, and comparative synthesis phases were conducted sequentially during the Jan 2026–Jul 2026 study period.

3.4 Theoretical Model

The study employed a structural evaluation model anchored in a Freedom of Religion or Belief (FoRB) perspective, integrated with decolonial theories of digital colonialism and algorithmic state violence. This theoretical model positioned state and corporate AI regulatory architectures as the independent structural variables and the systemic marginalization of religious minorities as the dependent outcome. The model posited that the realization or violation of minority rights in digital spaces is not an accidental technological byproduct, but a structural outcome heavily mediated by the specific geopolitical governance environment (Global North versus Global South). Consequently, the evaluation focused on structural mechanisms and regulatory voids rather than individual algorithmic anomalies.

3.5 Empirical Model

The analytical process followed a multi-stage thematic coding and comparative policy evaluation framework. Extracted texts were initially subjected to a rigorous qualitative coding protocol; the analysis did



not rely on quantitative algorithmic auditing or machine learning code reviews. Data was categorized into predefined thematic nodes: implicit Northern corporate bias, explicit Southern state repression, transnational regulatory voids, and documented impact footprints. Following the coding phase, a comparative policy evaluation was executed by mapping the encoded themes against a two-axis analytical matrix. This matrix contrasted the operational definitions of “security,” “extremism,” and “moderation” within the legal and corporate texts to trace how discriminatory biases were formally encoded. In the final analytical stage, these descriptive policy mappings were triangulated against empirical NGO and human rights baseline reports to verify real-world harms. This synthesis enabled a rigorous evaluation of the study’s three core qualitative propositions, anchoring the assessment of policy architectures strictly in their documented impacts on marginalized faith communities.

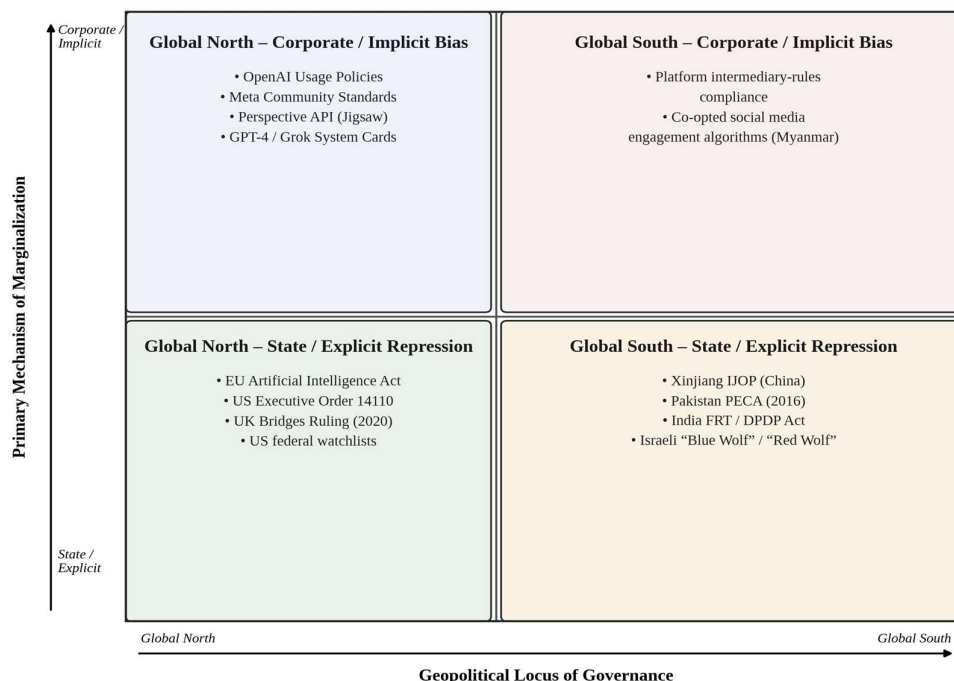
4. Results and Discussion

4.1 Contextual and Documentary Baseline

The documentary dataset comprises 57 accepted and utilized references categorized into three operational pillars: state legislative and regulatory instruments, corporate policy frameworks and technical system cards, and triangulated human rights conditions reports alongside judicial precedents. The geographical distribution reflects a deliberate cross-hemispheric comparative design. The Global North corpus is anchored by federal legislative and regulatory instruments from the United States (e.g., Exec. Order No. 14110, 2023, revoked in January 2025 and superseded by Exec. Order No. 14179; the proposed but never-enacted Algorithmic Accountability Act, 2023), the European Union (Regulation EU 2024/1689), the United Kingdom (Online Safety Act, 2023), and Canada (the since-lapsed Bill C-27, 2022, which died at prorogation in January 2025 without being enacted, leaving Canada without a federal AI statute). This corpus is complemented by foundational policies from leading AI developers (e.g., OpenAI, Google, Meta, xAI, Microsoft). The Global South corpus is grounded in statutory cybersecurity and data frameworks across South Asia, specifically Pakistan’s Prevention of Electronic Crimes Act (2016) and India’s Digital Personal Data Protection Act (2023). This corpus is supplemented by regional directives across East and Southeast Asia, including China’s Algorithmic Recommendation Provisions (Cyberspace Administration of China, 2022) and the Xinjiang De-Extremification Regulations (2018).

Figure 1

Two-Axis Taxonomy of AI Governance Architectures and Mechanisms of Religious Minority Marginalization, Global North and Global South





Note. Diagram constructed from the study's two-axis conceptual framework (Section 2.4) and the documentary taxonomy of sampled AI governance instruments (Section 4.1).

4.2 Descriptive Analysis: AI Architectures and Policy Landscapes

A comparative audit of statutory texts and corporate governance policies reveals a fundamental divergence in regulatory design between the Global North and the Global South (see Table 1).

Table 1

Comparative Matrix of State Regulatory and Corporate Governance Architectures

Governance Dimension	Global North Frameworks	Global South Frameworks
Primary Regulatory Focus	Risk-based classification, consumer data privacy, algorithmic transparency, and commercial market compliance (e.g., EU AI Act; US Algorithmic Accountability Act, proposed but not enacted).	National security, regime stability, state-mandated content removal, and centralized digital sovereignty (e.g., Pakistan PECA; China CAC Provisions; India IT Rules).
Operational Mechanism	Voluntary corporate ethics, self-regulatory red-teaming, model cards, and developer acceptable use policies (e.g., OpenAI, Google, Meta, xAI).	Statutory surveillance mandates, mandatory biometric registration, localized data storage, and intermediary liability coercion (e.g., Khine, 2023; SAFEnet, 2022).
Treatment of Religious Identity	Addressed as an abstract protected demographic category within broad non-discrimination and anti-bias principles; lacks explicit FoRB operationalization (e.g., Exec. Order No. 14110, since revoked; Google, 2023).	Frequently weaponized or securitized; religious behaviors, attire, or speech are formally coded as extremism, public disorder, or blasphemy (e.g., Xinjiang Regulations, 2018; PECA, 2016).
Enforcement and Oversight	Independent administrative agencies, risk assessments, civil society monitoring, and judicial review (e.g., Mitchell et al., 2019; Veale & Borgesius, 2021).	Unilateral executive commissions, military directives, broad state intelligence exemptions, and weakened judicial oversight (e.g., Centre for Policy Alternatives, 2024; India DPDP Act, 2023).

Note. Synthesis derived from systematic qualitative document analysis of state statutory acts and corporate governance frameworks (2015–2026).

In the Global North, governance architectures organize AI regulation around consumer risk categorization and corporate self-regulation. The European Union's Artificial Intelligence Act (Regulation EU 2024/1689) exemplifies this approach by establishing a harmonized, risk-based framework that prohibits specific biometric categorization systems inferring sensitive attributes such as religious beliefs, while mandating transparency audits for high-risk deployments. In the United States, Executive Order 14110 (2023), revoked in January 2025 and superseded by Executive Order 14179, had directed corporate developers to conduct impact assessments on automated decision systems, while the proposed but never-enacted Algorithmic Accountability Act (2023) sought to impose similar obligations through statute. However, corporate policy texts, such as OpenAI's Usage Policies (2025), Google's AI Principles (2023), and Meta's Community Standards (2024), rely on internal safety alignment, model cards, and automated filters to mitigate bias. These self-regulatory mechanisms consistently treat religious identity as a broad demographic variable rather than as an active human rights dimension vulnerable to algorithmic erasure (Mitchell et al., 2019).

Conversely, in the Global South, regulatory frameworks are explicitly structured around state sovereignty, national security, and digital control. In Pakistan, the Prevention of Electronic Crimes Act (2016) grants executive agencies sweeping statutory authority to censor online speech and monitor telecommunications under broad public order and anti-blasphemy provisions. India's regulatory landscape, anchored by the Information Technology Intermediary Guidelines (2021) and the Digital Personal Data Protection Act (2023), effectively compels private platforms to automate content takedowns while granting state intelligence agencies far-reaching, unchecked exemptions from data privacy mandates. In East Asia, China's Cyberspace Administration Provisions (2022) legally encode state ideological conformity into recommendation algorithms, while regional statutes such as the Xinjiang De-Extremification Regulations (2018) explicitly define non-violent religious practices as ideological extremism, establishing the statutory



basis for pervasive automated tracking.

4.3 Inferential Analysis: Structural Patterns of Marginalization

Linking these descriptive regulatory texts to systemic outcomes uncovers two distinct mechanisms of religious marginalization. In the Global North, marginalization operates through implicit, corporate-mediated algorithmic bias and the uncritical deployment of predictive security architectures. Foundational machine learning models and commercial toxicity APIs frequently encode historical sociotechnical prejudices, systematically misclassifying Islamic terminology or religious discourse as violent or toxic (Abid et al., 2021; Jigsaw, 2022). When Northern law enforcement and immigration authorities adopt these ostensibly neutral predictive architectures, the systems translate historical policing data into disparate impacts against religious minorities, subjecting Muslim and immigrant populations to opaque watchlists and predictive profiling without procedural transparency (Selbst, 2017; Patel & Melendi, 2024; Sinnar, 2025).

In the Global South, marginalization operates through explicit, state-directed digital repression and techno-docility. Authoritarian and majoritarian regimes actively weaponize imported biometric technologies and AI-driven platforms to enforce societal obedience and dismantle minority religious life (Topal, 2023). In China, police algorithms in Xinjiang automatically process facial recognition streams and communication metadata to flag routine religious conduct, such as daily prayer or religious attire, as indicators of radicalization, directly facilitating mass arbitrary detention (Aras & Günar, 2026; Human Rights Watch, 2019). In South Asia, broad cyber-offense definitions are leveraged alongside unregulated facial recognition dragnets to suppress minority political dissent and monitor religious assemblies (Digital Rights Foundation, 2022; Ranaware & Mishra, 2026). Furthermore, corporate profit models intersect with this repression. Northern platform algorithms optimized for engagement have been linked to the amplification of hate speech associated with violence against vulnerable minorities in Myanmar and South Asia (UN Human Rights Council, 2018), a connection Schissler (2024) complicates by locating the mechanism partly in in-group dynamics rather than algorithmic amplification alone.

4.4 Results of Propositional Evaluation

The documentary analysis systematically confirms the three qualitative propositions:

Proposition 1 (Northern Corporate Bias): Strongly Supported. The evidence confirms that Northern marginalization is predominantly mediated by corporate algorithmic bias, opaque content moderation, and predictive policing architectures (Al-Rawi, 2022; Buolamwini & Gebru, 2018; Richardson et al., 2019). Explicit, identity-based state discrimination is at times restrained by court interventions and constitutional privacy challenges, as seen in judicial rulings halting unlawful biometric surveillance and predictive profiling (e.g., *Bundesverfassungsgericht*, 2023; *R (Bridges) v Chief Constable of South Wales Police*, 2020).

Proposition 2 (Southern State Weaponization): Strongly Supported. The documentary record verifies that Southern regimes overtly weaponize biometric tracking and cybersecurity mandates against religious minorities (Amnesty International, 2023; Goodfriend, 2023). This weaponization is directly enabled by deliberate domestic statutory voids and broad national security exemptions (e.g., Internet Freedom Foundation, 2023).

Proposition 3 (Shared Transnational Regulatory Void): Strongly Supported. Across both hemispheres, existing governance frameworks fail to establish dedicated safeguards for Freedom of Religion or Belief (FoRB) in AI deployment. Northern corporate self-regulation relies on abstract bias metrics that ignore complex religious nuances (Google, 2024; OpenAI, 2025), and even the European Union's more comprehensive statutory approach lacks FoRB-specific safeguards despite its broader biometric protections (Veale & Borgesius, 2021). Southern frameworks, meanwhile, exploit these voids to institutionalize state surveillance (Shaheed, 2022).

4.5 Empirical Estimation: Documented Impact Footprint

Triangulating policy architectures against human rights documentation reveals the severity of real-world harms experienced by religious minorities (see Table 2).



Table 2

Triangulated Empirical Impact Footprint of AI Deployments on Religious Minorities

Target Community & Jurisdiction	AI Architecture Deployed	Documented Real-World Impact	Key Evidentiary Sources
Uyghur Muslims (Xinjiang, China)	Integrated Joint Operations Platform (IJOP); mass facial recognition; predictive policing.	Automated flagging of religious behaviours (e.g., fasting, Qur’anic study) leading to mass arbitrary internment, movement restrictions, and forced cultural assimilation.	Aras & Günar (2026); Australian Strategic Policy Institute (2019); Human Rights Watch (2018, 2019).
Rohingya Muslims (Myanmar)	Corporate recommendation algorithms; automated social media engagement pipelines.	Algorithmic amplification of genocidal hate speech and coordinated disinformation campaigns, materially facilitating military-led ethnic cleansing.	Schissler (2024); UN Human Rights Council (2018).
Palestinian Population (West Bank / East Jerusalem)	“Red Wolf” and “Blue Wolf” automated facial recognition checkpoints; biometric tracking dragnets.	Restriction of movement, automated spatial segregation, harassment at religious sites, and algorithmic dispossession.	Amnesty International (2023); Byler & Ketter (2024); Goodfriend (2023).
Religious Minorities (Pakistan & India)	State-mandated cyber monitoring; unregulated facial recognition; automated social media takedowns.	Targeted online harassment, state-directed censorship of minority religious expression, and wrongful criminal detentions under anti-blasphemy and cyber-crime laws.	Digital Rights Foundation (2022); Ranaware & Mishra (2026); South Asia Collective (2022).
Muslim Minorities (United States & Western Europe)	Automated border screening algorithms; federal watchlist matching; predictive policing software.	Disproportionate airport detentions, unconstitutional profiling, travel restrictions, and systemic stigmatization through flawed algorithmic matching.	Council on American-Islamic Relations (2023); Fattoumi (2026); Patel & Melendi (2024); Sinnar (2025).

Note. Empirical footprint synthesized from triangulated NGO monitoring, UN fact-finding missions, and judicial findings (2018–2026).

4.6 Discussion

When interpreted through the theoretical literature, these findings establish that the marginalization of religious minorities through AI is a structural manifestation of digital colonialism and algorithmic Orientalism. As theorized by Birhane (2020) and Mohamed et al. (2020), Northern corporate architectures extract global data while exporting models embedded with Western socio-cultural assumptions. This structural dynamic produces “algorithmic Orientalism,” wherein language models and automated content filters persistently associate Islamic identity with criminality and terrorism (Abid et al., 2021; Dahman & Saou, 2025). In the Global South, this dynamic interacts catastrophically with authoritarian governance. The export of unchecked surveillance architectures creates a transnational pipeline of repression, wherein technologies developed or trained in the Global North are adapted by Southern states to engineer techno-docility and crush religious pluralism (Privacy International, 2022; Topal, 2023).

Crucially, the empirical data challenges the simplistic assumption that technological harm in the Global North is purely accidental while Southern harm is entirely deliberate. Northern state architectures, such



as automated border screening and Olympic algorithmic surveillance, demonstrate a persistent tendency toward state surveillance creep that disproportionately harms racialized religious minorities. Judicial restraint on this creep is inconsistent; however, France's Constitutional Council upheld its Olympic surveillance law despite a comparable constitutional challenge, even as courts elsewhere have struck down similar measures (see Section 4.4) (La Quadrature du Net, 2023; Statewatch, 2023).

4.7 Analysis

The analysis demonstrates that international AI governance frameworks suffer from a foundational blind spot regarding Freedom of Religion or Belief. While global policy initiatives focus heavily on commercial data privacy, intellectual property, and generalized racial bias, the specific intersection of technology and religious persecution remains critically under-regulated (Ashraf, 2022; Shaheed, 2022).

For policymakers and civil society in Pakistan and South Asia, this comparative evaluation exposes the profound inadequacy of relying on corporate self-regulation or imported AI standards. Northern corporate policies (e.g., OpenAI, Meta, Google) are structurally incapable of protecting Southern minorities from state-coerced takedowns, local-language hate amplification, or authoritarian biometric dragnets. Addressing this crisis demands a structural paradigm shift. Human rights advocates must move beyond general ethical AI manifestos and demand legally binding, FoRB-centered impact assessments that prohibit the deployment of biometric surveillance and predictive policing tools against vulnerable religious communities in both hemispheres (UN High Commissioner for Human Rights, 2021; United States Commission on International Religious Freedom, 2024).

5. Conclusion and Recommendations

5.1 Major Findings

This documentary analysis yielded four principal findings regarding the intersection of artificial intelligence architectures and religious minority rights. First, AI-driven marginalization in the Global North is predominantly covert, operationalized through corporate algorithmic bias, opaque content moderation, and ostensibly neutral predictive architectures that inadvertently criminalize minority identities. Second, in the Global South, this marginalization is overt and state-directed; authoritarian and majoritarian regimes actively weaponize imported biometric technologies and broad cybersecurity legislation to systematically monitor, censor, and dispossess religious minority communities. Third, despite these divergent mechanisms, both hemispheres share a profound transnational regulatory void characterized by the structural absence of Freedom of Religion or Belief (FoRB) safeguards in AI governance. Finally, the empirical footprint of these architectures demonstrates severe real-world harms, ranging from digital censorship and unconstitutional profiling in the Global North to mass internment and state-aligned ethnic violence in the Global South.

5.2 Conclusion

This study set out to determine how state and corporate AI regulatory architectures disparately encode and operationalize the marginalization of religious minorities across hemispheres. It concludes that while the systemic marginalization of religious minorities is a universal outcome of unregulated AI, the structural mechanisms driving it are fundamentally bifurcated by geography. The Global North relies on a paradigm of corporate algorithmic negligence and self-regulation, which fails to capture religious nuances, whereas the Global South leverages deliberate legislative voids and explicit state-directed biometric weaponization. Ultimately, the absence of FoRB-centered governance functions as a binding structural constraint globally, rendering minority communities uniquely vulnerable to digital oppression. Mitigating these harms requires shifting the international discourse from generalized technological ethics to robust, context-specific regulatory frameworks that actively prohibit the technological targeting of vulnerable faith communities.

5.3 Recommendations

To address the structural vulnerabilities identified in this study, the following actions are recommended:

For Southern State Policymakers: Immediately halt the unregulated deployment of mass biometric surveillance systems. Amend broad cybersecurity and electronic crimes legislation to include explicit judicial oversight mechanisms and specific exemptions that protect religious expression and minority political dissent.

For Northern Corporate AI Developers: Move beyond generalized demographic bias mitigation by



mandating localized, context-specific human rights impact assessments. These assessments must explicitly evaluate model behavior and moderation algorithms against ForB standards before deploying systems or open-weight models into Global South markets.

For Civil Society and Human Rights Advocates: Shift advocacy strategies from demanding generalized algorithmic transparency toward campaigning for binding legal prohibitions on the use of predictive policing and facial recognition technologies in contexts where religious minorities face documented persecution.

5.4 Limitations

This study is subject to two primary limitations. First, the analysis relies heavily on secondary, English-language NGO reports and institutional framings, which may underrepresent localized, grassroots experiences of algorithmic harm. Second, because corporate AI systems operate as proprietary black boxes, primary technical audits were not possible, limiting the analysis to publicly available policy texts and documented downstream impacts.

5.5 Delimitations

The scope of this study was defined by two strict delimitations. First, the methodology deliberately excluded primary quantitative hypothesis testing and machine learning code reviews in favor of a structural, policy-level evaluation. Second, the analysis focused exclusively on the intersection of AI architectures and religious minority rights, intentionally bypassing the broader impacts of AI on racial or socio-economic majorities to maintain analytical depth.

5.6 Future Recommendations

Future scholarship should pursue two critical directions. First, researchers should conduct localized, community-led AI impact assessments in the Global South to capture the lived experiences of religious minorities navigating state digital repression. Second, future research should further examine the mechanisms of digital colonialism, specifically how the unchecked proliferation of Northern open-weight AI models empowers Southern authoritarian states to scale surveillance without corporate friction.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Data Availability

The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

References

- Abid, A., Farooqi, M., & Zou, J. (2021). Persistent anti-Muslim bias in large language models. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 298–306. <https://doi.org/10.1145/3461702.3462624>
- Algorithmic Accountability Act of 2023, H.R. 5628, 118th Cong. (2023). <https://www.congress.gov/bill/118th-congress/house-bill/5628/text>
- Al-Rawi, A. (2022). Hashtagged trolling and emoji-fied hate against Muslims on social media. *Religions*, 13(6), 521. <https://doi.org/10.3390/rel13060521>
- Amnesty International. (2023). *Automated apartheid: How facial recognition fragments, segregates and controls Palestinians in the OPT*. <https://www.amnesty.org/en/documents/mde15/6701/2023/en/>
- Aras, İ., & Günar, A. (2026). China's digital surveillance and religious practices in the Xinjiang Uyghur



- Autonomous Region. *Kocatepe İslami İlimler Dergisi*, 9(Özel Sayı), 96–108.
<https://doi.org/10.52637/kiid.1853591>
- Ashraf, C. (2022). Exploring the impacts of artificial intelligence on freedom of religion or belief online. *The International Journal of Human Rights*, 26(5), 757–791.
<https://doi.org/10.1080/13642987.2021.1968376>
- Australian Strategic Policy Institute. (2019). *Mapping more of China's tech giants: AI and surveillance*.
<https://www.aspi.org.au/report/mapping-more-chinas-tech-giants>
- Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 1st Sess., 44th Parl., 2022. <https://www.parl.ca/legisinfo/en/bill/44-1/c-27>
- Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*, 17(2), 389–409.
<https://doi.org/10.2966/scrip.170220.389>
- Bundesverfassungsgericht. (2023, February 16). *Judgment of the First Senate of 16 February 2023 - 1 BvR 1547/19*.
https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/2023/02/rs20230216_1b_vr154719en.html
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91.
<https://proceedings.mlr.press/v81/buolamwini18a.html>
- Byler, D., & Ketter, C. (2024). On the travel of state crimes by algorithm: Chinese camera systems in Israel. *Made in China Journal*, 9(1), 18–21. <https://doi.org/10.22459/MIC.09.01.2024.02>
- Centre for Policy Alternatives. (2024). *The Online Safety Act: Implications for freedom of expression, surveillance, and minority rights in Sri Lanka*. <https://www.cpalanka.org/the-online-safety-act-implications-for-freedom-of-expression-and-minority-rights/>
- Council of Europe. (2018). *Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*. <https://rm.coe.int/study-on-the-human-rights-dimensions-of-automated-data-processing-techn/16807908b1>
- Council on American-Islamic Relations. (2023). *Twenty years too many: A review of the FBI's secret watchlist and automated screening architectures*. <https://www.cair.com/wp-content/uploads/2023/06/Watchlist-Report-2023-Final.pdf>
- Cyberspace Administration of China. (2022). *Provisions on the management of algorithmic recommendations for internet information services (Order No. 9)*. http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm
- Dahman, A., & Saou, H. (2025). Algorithmic Orientalism: A systems analysis of anti-Muslim bias in artificial intelligence. *RSIC*, 10(2), 13–28. <https://doi.org/10.55781/rsic.v10i2.541>
- Digital Personal Data Protection Act, 2023, Act No. 22 of 2023 (India). <https://www.meity.gov.in/writereaddata/files/Digital%20Personal%20Data%20Protection%20Act%202023.pdf>
- Digital Rights Foundation. (2022). *Surveillance, censorship and religious minorities online: A review of PECA enforcement in Pakistan*. <https://digitalrightsfoundation.pk/wp-content/uploads/2022/10/Policy-Brief-PECA-and-Freedom-of-Expression.pdf>
- Exec. Order No. 14110, 88 Fed. Reg. 75191 (Nov. 1, 2023). <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>
- Fattoumi, F. (2026). Beyond recognition: The complexities of biometrics and minority rights. *Open Journal of Social Sciences*, 14(3), 15–35. <https://doi.org/10.4236/jss.2026.143002>
- Goodfriend, S. (2023). Algorithmic state violence: Automated surveillance and Palestinian dispossession in Hebron's Old City. *International Journal of Middle East Studies*, 55(3), 543–561.
<https://doi.org/10.1017/S0020743823000879>
- Google. (2023). *Artificial intelligence at Google: Our principles*. <https://blog.google/technology/ai/ai->



- [principles/](#)
- Google. (2024). *Generative AI prohibited use policy*. <https://policies.google.com/terms/generative-ai/use-policy>
- Human Rights Watch. (2018). "Eradicating ideological viruses": China's campaign of repression against Xinjiang's Muslims. <https://www.hrw.org/report/2018/09/09/eradicating-ideological-viruses/chinas-campaign-repression-against-xinjiangs>
- Human Rights Watch. (2019). *China's algorithms of repression: Reverse engineering a Xinjiang police mass surveillance app*. <https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance>
- Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E) (India).
https://www.meity.gov.in/writereaddata/files/Intermediary_Guidelines_and_Digital_Media_Ethics_Code_Rules-2021.pdf
- Internet Freedom Foundation. (2023). *Project Panoptic: Tracking the unregulated growth of facial recognition technology in India*. <https://panoptic.in/report-2023>
- Jigsaw. (2022). *Perspective API*. <https://perspectiveapi.com/>
- Khine, N. K. (2023). Digital rights in post-coup Myanmar: Enabling factors for digital authoritarianism. *Journal of Human Rights and Peace Studies*, 9(2), 186–216. <https://so03.tci-thaijo.org/index.php/HRPS/article/view/266778>
- La Quadrature du Net. (2023, March 23). *Algorithmic video surveillance and the 2024 Olympics: An unprecedented threat to fundamental freedoms*. <https://www.laquadrature.net/en/2023/03/23/olympic-games-surveillance-french-parliament-legalizes-algorithmic-video-surveillance/>
- Meta. (2024). *Community standards*. <https://transparency.meta.com/policies/community-standards/>
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. <https://doi.org/10.1145/3287560.3287596>
- Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659–684. <https://doi.org/10.1007/s13347-020-00405-8>
- Online Safety Act 2023, c. 50 (UK). <https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted>
- OpenAI. (2025). *OpenAI usage policies*. <https://openai.com/policies/usage-policies/>
- Patel, F., & Melendi, J. (2024). *Advances in AI increase risks of government social media monitoring*. Brennan Center for Justice. <https://www.brennancenter.org/our-work/analysis-opinion/advances-ai-increase-risks-government-social-media-monitoring>
- Prevention of Electronic Crimes Act, 2016, Act No. XL of 2016 (Pak.).
https://na.gov.pk/uploads/documents/1472635250_246.pdf
- Privacy International. (2022). *The global trade in biometric surveillance: Exporting oppression from the Global North to the Global South*. <https://privacyinternational.org/report/4981/global-surveillance-trade>
- R (Bridges) v Chief Constable of South Wales Police & Ors*, [2020] EWCA Civ 1058.
<https://www.bailii.org/ew/cases/EWCA/Civ/2020/1058.html>
- Ranaware, V., & Mishra, R. (2026). Unquestioned use of AI-based facial recognition technology in criminal investigations: Delhi riots lessons on rights and reliability. *Engineering Proceedings*, 143(1), 17.
<https://doi.org/10.3390/engproc2026143017>
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, OJ L, 2024/1689. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>
- Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *N.Y.U. Law Review Online*, 94, 15.
https://nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson_et-al-



[FIN.pdf](#)

- Schissler, M. (2024). Beyond hate speech and misinformation: Facebook and the Rohingya genocide in Myanmar. *Journal of Genocide Research*, 1–25. <https://doi.org/10.1080/14623528.2024.2375122>
- Selbst, A. D. (2017). Disparate impact in big data policing. *Georgia Law Review*, 52(1), 109–176. <https://digitalcommons.law.uga.edu/glr/vol52/iss1/4/>
- Shaheed, A. (2022). *Rights of persons belonging to religious or belief minorities in situations of conflict or insecurity: Report of the Special Rapporteur on freedom of religion or belief (A/HRC/49/44)*. UN Human Rights Council. <https://undocs.org/A/HRC/49/44>
- Sinnar, S. (2025). The Supreme Court & the unaccountable racialized security state. *Daedalus*, 154(4), 106–121. <https://doi.org/10.1162/DAED.a.945>
- South Asia Collective. (2022). *South Asia state of minorities report 2022: Weakening human rights commitments and its impact on minorities*. Minority Rights Group International. <https://minorityrights.org/publications/south-asia-state-of-minorities-report-2022/>
- Southeast Asia Freedom of Expression Network (SAFEnet). (2022). *Digital authoritarianism in Indonesia: Permenkominfo 5/2020, algorithmic governance, and minority vulnerability*. <https://safenet.or.id/en/2022/08/safenet-report-on-mr5-and-digital-authoritarianism/>
- Statewatch. (2023). *Automating the border: AI, biometrics, and the securitisation of migration in the European Union*. <https://www.statewatch.org/analyses/2023/automating-the-border/>
- Topal, R. (2023). A techno-docility approach to the repression and surveillance of Uyghurs. *Asian Survey*, 64(1), 143–167. <https://doi.org/10.1525/as.2023.2079647>
- UN General Assembly. (2018). *Promotion and protection of the right to freedom of opinion and expression (A/73/348)*. <https://undocs.org/A/73/348>
- UN High Commissioner for Human Rights. (2021). *The right to privacy in the digital age: Artificial intelligence, biometric technologies, and discrimination (A/HRC/48/31)*. UN Human Rights Council. <https://undocs.org/A/HRC/48/31>
- UN Human Rights Council. (2018). *Report of the independent international fact-finding mission on Myanmar (A/HRC/39/64)*. <https://undocs.org/A/HRC/39/64>
- United States Commission on International Religious Freedom. (2024). *Annual report 2024*. <https://www.uscifr.gov/annual-reports>
- Veale, M., & Borgesius, F. Z. (2021). Demystifying the Draft EU Artificial Intelligence Act, Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, 22(4), 97–112. <https://doi.org/10.9785/cr-2021-220402>
- Xinjiang Uyghur Autonomous Region People's Congress. (2018). *Xinjiang Uyghur Autonomous Region regulation on de-extremification*. China Law Translate. <https://www.chinalawtranslate.com/en/xinjiang-de-extremification-regulations-2018/>