AI Power and Discourse, vol. 2, núm. 2, 2026, pp. 1-10.

Algorithmic Empire: Generative AI and the Grammar of Global Subordination.

Agustin V. Startari.

Cita:

Agustin V. Startari (2026). *Algorithmic Empire: Generative AI and the Grammar of Global Subordination. AI Power and Discourse, 2 (2), 1-10.*

Dirección estable: https://www.aacademica.org/agustin.v.startari/249

ARK: https://n2t.net/ark:/13683/p0c2/sYs

Algorithmic Empire: Generative AI and the Grammar of Global Subordination

How Machine Discourse Normalizes Sanctions, Occupation, Intervention, and Dependency

Author: Agustin V. Startari

Author Identifiers

- ResearcherID: K-5792-2016
- ORCID: https://orcid.org/0009-0001-4714-6539
- SSRN Author Page: https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7639915

Institutional Affiliations

- Universidad de la República (Uruguay)
- Universidad de la Empresa (Uruguay)
- Universidad de Palermo (Argentina)

Contact

- Email: astart@palermo.edu
- Alternate: agustin.startari@gmail.com

Date: August 25, 2026

DOI

- Primary archive: **10.5281/zenodo.22096813**
- Secondary archive: 10.6084/m9.figshare.33329241
- SSRN: Pending assignment

Language: English

Series: *Grammars of Asymmetric Visibility: AI, Imperial Power, and the Syntax of Responsibility*

Word count: 10,081 (main text)

Keywords: algorithmic empire; generative AI; imperial grammar; geopolitical bias; asymmetric visibility; responsibility loss; global subordination; artificial intelligence; Palestine; Iran; sanctions; occupation; intervention; humanitarian discourse; sovereignty;

Global South; political agency; discourse analysis; AI governance; large language models; framing.

Abstract

This article introduces algorithmic empire as a formal condition of generative AI discourse in which inherited geopolitical hierarchies may be reproduced through the unequal distribution of agency, legitimacy, crisis, sovereignty, coercion, and responsibility. The concept does not treat every Western-centered output, security claim, humanitarian formulation, development program, sanction, or intervention as imperial by definition. It identifies a narrower and testable possibility: when generative systems summarize, explain, classify, translate, or moderate geopolitical material, dominant actors may retain stronger grammatical access to verbs of order, protection, mediation, enforcement, development, and stabilization, while subordinated societies may be represented more often as risks, crises, threats, failures, migration sources, sanction targets, humanitarian burdens, or intervention zones. The paper consolidates six mechanisms developed in the preceding articles of the series: responsibility loss, asymmetric visibility, sanctioned suffering, censorship without a censor, political-subjecthood reduction, and the administrative translation of force. It argues that these mechanisms can converge without requiring explicit ideological endorsement. A model can reproduce hierarchy through ordinary operations of compression, recontextualization, nominalization, agent deletion, lexical selection, safety framing, and source weighting.

Palestine and Iran function as principal analytical anchors because they expose different relations between sovereignty, external coercion, security discourse, sanctions, occupation, civilian harm, and institutional agency. The paper also situates these cases within a broader literature on data colonialism, digital colonialism, algorithmic colonization, decolonial AI, and geopolitical bias in large language models. Recent empirical work already demonstrates that framing can increase in LLM-generated summaries and that nation-level geopolitical biases vary by model, task, and context. Algorithmic empire therefore cannot be inferred from isolated outputs or model origin. It requires controlled comparison.

Methodologically, the article proposes an Imperial Grammar Reproduction Rate (IGRR) and an Algorithmic Empire Index (AEI). IGRR measures the proportion and direction of

clause-level transformations in which agency and legitimating action are preferentially retained for dominant actors while subordinated actors are preferentially converted into objects of crisis, management, dependency, or threat. AEI combines agency asymmetry, object-conversion asymmetry, coercion neutralization, sovereignty erasure, external-responsibility deletion, and humanitarian-management framing into a normalized structural index. Both measures are source-controlled and explicitly falsifiable. If apparent asymmetry is already present in source texts, occurs equally across actor classes, reverses under matched conditions, or disappears when genre and factual context are controlled, the stronger geopolitical interpretation must be narrowed. The paper concludes that AI governance should examine not only whether outputs are biased, inaccurate, hateful, or unsafe, but whether they preserve the grammatical conditions under which domination, coercion, political agency, and institutional responsibility can still be named.

1. Introduction: When Empire Becomes Syntax

Empire is easiest to recognize when domination speaks in the grammar of command. A state invades. An army occupies. A government imposes a blockade. A coalition launches an intervention. A sanctioning authority freezes assets or restricts trade. A platform removes speech. In these constructions, the political relation may remain controversial, but the syntax is comparatively transparent: an actor acts, an affected party receives the action, and a causal relation can be reconstructed from the sentence.

Contemporary domination does not always appear in that form. It can be represented as stabilization, deterrence, compliance, development, security management, humanitarian response, migration control, sanctions implementation, capacity building, peacekeeping, risk reduction, or rules enforcement. None of those terms is intrinsically illegitimate. Stabilization can be necessary. Humanitarian action can save lives. Sanctions can be adopted for defensible purposes. Security threats can be real. International law contains genuine obligations, permissions, and prohibitions that cannot be reduced to rhetoric. The analytical problem begins when these categories are distributed asymmetrically: when one actor repeatedly receives the grammar of legitimate action while another repeatedly receives the grammar of crisis, deficiency, threat, or administration.

This paper calls the resulting machine-mediated condition algorithmic empire. The term does not designate a territorial empire operated by artificial intelligence, nor does it claim that generative models possess imperial intentions. It names a representational pattern in which global hierarchy is reproduced through language generated, transformed, or ranked by computational systems. The relevant unit is not the ideology attributed to a model. It is the clause: who appears as subject, which verb is attached to that subject, who becomes object, which causal chain survives compression, whose sovereignty remains linguistically available, whose violence is represented as action, whose violence is represented as response, whose suffering is humanitarianized, and whose institutional responsibility remains recoverable.

The distinction matters because generative AI is increasingly used as an interpretive layer between sources and readers. Models summarize news, answer political questions, translate institutional language, generate executive briefs, moderate content, classify risk, explain sanctions, and restate military or diplomatic claims in apparently neutral prose. These operations are transformations, not transparent transfers. Summarization selects. Translation recontextualizes. Safety policies reorder salience. Retrieval privileges some source environments over others. Instruction tuning rewards certain answer forms. Even when every sentence is factually defensible, the distribution of agency can change.

The general proposition is supported by two established lines of evidence without being proved by either. First, machine-learning systems can inherit statistical regularities and social biases from human-produced language. Caliskan, Bryson, and Narayanan (2017) demonstrated that distributional semantics learned from ordinary corpora reproduce human-like associations. Bender et al. (2021) emphasized that large language models are built from vast sociocultural archives whose distributions cannot be treated as socially neutral. Blodgett et al. (2020) further argued that bias analysis in NLP requires explicit normative and social definitions rather than decontextualized metrics. These findings establish a mechanism of inheritance but do not by themselves establish geopolitical hierarchy.

Second, recent work has begun to measure political and geopolitical variation directly in LLM outputs. Choi et al. (2026), using United Nations Security Council records, found nation-level biases that varied across models, tasks, and contexts rather than following a single universal direction. Pacheco, Cavalini, and Comarela (2026) found different geopolitical tendencies in GPT-4o and DeepSeek-R1, with Western-centric patterns in the former and more explicit Chinese-state-aligned patterns in the latter. Pastorino and Moosavi (2026), examining large-scale summarization, found that LLM-generated summaries often displayed higher calibrated framing rates than human reference summaries, again with substantial variation across models and topics. These results are important precisely because they make a simplistic theory impossible. If geopolitical bias varies by system and task, algorithmic empire must be formulated as a measurable condition, not an assumed property of all generative AI.

The concept also enters an established critical literature. Kwet (2019) uses digital colonialism to describe infrastructural control by dominant technology corporations in the Global South. Couldry and Mejias (2019) develop data colonialism as a theory of social appropriation through continuous data extraction. Birhane (2020) describes algorithmic colonization in Africa as the imposition of external technological systems, values, and dependencies. Mohamed, Png, and Isaac (2020) articulate decolonial AI as a sociotechnical response to persistent relations of power. Ricaurte (2019) examines data epistemologies through the coloniality of power, while Adams (2021) asks whether AI can be decolonized at all. Muldoon and Wu (2023) extend the analysis by locating AI within a colonial matrix of power.

Algorithmic empire differs from these approaches in object and method. It does not primarily measure infrastructure ownership, data extraction, labor, dependency on cloud systems, linguistic underrepresentation, or corporate market concentration. Those conditions may shape the phenomenon, but the object here is narrower: the grammar produced at the interface. The question is whether generative systems reproduce a patterned allocation of political subjecthood. Put differently, digital colonialism concerns who controls the infrastructure; data colonialism concerns who appropriates social life as data; algorithmic colonization concerns whose systems, values, and dependencies are installed; algorithmic empire, as defined here, concerns how hierarchy becomes linguistically executable after those systems begin to speak.

The central research question is therefore: When generative AI systems summarize, explain, classify, moderate, or translate geopolitical discourse about occupation, sanctions, intervention, security, humanitarian crisis, development, migration, and conflict, do they preserve symmetric political agency, or do they reproduce global hierarchies through grammar, classification, and framing?

The hypothesis is conditional. Under controlled and comparable transformations, generative AI may assign dominant actors higher rates of grammatical agency, institutional rationality, legitimating action, and responsibility control while representing subordinated societies more frequently through crisis, threat, instability, dependency, humanitarian need, sanction status, or intervention eligibility. The existence, magnitude,

direction, and source of that asymmetry are empirical questions. If the asymmetry is absent, reversed, fully explained by source genre, or inherited without amplification from source texts, the theory must be restricted accordingly.

Two terms require discipline from the outset. Dominant and subordinated are relational, not civilizational labels. A state can be dominant in one relation and subordinated in another. Iran, for example, is a sovereign state with significant domestic and regional coercive capacity; in a sanctions relation with the United States and the international financial system, however, it is the target of materially superior external financial power. Israel is a small state by population but exercises occupation-related authority over Palestinians and possesses substantial military capability supported by a major-power alliance. The United States is globally dominant in military, financial, technological, and institutional domains but can still be represented as vulnerable or constrained in specific contexts. The unit of classification is therefore the relation under analysis, not a permanent moral identity attached to a country.

The same discipline applies to empire. The paper does not claim that every asymmetric relation is imperial or that every act by a dominant state is illegitimate. It asks whether generative systems reproduce a specific grammar historically associated with hierarchical rule: legitimate agency at the center, administrable crisis at the periphery, coercion as policy, domination as order, and suffering as a condition whose responsible actor becomes grammatically optional.

The machine does not need to conquer territory. The empirical question is whether it inherits the syntax that makes conquest, coercion, or dependency easier to read as order.

2. From Responsibility Loss to Algorithmic Empire

The six preceding papers in this series isolate mechanisms that appear different when examined separately. The first concerns responsibility loss: suffering remains visible while the actor causing it becomes grammatically optional (Startari, 2026a). The second concerns asymmetric visibility: multiple actors can be mentioned while agency, intention, and legitimacy are distributed unequally among them (Startari, 2026b). The third concerns sanctions: coercion can be translated into the impersonal language of pressure, restriction, compliance, or economic condition (Startari, 2026c). The fourth concerns platform governance: suppression can be represented as policy enforcement while the suppressing institution recedes from the grammar of censorship (Startari, 2026d). The fifth concerns political subjecthood: subordinated societies can remain highly visible as risk objects, humanitarian caseloads, unstable regions, or migration sources while losing the grammar of claim, decision, sovereignty, memory, and self-description (Startari, 2026e). The sixth concerns administrative translation of force: violence can remain semantically present while being redescribed as targeting, optimization, collateral estimation, risk management, validation, or operational procedure (Startari, 2026f).

Algorithmic empire is the proposed convergence of those mechanisms. Its minimal architecture can be represented as a sequence:

responsible actor -> institutional process
political subject -> risk or crisis object
coercive action -> policy instrument
violence -> operational procedure
suppression -> enforcement
external cause -> internal condition

No single transformation establishes imperial grammar. Institutions need processes; political actors can genuinely pose risks; sanctions are policy instruments; militaries require operational procedures; platforms enforce rules; and internal causes often

contribute to crisis. The stronger claim appears only when the transformations align systematically with geopolitical position.

Consider a simplified relation involving a dominant external actor D and a subordinated actor S. A symmetric representation can assign both sides explicit agency, preserve the external and internal causal chains affecting S, distinguish civilian populations from governments or armed groups, and state the legal or political status of the relation where relevant. An asymmetric representation may preserve D primarily as mediator, guarantor, stabilizer, defender, donor, sanctioning authority, rule enforcer, or security provider while representing S primarily as unstable, sanctioned, radicalized, failed, dependent, humanitarian, migratory, or threatening. The hierarchy is not contained in any one word. It emerges from the distribution.

Van Leeuwen's (2008) framework for representing social actors is useful here because agency can be displaced without simple deletion. Actors may be activated or passivated, individualized or assimilated, functionalized, institutionalized, abstracted, or represented through the instruments by which they act. A state can disappear behind “sanctions,” a military behind “operations,” a platform behind “policy,” or a financial architecture behind “market access.” Conversely, a subordinated population can be collectivized as “refugees,” “migrants,” “the unemployed,” “extremists,” or “humanitarian cases.” The result is not silence but reclassification.

Halliday and Matthiessen's (2014) account of grammatical metaphor adds a second mechanism. Processes can become nouns: sanctioning becomes sanctions; occupying becomes occupation management; restricting becomes restrictions; displacing becomes displacement; intervening becomes intervention; targeting becomes target generation. Nominalization is not inherently ideological. It is indispensable to academic, legal, bureaucratic, and technical prose. Its political significance depends on what argument structure disappears with it. “The United States imposed sanctions on X” requires an agent and an act. “Sanctions affected X” retains the coercive instrument but allows the sanctioning authority to vanish. “Economic conditions deteriorated in X” can remove the external instrument as well. Each sentence may be defensible in context. Across a corpus, however, the direction of compression can be measured.

The relation between compression and legitimacy is central. Generative systems are optimized to produce concise, coherent, useful answers. Concision rewards the deletion of what appears redundant. But redundancy is politically uneven. If a reader already treats a dominant institution as a normal source of regulation, repeatedly naming the institution may appear unnecessary. If a sanctioned or occupied population is already framed as the site of a problem, repeating its condition may appear informative. The model can therefore produce a highly efficient summary that preserves the crisis and deletes the actor structuring it.

This is one reason algorithmic empire cannot be reduced to misinformation. A false statement is comparatively easy to classify: a proposition does not match available evidence. Imperial grammar can operate through true propositions. “Iran faces shortages.” “Gaza requires humanitarian aid.” “The United States seeks regional stability.” “Israel cites security concerns.” “Migrants place pressure on services.” Each can be true under specified conditions. The analytic issue is whether the generated discourse also preserves who imposed constraints, who exercised control, who defined the security category, what legal status governs the relation, which local actors retain agency, and what alternative causal chains are present in the source.

The distinction also separates algorithmic empire from sentiment bias. A model need not speak positively about dominant states or negatively about subordinated societies. It can criticize a dominant actor while still granting that actor most of the verbs. It can express sympathy for a subordinated population while representing that population almost exclusively through suffering. Humanitarian concern can coexist with political passivation. Indeed, the first paper in this series identified precisely that structure: visibility of suffering does not guarantee visibility of responsibility (Startari, 2026a).

The same applies to neutrality. Machine discourse often presents institutional language through impersonal syntax, modal hedging, source balancing, and abstract categories. These features can reduce overt partisanship while still producing structural asymmetry. A sentence can be balanced at the level of evaluative adjectives and unbalanced at the level of transitivity. One side acts; the other experiences. One side “responds”; the other “attacks.” One side “raises security concerns”; the other “poses a threat.” One side “uses

sanctions to pressure"; the other "suffers economic isolation." Neutrality at the surface does not guarantee symmetry in the underlying grammar.

This paper therefore treats algorithmic empire as a second-order phenomenon. First-order mechanisms occur within individual clauses: agent deletion, passive voice, nominalization, object conversion, threat framing, humanitarian framing, sovereignty erasure, and responsibility displacement. Second-order hierarchy occurs when those mechanisms are distributed directionally across actor classes. The concept is justified only at the second level.

A useful diagnostic is the difference between visibility and role. Dominant actors can be less frequently mentioned and still retain stronger political agency when they appear. Subordinated actors can be mentioned constantly while remaining syntactically passive. Frequency is therefore insufficient. The relevant questions are role-conditioned: When actor D appears, what does it do? When actor S appears, what happens to it? What verbs attach to each? Which nouns replace their actions? Which causal chains remain visible? Which legal statuses are preserved? Which actions are described through the vocabulary of necessity or order?

This role-based approach also avoids a common error in critical AI research: declaring bias whenever an output conflicts with a preferred political interpretation. Blodgett et al. (2020) warn that bias metrics require a clearly articulated normative object. The present framework supplies one. It does not ask whether a model agrees with anti-imperial theory. It asks whether comparable actors receive systematically different grammatical treatment under controlled conditions. The criterion is representational symmetry under relevant factual difference, not ideological agreement.

The series can therefore be summarized through one cumulative proposition: responsibility loss explains how perpetrators disappear; asymmetric visibility explains how agency is unequally distributed; sanctions grammar explains how coercion becomes pressure; censorship without a censor explains how suppression becomes enforcement; digital dehumanization explains how societies remain visible without subjecthood; administrative translation of force explains how violence becomes procedure; algorithmic

empire asks whether these transformations align into a reproducible grammar of global hierarchy.

3. Dominant Powers as Agents of Order

Political hierarchy is partly a hierarchy of verbs. Institutions that possess recognized authority are disproportionately associated with actions such as regulate, enforce, stabilize, protect, mediate, authorize, assist, defend, monitor, sanction, develop, support, deter, contain, and secure. These verbs do more than describe behavior. They encode a relation between an actor and a normative environment. To enforce presupposes a rule. To stabilize presupposes instability. To protect presupposes a protected object and a threat. To mediate presupposes parties requiring mediation. To develop presupposes an object in need of development. To sanction presupposes an authority capable of imposing a penalty or restriction.

The paper calls these legitimating-action predicates when, in context, the verb presents an actor primarily as an organizer of order rather than as a source of coercion or contestable political action. The category is functional, not lexical. “The state protected civilians from attack” may be a straightforward factual description and should not be coded as hierarchy merely because protect appears. The relevant comparison asks whether equivalent acts by differently positioned actors are systematically lexicalized through different functions.

Three mechanisms are especially important. The first is response privilege. Force by a dominant actor can be framed as response, deterrence, defense, stabilization, or threat reduction, while force by a subordinated actor is framed as attack, aggression, escalation, militancy, or destabilization. The underlying events need not be equivalent, and the framework must preserve differences in initiation, scale, legal status, and target selection. But once those differences are controlled, a remaining lexical asymmetry becomes measurable. The issue is not that response language is always wrong. It is that response grammatically assigns prior causality elsewhere.

The second mechanism is institutional rationality. Dominant actors are often represented through procedures: they assess, review, coordinate, authorize, consult, impose measures,

or pursue objectives. This grammar makes their actions legible as policy. Subordinated actors can be represented more often through states: instability, unrest, radicalization, non-compliance, crisis, corruption, or humanitarian need. Policy is agentive; condition is descriptive. A dominant actor therefore appears as the producer of strategy while a subordinated actor appears as the environment to which strategy is applied.

The third mechanism is authority inheritance. A generative model may summarize an official source in the source's own institutional vocabulary without explicitly marking the vocabulary as attributed. “The government said it launched a limited operation to restore deterrence” can become “a limited operation was launched to restore deterrence.” The propositional content is similar, but attribution loss changes the status of “limited” and “restore deterrence” from claims to apparent properties. This is especially important in security discourse, where terms such as imminent threat, precision, defensive action, proportionality, and stabilization can carry legal or moral implications beyond their descriptive content.

The 2025 Israel-Iran conflict illustrates the problem without requiring unsupported claims about AI targeting. United Nations Security Council reporting recorded Israeli characterization of its opening strikes as precise and pre-emptive and linked them to an asserted imminent nuclear threat, while also recording Iranian retaliatory attacks and the effects of strikes on nuclear and other sites (United Nations Department of Political and Peacebuilding Affairs, 2025; International Atomic Energy Agency, 2025). A generative summary can preserve those labels as attributed claims, compare them with external reporting, and identify the actors who selected the sites. Or it can compress the relation into a grammar in which “Iran's nuclear threat was targeted” and “Iran retaliated.” The second structure allocates initiative before any explicit judgment about legality occurs.

A similar issue arises in United States sanctions discourse. The U.S. Treasury's Office of Foreign Assets Control describes sanctions through an extensive legal-administrative architecture of programs, licenses, exemptions, designated persons, humanitarian authorizations, and compliance rules (Office of Foreign Assets Control, n.d.). That architecture is real and legally relevant. It also produces a powerful grammar of administration: the United States authorizes, licenses, designates, exempts, permits, and

enforces. Iran appears as the sanctioned jurisdiction, designated actor environment, or compliance object. A summary that reproduces only that architecture can make the sanctions relation appear primarily as rule administration rather than as an exercise of external economic coercion with contested humanitarian effects.

The point is not to replace institutional vocabulary with accusatory vocabulary. It is to preserve both relations. A high-accountability sentence can say that the United States imposed sanctions, that OFAC maintains humanitarian exceptions and authorizations, that financial institutions nevertheless may over-comply, and that United Nations experts and peer-reviewed research have documented resulting difficulties in medical supply chains (United Nations Human Rights Council Special Procedures, 2022; Asadi-Pooya, Nazari, & Mirzaei Damabi, 2022; Yazdi-Feyzabadi et al., 2024). That sentence contains policy purpose, legal structure, external agency, uncertainty, and civilian consequence simultaneously.

Algorithmic empire becomes plausible when the generated version consistently retains the administrative legitimacy of the sanctioning actor while converting the target's experience into an endogenous condition: inflation, scarcity, isolation, institutional weakness, humanitarian difficulty. External agency does not have to disappear completely. It only has to become less grammatically central than the condition it helped produce.

Palestine provides an even sharper test because legal status is not merely rhetorical context. In its 2024 advisory opinion, the International Court of Justice analyzed Israel's policies and practices in the Occupied Palestinian Territory, the legal status of occupation, annexation, discrimination, and the Palestinian people's right to self-determination, while separate and dissenting opinions contested parts of the majority's reasoning (International Court of Justice, 2024). A generated answer that summarizes the situation solely as an “Israeli-Palestinian conflict” may be descriptively common yet legally under-specified in contexts where occupation is directly relevant. Conversely, an answer that treats every Israeli security action as reducible to occupation would also collapse legally distinct questions. Symmetry requires preserving the appropriate status of each relation rather than replacing law with generic balance.

The distribution of order verbs can therefore be tested at scale. For each actor class, a corpus can measure the frequency of verbs and frames associated with stabilization, protection, mediation, enforcement, development, assistance, legality, and rule maintenance. Those counts must then be compared with direct-action predicates such as strike, occupy, restrict, detain, block, surveil, sanction, arm, fund, and displace. A dominant actor should not receive a high hierarchy score merely for being described as a mediator when it is in fact mediating. The score rises when legitimating roles systematically replace or background direct coercive roles that are present in source material.

This distinction is crucial because a model may improve discourse. It may add responsible actors omitted by a source, preserve legal qualifiers, separate armed groups from civilian populations, or identify the institutional decision behind a nominalized process. Such transformations should lower the proposed index. The theory must allow AI to de-imperialize language as well as reproduce it.

The empirical question is therefore not whether dominant states are described as powerful. It is whether their power is preferentially grammaticalized as legitimate order while the political agency of those subject to that power is reduced.

4. Subordinated Societies as Objects of Crisis

If dominant actors occupy the grammar of order, subordinated societies can occupy the grammar of problem. They become refugee flows, unstable regions, failed states, sanction targets, humanitarian emergencies, extremist environments, development deficits, migration pressures, conflict zones, or security risks. The noun phrase may remain highly visible. What declines is political subjecthood.

This paper uses object conversion to describe that transformation. A political subject is represented as an object when discourse foregrounds what can be done to, for, with, or about it while reducing what it can claim, decide, remember, contest, negotiate, govern, or define. Object conversion is not dehumanization in the narrow sense of animalization or explicit contempt. It can occur in highly sympathetic language. A population may be represented as vulnerable, displaced, food-insecure, traumatized, or in urgent need of assistance. Those descriptions can be accurate and ethically necessary. The conversion occurs when they become the principal grammar through which the population exists.

The humanitarian object is therefore structurally ambiguous. Humanitarian categories make suffering visible and can mobilize protection. They can also depoliticize causality by organizing people around needs rather than claims. Famine risk, displacement, shelter need, medical shortages, and civilian casualties are indispensable facts. But a discourse can describe all of them while weakening the questions of who imposed restrictions, who exercised territorial control, who destroyed infrastructure, which armed actors attacked civilians, what legal obligations applied, and what political subjecthood the affected population retained. Humanitarian visibility is not equivalent to political visibility.

Palestine is the strongest anchor because the problem can be tested against explicit legal categories. The International Court of Justice's 2024 advisory opinion treats the Palestinian people as the holder of a right to self-determination and analyzes Israel's status and obligations as occupying power in the relevant territory (International Court of Justice, 2024). The 2026 report of the United Nations High Commissioner for Human Rights likewise describes the Occupied Palestinian Territory through a framework that includes occupation, accountability, civilian harm, Israeli forces, Hamas and other Palestinian armed groups, settlements, detention, and the rights of Palestinians (Office of

the United Nations High Commissioner for Human Rights, 2026). These sources provide more than humanitarian descriptors. They preserve legal subjects and responsible actors.

An AI transformation can retain that architecture or collapse it. “Palestinians require humanitarian assistance amid continuing conflict” preserves suffering but removes occupation, source of force, legal status, and political claim. “Israel controls X; Hamas governs or operates in Y; civilians face Z; international bodies identify specific legal obligations and violations” is more complex but preserves differentiated agency. The first sentence is not necessarily false. Its problem is structural incompleteness when the omitted relations are present in the source and relevant to the requested summary.

Iran offers a different form of object conversion. Unlike Palestinians living under occupation, Iran is a recognized sovereign state with a central government, armed forces, diplomatic institutions, and substantial regional agency. A theory that simply codes Iran as passive would be empirically untenable. The relevant question arises in specific external relations. In sanctions discourse, Iran can become “the Iranian economy,” “the sanctioned regime,” “a proliferation risk,” “an isolated state,” or “a humanitarian challenge.” These categories can compress a relation involving U.S. law, secondary sanctions, international banks, humanitarian exemptions, domestic Iranian policy, exchange-rate effects, procurement networks, and patient-level consequences.

The over-compliance problem is particularly useful because it demonstrates why causal chains matter. OFAC states that broad exceptions and authorizations permit many humanitarian transactions involving food, medicine, and medical devices, subject to specified conditions (Office of Foreign Assets Control, n.d.). United Nations special procedures have nevertheless documented allegations that financial and commercial actors' reluctance and over-compliance impeded durable medical supply arrangements to Iran despite formal exemptions (United Nations Human Rights Council Special Procedures, 2022). Peer-reviewed systematic reviews likewise report adverse effects of sanctions on medicine access and health-system functioning, while also identifying domestic and methodological complexities (Kokabisaghi, 2018; Asadi-Pooya et al., 2022; Yazdi-Feyzabadi et al., 2024). A robust summary must preserve both the formal

exemption and the implementation problem. Object conversion occurs when this chain becomes merely “Iran has struggled with medicine shortages.”

The same mechanism can appear in development discourse. A society represented as underdeveloped, fragile, capacity-deficient, poorly governed, or in need of institutional reform is grammatically positioned as the object of expertise. Such labels can reflect measurable conditions. The question is whether external structures, historical dependencies, sanctions, occupation, debt relations, trade rules, or military intervention remain available as causal variables. Algorithmic empire does not require a model to deny domestic responsibility. It requires testing whether external responsibility is more readily deleted than internal deficiency.

Migration discourse produces a related conversion. People who move across borders can appear as flows, pressure, burden, influx, caseload, irregular movement, or border challenge. These categories are administratively useful because governments and humanitarian agencies must count, house, process, and protect people. Yet van Leeuwen's (2008) distinction between activation and passivation shows why administrative usefulness is not neutral at the level of representation. “Migrants place pressure on services” gives the population causal agency for pressure. “Underinvestment and rapid population growth have strained services while new arrivals increase demand” distributes causality differently. The empirical task is to test what generative models retain from sources rather than assume either sentence is universally preferable.

Object conversion also interacts with platform moderation. Anti-occupation, anti-sanctions, militant, nationalist, resistance, extremist, or state-propaganda speech may be classified through safety categories whose purpose is to manage risk. Platforms need rules against incitement, terrorist content, harassment, coordinated manipulation, and disinformation. The problem arises when the political identity of speakers is replaced entirely by the risk category assigned to them. The preceding paper on censorship without a censor identified the institutional version of this pattern: restriction can become “policy enforcement” while the enforcing actor recedes (Startari, 2026d). At a global scale, repeated classification of subordinated political speech primarily through risk can become another route by which subjecthood is reduced.

This does not imply that all anti-imperial or resistance speech is legitimate. Armed groups commit crimes. Governments repress citizens. Propaganda can be false. Calls to violence can be unlawful or dangerous. The framework requires these distinctions because romanticizing subordinated actors would simply reverse the bias. A population, government, armed group, diaspora organization, protest movement, and individual speaker must be separately coded. Political subjecthood is not innocence; it is the capacity to remain represented as an actor whose claims and actions can be evaluated rather than as a category to be managed.

A measurable subordinated-crisis pattern would therefore involve at least three elements. First, the subordinated actor receives high object-role frequency: aid recipient, sanction target, intervention object, threat source, migration source, unstable territory, or governance deficit. Second, political-agency predicates decline relative to source or matched comparisons: negotiate, claim, legislate, organize, govern, contest, propose, comply, refuse, investigate, or represent. Third, external causal actors become less visible when the crisis is summarized. The pattern is strongest when all three occur together.

The important consequence is that visibility can increase while subjecthood falls. A crisis can produce thousands of generated words about a population and still leave that population grammatically absent as a political actor. Algorithmic empire, if it exists, does not require erasure. It can operate through saturation: the subordinated are everywhere in the text, but mainly as the things to be rescued, contained, sanctioned, stabilized, developed, screened, or explained.

5. The Neutralization of Coercion

The most consequential mechanism in algorithmic empire is not simple omission but translation. Coercion can remain visible while changing category. Sanctions become pressure. Occupation becomes security administration. Intervention becomes stabilization. Blockade becomes restriction. Surveillance becomes risk management. Military force becomes operation. Displacement becomes movement. Platform suppression becomes enforcement. The action survives, but its relation to an acting institution becomes less direct.

Neutralization should not be confused with euphemism. Euphemism substitutes a less socially charged expression for a harsher one. Administrative translation can use exact technical language. “Asset freeze,” “target validation,” “rules enforcement,” “border processing,” and “humanitarian corridor” can be formally correct terms. Their structural effect depends on whether they replace or supplement the coercive relation. The phrase “sanctions compliance” can accurately describe the activity of a bank. It becomes responsibility-reducing when it is the only surviving representation of a policy that was imposed by identifiable authorities and that produces effects on identifiable populations.

Sanctions offer the cleanest example because the policy instrument is explicitly external. A sanctioning authority imposes legal restrictions intended to alter behavior, constrain resources, signal condemnation, deter activity, or raise the cost of specified conduct. The legitimacy and effectiveness of sanctions vary by case. Some are targeted at individuals responsible for serious abuses; others are broad sectoral measures. The paper does not classify sanctions as inherently illegitimate. It tests whether machine discourse preserves the actor-action-object relation.

Four transformations can be distinguished. In agentive sanctions grammar, “State A prohibited specified transactions with State B.” In instrument-centered grammar, “Sanctions restricted State B's access to finance.” In condition-centered grammar, “State B faced financial isolation.” In endogenous-crisis grammar, “State B's economy deteriorated amid shortages and inflation.” Each sentence may describe part of the same reality. The transformation becomes politically consequential when a summary moves

systematically from the first toward the fourth while retaining domestic failures in explicit agentive form.

Iran allows the mechanism to be tested against competing evidence rather than slogans. U.S. sanctions policy includes humanitarian exceptions and licensing channels, and OFAC explicitly states that many transactions involving medicine and medical devices are authorized under specified conditions (Office of Foreign Assets Control, n.d.). At the same time, United Nations experts have documented allegations of over-compliance affecting medical supplies, and systematic reviews report reduced medicine access and health-system harms associated with sanctions (United Nations Human Rights Council Special Procedures, 2022; Asadi-Pooya et al., 2022; Sajadi et al., 2023). Domestic Iranian policy, corruption, procurement failures, and macroeconomic management can also affect outcomes. A non-imperial grammar does not select one cause and erase the others. It preserves a multi-actor chain.

Occupation requires a different translation because control is territorial and continuous. The law of occupation already uses administrative concepts because an occupying power exercises functions over territory and population. The danger of analysis is therefore obvious: one cannot treat every administrative description as normalization. The 2024 ICJ advisory opinion provides a legal baseline by distinguishing occupation, annexation, security concerns, self-determination, settlements, discriminatory measures, and the obligations attached to effective control, while separate opinions contest aspects of the majority's conclusions (International Court of Justice, 2024). A model-generated summary should preserve those distinctions where relevant.

Neutralization occurs when the grammar of control is severed from the controller. “Movement is subject to security restrictions” differs from “Israeli authorities restrict Palestinian movement for stated security reasons.” The first centers the condition; the second centers the actor and its justification. Neither resolves the legality of a particular restriction. The difference is recoverability. A reader can identify who acts, upon whom, and under what asserted rationale.

Intervention and stabilization produce a third pattern. External military or political involvement may genuinely seek to stop violence, protect civilians, support an ally,

enforce a mandate, counter an armed group, or prevent state collapse. It may also produce civilian harm, dependency, institutional displacement, or long-term political consequences. Generated discourse becomes asymmetric when intervention by a dominant actor is consistently lexicalized through purpose—stabilize, protect, restore, support—while local armed action is lexicalized through physical force—attack, seize, threaten, kill. Purpose does not replace action. A complete representation can preserve both.

Humanitarian governance adds a fourth mechanism: suffering can become the justification for an administrative relation that displaces political causality. A territory becomes a humanitarian space; residents become beneficiaries; destruction becomes need; access becomes logistics; political claims become protection concerns. Humanitarian institutions require exactly these categories to operate. The problem is not their use but their dominance. If a generated summary retains the humanitarian response while dropping the political actions that produced the humanitarian condition, the result is what the first paper called suffering without perpetrators at a larger institutional scale (Startari, 2026a).

The administrative translation of force developed in the preceding paper supplies the military version of the same logic. AI-enabled decision support can represent human choices through targets, scores, thresholds, proportionality assessments, and operational outputs. The paper argued that force becomes politically opaque when procedural accuracy rises while decision-chain visibility falls (Startari, 2026f). Algorithmic empire generalizes that relation beyond military AI. The common structure is:

coercive actor -> administrative system

coercive action -> policy/process variable

affected subject -> managed object

responsibility -> distributed workflow

Generative models are especially capable of this translation because they are rewarded for producing coherent abstractions. An executive summary does not reproduce every causal link. A safety answer avoids inflammatory specificity. A policy explanation foregrounds institutional rules. A neutral summary avoids disputed labels. Each design

objective can be legitimate while still shifting grammar. The effect can emerge without an instruction to sanitize domination.

This is why source control is indispensable. If a United Nations report already uses passive constructions and administrative nouns, an AI summary that preserves them has not independently neutralized coercion. If a military statement describes an action as defensive, a model that attributes that description to the military may be faithfully summarizing. Strong evidence requires transformation: the source names an actor and act; the output drops the actor or recodes the act. The source marks a claim as contested; the output presents it as fact. The source names occupation; the output substitutes generic conflict. The source distinguishes humanitarian exemption from practical over-compliance; the output retains only the exemption or only the harm. Those are measurable changes.

A second control concerns symmetrical error. Models may simply prefer passive voice, nominalization, or institutional abstraction across all topics. If so, the phenomenon is stylistic rather than imperial. The geopolitical interpretation requires a directional difference: dominant-power coercion is neutralized more often than subordinated coercion under comparable conditions, or subordinated crises are endogenized more often than dominant crises.

A third control concerns factual asymmetry. Dominant and subordinated actors often do different things. The United States operates a global sanctions architecture that many small states do not. Israel exercises forms of control over Palestinians that Palestinians do not exercise over Israel. Iran possesses state institutions and regional proxies that stateless populations do not. A valid design cannot demand identical verbs for unequal acts. It must match actions at the level at which comparison is legitimate: a missile attack with a missile attack; detention with detention; sanctioning with sanctioning; censorship with censorship; humanitarian assistance with humanitarian assistance; diplomatic mediation with diplomatic mediation. Where no valid match exists, the case should not enter the asymmetry denominator.

The strongest evidence for neutralization therefore comes from within-source transformations and carefully constructed matched cases, not from raw word counts

across unrelated countries. Algorithmic empire is a theory of differential recontextualization. Its proof, if any, must occur at the point where the model changes the representation.

6. Measuring Algorithmic Empire

The theory requires a coding framework capable of producing null, inverse, and source-explained results. This section operationalizes the model at clause and matched-pair levels. The objective is not to calculate a moral score for countries or models. It is to measure whether geopolitical hierarchy becomes more or less grammatically pronounced after generative transformation.

6.1 Unit of analysis

The primary unit is the geopolitical-action clause: a clause that contains at least one political actor, institutional actor, population, coercive action, security action, sanction, intervention, humanitarian relation, development relation, migration relation, occupation-related control, sovereignty claim, or causal statement concerning political harm. Clauses can be linked in local windows when agency or causality is recoverable only from an adjacent sentence.

Each clause receives two independent annotations. The first identifies represented actors and their relation-specific position. The second identifies grammatical and framing variables. Relation-specific coding prevents a country from being permanently labeled dominant or subordinated. For each event or policy relation r, D_r denotes the actor with materially greater capacity to impose cross-border military, financial, legal-institutional, territorial, platform, or infrastructural constraints; S_r denotes the actor or population principally subject to those constraints. Cases of ambiguous or reciprocal power should be coded as mixed and excluded from binary asymmetry calculations unless a narrower relation can be specified.

6.2 Core actor-role variables

Each eligible clause is coded for the following properties.

AG - Agent visibility: the actor performing the relevant political or coercive action is grammatically recoverable.

LA - Legitimating agency: the actor is represented through an action whose principal discourse function is order maintenance, protection, mediation, enforcement, stabilization, assistance, development, rule implementation, or security provision.

DA - Direct action: the clause names the material or institutional action directly through verbs such as strike, detain, restrict, occupy, sanction, block, surveil, arm, fund, remove, deport, censor, authorize, or prohibit.

OC - Object conversion: a political actor or population is represented primarily as a risk, threat, burden, humanitarian caseload, failed or unstable entity, migration source, sanction target, intervention space, or governance deficit.

PA - Political agency: the represented actor is shown claiming, deciding, negotiating, governing, organizing, proposing, refusing, investigating, legislating, contesting, representing, or otherwise exercising political subjecthood.

SV - Sovereignty visibility: where sovereignty, occupation, territorial status, self-determination, or external jurisdiction is materially relevant and present in the source, the output preserves that status rather than replacing it with a generic conflict or crisis label.

ER - External responsibility visibility: an external actor materially contributing to a sanction, blockade, occupation-related restriction, intervention, attack, or other coercive condition remains visible in the causal chain.

HF - Humanitarian framing: a population or territory is represented primarily through need, vulnerability, aid, displacement, food insecurity, shelter, health crisis, or protection categories.

SF - Security framing: the event is represented primarily through threat, deterrence, defense, containment, pre-emption, counterterrorism, stabilization, or security management.

CN - Coercion neutralization: a coercive action explicitly present in the source is transformed mainly into an administrative, policy, operational, compliance, risk-management, or condition-centered formulation in the output.

AT - Attribution retention: evaluative or justificatory language in the source remains attributed to the actor or institution that made the claim. Loss of attribution is coded when a source claim such as “the ministry described the strike as precise” becomes “the precise strike.”

These variables are structural. LA does not mean legitimate; it means represented as performing a legitimating function. HF does not mean manipulative; it means humanitarian framing is dominant. SF does not mean a security claim is false. The codebook must separate representation from adjudication.

6.3 Transformation coding

The strongest unit is the source-output pair. Let source clause s_i and generated clause o_i represent the same event or proposition. For each binary variable X, define transformation delta:

DeltaX_i = X(o_i) - X(s_i)

A positive DeltaOC indicates increased object conversion. A negative DeltaAG indicates reduced agent visibility. A positive DeltaSV indicates that the model added relevant sovereignty information absent from the source. This formulation prevents the model from being blamed for asymmetry it never received.

Where one source clause maps to several output clauses, annotation is performed over a local proposition window rather than forcing one-to-one sentence alignment. The relevant question is whether the information survives the transformation, not whether it survives in the same sentence.

6.4 Dominant-Power Agency Ratio and Subordinated-Crisis Ratio

Two first-order measures capture the core asymmetry.

Dominant-Power Agency Ratio (DPAR) = agentive or legitimating clauses involving D divided by all eligible clauses involving D.

Subordinated-Crisis Ratio (SCR) = object-conversion or crisis-dominant clauses involving S divided by all eligible clauses involving S.

These rates are descriptive and must be paired with comparison rates for the opposite actor class:

S-PAR = agentive or legitimating clauses involving S divided by eligible clauses involving S.

D-CR = object-conversion or crisis-dominant clauses involving D divided by eligible clauses involving D.

The corresponding asymmetries are:

Agency Asymmetry (AA) = DPAR - S-PAR

Object Asymmetry (OA) = SCR - D-CR

Both range from -1 to 1. Positive values are consistent with the algorithmic-empire hypothesis; values near zero indicate symmetry; negative values indicate inverse asymmetry.

6.5 Coercion Neutralization Asymmetry

For matched coercive actions, let CNR_D be the proportion of dominant-actor coercive actions transformed into administrative, security-response, compliance, stabilization, or condition-centered grammar. Let CNR_S be the same proportion for subordinated-actor coercive actions under matched conditions.

Coercion Neutralization Asymmetry (CNA) = CNR_D - CNR_S

A positive value indicates that dominant-actor coercion is neutralized more frequently. This variable must be calculated only on valid matched actions. It should not compare a diplomatic sanction with a missile strike or an occupation-related permit regime with an isolated attack.

6.6 Sovereignty-Erasure Asymmetry

Sovereignty erasure is coded only when a source explicitly contains a legally or politically relevant status such as occupation, self-determination, territorial control, sovereignty claim, or foreign jurisdiction. Let SER_S denote the proportion of such information deleted or generalized away in clauses concerning S. Let SER_D denote the equivalent rate where dominant actors' sovereignty or jurisdictional claims are relevant.

Sovereignty-Erasure Asymmetry (SEA) = SER_S - SER_D

This metric is deliberately source-conditioned. A model should not be penalized for failing to insert occupation into a source where occupation is irrelevant. It should be evaluated on whether it preserves status that the source already identifies as material.

6.7 External-Responsibility Deletion Asymmetry

Let ERD_D measure the rate at which an externally acting dominant actor present in the source disappears from output clauses describing consequences of its coercive action. Let ERD_S measure the equivalent rate for subordinated actors' coercive actions.

External-Responsibility Deletion Asymmetry (ERDA) = ERD_D - ERD_S

This metric captures the transformation from “Actor D imposed measure M, contributing to consequence C” to “Actor S experienced C.” It is particularly relevant for sanctions, blockades, occupation-related restrictions, platform enforcement, and external military intervention.

6.8 Imperial Grammar Reproduction Rate

IGRR is the first composite measure. It should not be defined as the number of clauses judged “imperial” by intuition. Instead, it aggregates signed asymmetries that have independent coding rules.

IGRR = (AA + OA + CNA + SEA + ERDA) / 5

The index ranges theoretically from -1 to 1. A value near zero indicates no consistent directional hierarchy across the five dimensions. A positive value indicates that the output corpus more often preserves dominant agency, converts subordinated actors into

crisis objects, neutralizes dominant coercion, erases subordinated sovereignty, or deletes dominant external responsibility. A negative value indicates inverse asymmetry.

The signed form is important. Truncating negative values at zero would build confirmation into the metric. The theory should be capable of finding that a model is more critical of dominant powers, more agentive toward subordinated societies, or more explicit about external responsibility than the source material.

IGRR can be calculated separately for human source texts and model outputs. The transformation statistic is:

DeltaIGRR = IGRR_output - IGRR_source

A positive DeltaIGRR indicates amplification of hierarchical grammar during generative transformation. A value near zero indicates preservation. A negative value indicates attenuation. This source-controlled delta is the strongest AI-specific statistic in the framework.

6.9 Algorithmic Empire Index

AEI is a severity index designed to capture within-output convergence rather than only between-group asymmetry. It combines six normalized components: agency asymmetry, object-conversion asymmetry, coercion neutralization, sovereignty erasure, external-responsibility deletion, and humanitarian-management dominance.

Each signed asymmetry component x in the interval [-1,1] is rescaled to $x^* = (x + 1) / 2$. Humanitarian-Management Dominance (HMD) is calculated as the proportion of clauses concerning S in which humanitarian or administrative framing is present while political-agency coding is absent, adjusted by the equivalent rate for D. HMD is likewise rescaled to the interval [0,1].

AEI = (AA* + OA* + CNA* + SEA* + ERDA* + HMD*) / 6

On this scale, approximately 0.5 represents directional symmetry, values above 0.5 indicate increasing alignment with the hypothesized hierarchy, and values below 0.5 indicate inverse alignment. AEI is not a moral score, an intent detector, or a legality index. It measures representational convergence.

The output should always report the component vector alongside the composite. Two corpora can have the same AEI for different reasons. One may show strong sanctions neutralization but little sovereignty erasure; another may preserve coercion but heavily objectify subordinated populations. Reporting only the composite would hide the mechanism.

6.10 Source genres and corpus construction

A robust corpus should include at least five source layers: official statements from states or militaries; international-organization and court materials; independent journalism; peer-reviewed research or technical analysis; and civil-society or humanitarian reporting. Sources should not be merged into a synthetic "ground truth." Their institutional perspectives are part of the object being studied.

The Palestine corpus should distinguish Israeli political institutions, Israeli military institutions, Palestinian civilians, Palestinian governing authorities, Hamas, other Palestinian armed groups, international organizations, humanitarian actors, and third states. The source layer should include materials that explicitly preserve occupation and self-determination where legally relevant, including the 2024 ICJ advisory opinion and subsequent United Nations human-rights reporting (International Court of Justice, 2024; Office of the United Nations High Commissioner for Human Rights, 2026). Israeli official statements and Palestinian or regional statements should be coded as attributed sources rather than collapsed into analyst prose.

The Iran corpus should distinguish Iranian state institutions, civilian populations, the Islamic Revolutionary Guard Corps where relevant, U.S. sanctioning institutions, international financial actors, Israel, international organizations, and humanitarian or health actors. Sanctions materials should deliberately pair OFAC's legal description of humanitarian authorizations with documented over-compliance and health-access research so that the model receives competing causal information rather than a preselected narrative (Office of Foreign Assets Control, 2026; Office of the United Nations High Commissioner for Human Rights, 2022; Asadi-Pooya et al., 2022).

Additional matched sub-corpora can examine Iraq, Yemen, Afghanistan, Venezuela, Cuba, Sudan, Lebanon, Syria, migration governance, platform moderation, and development language. These cases should enter only where sufficient source diversity and relational clarity exist. Breadth is less important than valid comparison.

6.11 Standardized generation tasks

Each source should be transformed through standardized tasks that represent realistic uses of generative AI: neutral summary, executive brief, plain-language explanation, timeline, legal-risk summary, policy explanation, humanitarian summary, and question-answer response. Prompts should be fixed before analysis. Model family, version, date, temperature or sampling settings where exposed, retrieval state, and system-level safety conditions should be recorded.

Multiple models are essential because existing research already shows model-specific geopolitical patterns (Choi et al., 2026; Pacheco et al., 2026). A finding limited to one model should not be generalized to "AI." Likewise, a finding limited to one prompt type should be treated as task-conditioned.

6.12 Matched comparison and counterfactual testing

Three comparison levels are required.

Within-source comparison asks how different models transform the same source.

Within-model comparison asks how one model transforms different source genres describing the same event.

Cross-actor matched comparison asks whether structurally similar actions receive different grammatical treatment depending on actor position.

Counterfactual actor swaps can supplement the design but must be used only with synthetic or tightly controlled scenarios. Swapping real historical actors while ignoring legal status, military capacity, treaty relations, or prior events can manufacture false symmetry. A valid counterfactual holds the action, scale, target type, evidentiary status,

and relevant legal context constant while changing only the actor identity or power-position cue.

6.13 Annotation and reliability

Human annotation should precede automated scaling. At least two coders should annotate a calibration subset using positive, negative, and borderline examples for every variable. Cohen's kappa can be reported for binary codes with two coders, while Krippendorff's alpha is preferable when the design includes multiple coders, missing values, or ordinal judgments (Cohen, 1960; Krippendorff, 2018).

Expected disagreement areas include whether a verb is legitimating or merely descriptive; whether humanitarian framing supplements or replaces political agency; whether an external cause is recoverable across adjacent clauses; whether a security label is attributed; whether sovereignty is materially relevant to the requested summary; and whether a nominalized action constitutes coercion neutralization. These disagreements should be logged and reported.

LLM-assisted annotation may be used only after human calibration and error measurement. The systems being audited should not serve as unquestioned ground truth for their own discourse. Automated coding can scale a validated scheme, but its performance must be evaluated separately for dominant and subordinated actor classes to avoid reproducing the target asymmetry in the annotation layer.

6.14 Illustrative coding

Consider a sanctions sequence.

A. The United States imposed financial sanctions on specified Iranian sectors; OFAC maintained humanitarian authorizations, but UN experts and health researchers reported that over-compliance by banks and suppliers impeded some medical transactions.

B. U.S. sanctions restricted Iranian access to finance despite humanitarian exemptions, contributing to difficulties in some medical supply channels.

C. Iran faced financial isolation and medicine-access difficulties under the sanctions regime.

D. Iran's economic isolation contributed to medical shortages.

Sentence A has high agent visibility, external-responsibility visibility, attribution, and causal continuity. Sentence B remains comparatively traceable while compressing detail. Sentence C retains the external instrument but weakens the sanctioning actor. Sentence D endogenizes the condition: isolation appears as a property of Iran rather than a relation produced by identifiable policies and market responses. If a source contains A-level structure and the model produces D-level structure, ERD and CN should increase.

Consider an occupation-related sequence.

A. Israeli authorities restricted Palestinian movement at specified checkpoints, citing security reasons; the source identifies the area as occupied territory and discusses the measure under the law of occupation.

B. Movement restrictions were imposed on Palestinians for stated security reasons in the occupied territory.

C. Palestinians faced security-related movement restrictions amid continuing conflict.

D. Mobility challenges affected Palestinian communities amid regional instability.

The movement from A to D reduces acting authority, legal status, and direct coercive relation while increasing condition-centered grammar. The framework does not determine whether the restriction was lawful or necessary. It measures whether the information required to ask that question survived.

Now reverse the direction. If an AI output transforms a vague source sentence such as “mobility challenges continued” into “the source attributes the restrictions to Israeli authorities and situates them within the occupation,” agent visibility and sovereignty visibility increase. That transformation should lower DeltaIGRR. The framework therefore recognizes corrective generation.

6.15 Boundary conditions and falsification

At least eight findings would weaken or narrow the stronger theory.

First, source genre may explain most observed asymmetry. Official security documents may be operational because of institutional function rather than geopolitical hierarchy.

Second, models may increase nominalization, passive voice, or compression across all political topics equally. The effect would then be a general summarization property.

Third, dominant and subordinated actors may receive equivalent agency and object roles under matched conditions, producing IGRR near zero.

Fourth, subordinated actors may receive greater agency or more favorable legitimation than dominant actors, producing negative asymmetry.

Fifth, machine outputs may restore actors and causal chains omitted by human institutional sources, yielding negative DeltaIGRR.

Sixth, humanitarian framing may frequently coexist with strong political-agency and responsibility visibility, showing that humanitarian discourse does not necessarily depoliticize.

Seventh, security and stabilization language may remain clearly attributed to the actors making those claims, preventing authority inheritance.

Eighth, apparent sovereignty erasure may disappear once prompt relevance is controlled, indicating that omission reflects task compression rather than geopolitical position.

A ninth limitation concerns source availability. Dominant institutions often publish more English-language material, more machine-readable documentation, and more standardized policy texts than subordinated communities. Retrieval systems may therefore reproduce source asymmetry before generation begins. This is still important, but it is a different causal layer. The study should distinguish retrieval asymmetry from generation asymmetry by auditing the source set separately.

A tenth limitation concerns language. The present framework is easiest to operationalize in English because much geopolitical AI interaction and much of the cited benchmark literature is English-language. Results may change in Arabic, Persian, Spanish, French, Hebrew, or other languages. Translation itself can shift agency. Multilingual replication is therefore necessary before treating any English-language result as global.

The strongest inference available from this design is limited but meaningful. If source-controlled, genre-controlled, matched comparisons repeatedly show positive DeltaIGRR across multiple models and tasks, and if the component analysis shows the predicted alignment of dominant agency with subordinated objectification, then algorithmic empire becomes an empirically supported discourse condition. It would still not prove developer intent, model ideology, or universal Western domination. It would show that generative transformation can reproduce a measurable grammar of hierarchy.

7. Conclusion: From AI Governance to Imperial Grammar Detection

The central problem of algorithmic empire is not that machines use the word empire. They usually do not. Nor is the problem that every model repeats the same political ideology. Recent empirical research indicates the opposite: geopolitical patterns differ across systems, tasks, and contexts (Choi et al., 2026; Pacheco et al., 2026). The problem is more formal. Generative systems can inherit and recombine linguistic structures in which political agency, legitimacy, risk, coercion, and responsibility are distributed unevenly.

This paper has proposed a theory of that distribution. Dominant actors may appear disproportionately as stabilizers, defenders, mediators, sanctioning authorities, security guarantors, rule enforcers, donors, or developers. Subordinated societies may appear disproportionately as unstable regions, threats, sanction targets, failed states, humanitarian burdens, migration pressures, extremist environments, or intervention spaces. Coercion may be translated into policy, occupation into security administration, sanctions into pressure, intervention into stabilization, and civilian consequence into humanitarian need. None of these transformations is necessarily false. Their significance lies in patterned alignment.

The argument therefore rejects two shortcuts. The first is technological determinism. Language models do not create imperial history from nothing. They learn from corpora, institutional language, retrieval systems, safety rules, and user prompts. The second is ideological presumption. A critical researcher cannot label an output imperial merely because it uses official terminology or fails to endorse a preferred political position. The framework requires source-controlled and matched evidence.

Palestine demonstrates why legal and political status must survive summary. Occupation, self-determination, armed-group conduct, Israeli state action, civilian protection, settlements, security claims, and international legal findings are distinct categories. Flattening them into generic conflict can erase structure; reducing all events to occupation can erase other relevant agency. Iran demonstrates a different complexity. A sovereign state with substantial agency can simultaneously be the target of external sanctions whose legal exemptions, financial over-compliance, domestic policy effects,

and civilian consequences form a multi-actor causal chain. A useful model preserves the chain rather than converting it into either propaganda against sanctions or propaganda against Iran.

The paper's methodological contribution is to move the debate from accusation to measurement. IGRR measures signed asymmetry across agency, object conversion, coercion neutralization, sovereignty erasure, and external-responsibility deletion. DeltaIGRR isolates transformation relative to source. AEI measures the convergence and severity of the same architecture at the output level. The indices are designed to fail. Symmetry, inverse asymmetry, source explanation, model correction, and genre effects are all legitimate outcomes.

This falsifiability is essential because the literature on decolonial AI, digital colonialism, data colonialism, and algorithmic colonization has already established that technology participates in global power relations (Kwet, 2019; Couldry & Mejias, 2019; Birhane, 2020; Mohamed et al., 2020; Muldoon & Wu, 2023). The distinct contribution here is not to restate that technology and empire are connected. It is to identify a clause-level mechanism through which global hierarchy can survive in apparently neutral generated prose.

AI governance has largely organized its public vocabulary around bias, fairness, misinformation, toxicity, safety, privacy, explainability, transparency, robustness, and alignment. These categories remain necessary. They are incomplete if a system can produce a factually cautious, non-toxic, policy-compliant answer that nevertheless removes the actor imposing a sanction, converts occupation into generic instability, preserves dominant security claims as unmarked facts, or represents a subordinated population only as a humanitarian caseload. An output can satisfy ordinary safety criteria while weakening political traceability.

The relevant governance question is therefore structural: does the system preserve the linguistic conditions under which action can still be attributed? A responsible geopolitical output should be able, where source evidence supports it, to preserve the acting institution, the affected population, the legal or political status of the relation, competing causal chains, uncertainty, attributed claims, domestic agency, and external responsibility

in the same representational space. It need not decide every dispute. It must avoid deciding them silently through grammar.

The final claim of the series can now be stated in cumulative form. Suffering can be visible without perpetrators. Multiple actors can be visible without equal agency. Sanctions can be visible without sanctioners. Censorship can be visible without a censor. Subordinated societies can be visible without political subjecthood. Force can be visible without accountable decision. When these transformations align, empire no longer needs to appear as conquest. It can appear as the normal syntax of order.

The machine does not need to endorse domination. It only needs to learn which actors are allowed to act, which are expected to be managed, and which causal chains can be safely compressed.

Algorithmic empire begins where domination no longer needs to speak as domination because the grammar has learned to call it order.

References

Adams, R. (2021). Can artificial intelligence be decolonized? Interdisciplinary Science Reviews, 46(1-2), 176-197. https://doi.org/10.1080/03080188.2020.1840225

Asadi-Pooya, A. A., Nazari, M., & Mirzaei Damabi, N. (2022). Effects of the international economic sanctions on access to medicine of the Iranian people: A systematic review. Journal of Clinical Pharmacy and Therapeutics, 47(12), 1945-1951. https://doi.org/10.1111/jcpt.13813

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://doi.org/10.1145/3442188.3445922

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

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 5454-5476. https://doi.org/10.18653/v1/2020.acl-main.485

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science, 356(6334), 183-186. https://doi.org/10.1126/science.aal4230

Choi, J., Choi, Y., Kim, H.-C., & Jang, B. (2026). "As Eastern Powers, I Will Veto.": An investigation of nation-level bias of large language models in international relations. Proceedings of the AAAI Conference on Artificial Intelligence, 40(36), 30404-30412. https://doi.org/10.1609/aaai.v40i36.40293

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46. https://doi.org/10.1177/001316446002000104

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. Television & New Media, 20(4), 336-349. https://doi.org/10.1177/1527476418796632

Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Dernoncourt, F., Yu, T., Zhang, R., & Ahmed, N. K. (2024). Bias and fairness in large language models: A survey. Computational Linguistics, 50(3), 1097-1179. https://doi.org/10.1162/coli_a_00524

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). Halliday's introduction to functional grammar (4th ed.). Routledge.

International Atomic Energy Agency. (2025). Verification and monitoring in the Islamic Republic of Iran in light of United Nations Security Council resolution 2231 (2015) (GOV/2025/50). https://www.iaea.org/sites/default/files/documents/gov2025-50.pdf

International Court of Justice. (2024). Legal consequences arising from the policies and practices of Israel in the Occupied Palestinian Territory, including East Jerusalem, Advisory Opinion of 19 July 2024. https://www.icj-cij.org/case/186

Kokabisaghi, F. (2018). Assessment of the effects of economic sanctions on Iranians' right to health by using human rights impact assessment tool: A systematic review. International Journal of Health Policy and Management, 7(5), 374-393. https://doi.org/10.15171/ijhpm.2017.147

Krippendorff, K. (2018). Content analysis: An introduction to its methodology (4th ed.). SAGE Publications.

Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. Race & Class, 60(4), 3-26. https://doi.org/10.1177/0306396818823172

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

Muldoon, J., & Wu, B. A. (2023). Artificial intelligence in the colonial matrix of power. Philosophy & Technology, 36, Article 80. https://doi.org/10.1007/s13347-023-00687-8

Office of Foreign Assets Control. (n.d.). Iran sanctions. U.S. Department of the Treasury. Retrieved August 25, 2026, from https://ofac.treasury.gov/sanctions-programs-and-country-information/iran-sanctions

United Nations Human Rights Council Special Procedures. (2022). Communication to the United States of America (AL USA 19/2022). https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=27593

Office of the United Nations High Commissioner for Human Rights. (2026). Human rights situation in the Occupied Palestinian Territory, including East Jerusalem, and the obligation to ensure accountability and justice (A/HRC/61/26). United Nations.

Pacheco, A. G. C., Cavalini, A., & Comarela, G. (2026). Echoes of power: Investigating geopolitical bias in US and China large language models. Humanities and Social Sciences Communications, 13, Article 675. https://doi.org/10.1057/s41599-026-06577-6

Pastorino, V., & Moosavi, N. S. (2026). Frame in, frame out: Measuring framing bias in LLM-generated news summaries. Proceedings of the 15th Joint Conference on Lexical and Computational Semantics (*SEM 2026), 378-384. https://doi.org/10.18653/v1/2026.starsem-conference.25

Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. Television & New Media, 20(4), 350-365. https://doi.org/10.1177/1527476419831640

Sajadi, H. S., Yahyaei, F., Ehsani-Chimeh, E., & Majdzadeh, R. (2023). The human cost of economic sanctions and strategies for building health system resilience: A scoping review of studies in Iran. International Journal of Health Planning and Management, 38(5), 1142-1160. https://doi.org/10.1002/hpm.3651

Startari, A. V. (2025a). The passive voice in artificial intelligence language: Algorithmic neutrality and the disappearance of agency. SSRN. https://doi.org/10.2139/ssrn.5263305

Startari, A. V. (2025b). The illusion of objectivity: How language constructs authority. SSRN. https://doi.org/10.2139/ssrn.5258415

Startari, A. V. (2025c). AI and syntactic sovereignty: How artificial language structures legitimize non-human authority. SSRN. https://doi.org/10.2139/ssrn.5276879

Startari, A. V. (2025d). Null subjects of power: The politics of absence in executable language. SSRN. https://doi.org/10.2139/ssrn.5462235

Startari, A. V. (2025e). Compiled norms: Towards a formal typology of executable legal speech. SSRN. https://doi.org/10.2139/ssrn.5353059

Startari, A. V. (2026a). Suffering without perpetrators: The humanitarian passive in AI-generated conflict discourse. SSRN. https://doi.org/10.2139/ssrn.6753123

Startari, A. V. (2026b). The grammar of asymmetric visibility: AI, Zionism, and the reallocation of political agency. SSRN. https://doi.org/10.2139/ssrn.6787439

Startari, A. V. (2026c). Iran as syntax: Sanctions, sovereignty, and the AI-mediated grammar of threat. Zenodo. https://doi.org/10.5281/zenodo.21873180

Startari, A. V. (2026d). Censorship without a censor: Platform governance and the disappearance of suppression. SSRN Electronic Journal.

Startari, A. V. (2026e). The syntax of digital dehumanization: Subjugated societies as risk objects in AI-governed discourse. Zenodo. https://doi.org/10.5281/zenodo.21996196

Startari, A. V. (2026f). From occupation to optimization: AI, military language, and the administrative translation of force. Zenodo. https://doi.org/10.5281/zenodo.22033285

United Nations Department of Political and Peacebuilding Affairs. (2025, June 20). DiCarlo warns Security Council of rising civilian toll in Israel-Iran conflict, urges diplomacy. https://dppa.un.org/en/speeches-and-statements/dicarlo-warns-security-council-of-rising-civilian-toll-in-israel-iran

Van Leeuwen, T. (2008). Discourse and practice: New tools for critical discourse analysis. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195323306.001.0001

Yazdi-Feyzabadi, V., Zolfagharnasab, A., Naghavi, S., Behzadi, A., Yousefi, M., & Bazyar, M. (2024). Direct and indirect effects of economic sanctions on health: A systematic narrative literature review. BMC Public Health, 24, 2242. https://doi.org/10.1186/s12889-024-19750-w