

## THESIS BRIEF — THEORY-FIRST RESEARCH

Edition: 2025-11-13 | Peer-review pending (Theory-First)

Smart Technology Investments

# Cognitive Wars: the AI Industrialization of Influence

Sources: 11 | Anchor Status: Anchor-Absent | Report Type: Theoretical Research | Horizon: Near-term | Confidence: 0.260 \*

SD	AC	MT	RR
0.06	0.00	0.45	0.65

Alignment: 6.0 Theory Depth: 6.0 Clarity: 7.0



*Image generated with OpenAI dall-e-3*

## Confidence Provenance

Source Diversity **0.06**

Anchor Coverage **0.00**

Method Transparency **0.45**

Replication Readiness **0.65**

Anchor coverage enforced at 0.00.

**Route:** THESIS — MarketScore 0.000 | Fresh 0/0, Unique domains 0, Anchors 0, Canonical 0

**Disclosure & Method Note:** This is a *theory-first* brief. Claims are mapped to evidence using a CEM grid; quantitative effects marked **Illustrative Target** will be validated via the evaluation plan. Where anchors are scarce, this brief is labeled **\*\*Anchor-Absent\*\*** and any analogical inferences are explicitly bounded.

## Abstract & Theory-First Framing.

<p><strong>Key Takeaways:</strong></p> <blockquote> <p><strong>Disclosure & Method Note.</strong> This is a <em>theory-first</em> brief. Claims are mapped to evidence using a CEM grid; quantitative effects marked <strong>Illustrative Target</strong> will be validated via the evaluation plan. <strong>Anchor Status:</strong> Anchor-Absent.</p> </blockquote> <ul> <li>This brief advances a theory-first argument that the industrialization of information — most recently catalyzed by AI and automation — has materially altered the cognitive foundations of warfare, producing a distinct set of dynamics I label "cognitive wars." Industrialization reshapes information production, distribution, and decision architectures, which in turn reconfigure collective beliefs, organizational heuristics, and escalation pathways</li> <li>The causal link advanced is: industrialization (automation, scale, algorithmic mediation) -> transformation of military and societal cognition (attention allocation, routinized inference, throughput heuristics) -> observable changes in the conduct and outcomes of wars (tempo, escalation dynamics, attribution ambiguity).</li> </ul>

### Outline

- [Abstract](#)
- [Introduction: cognitive wars and the puzzle](#)
- [Foundations](#)
- [Theoretical Grounding and Conceptual Framework](#)
- [Theory-first Framework: cognitive processes, wars, and industrialization influence](#)
- [Literature Review: cognitive, wars, and industrialization debates](#)
- [Hypotheses and Propositions](#)
- [Methodology: operationalizing cognitive change and industrialization](#)
- [Analytical Strategy and Evidence](#)
- [Mechanisms: how industrialized influence reshapes cognition](#)
- [Applications: parameterized vignettes and metrics](#)
- [Case Studies: industrialization-driven shifts in warfare cognition](#)
- [Discussion: theoretical implications and influence pathways](#)
- [Conclusion and Future Research](#)
- [Assumptions Ledger](#)
- [Notation](#)
- [Claim-Evidence-Method \(CEM\) Grid](#)
- [References \(selected cited sources\)](#)
- [Sources](#)

## **Abstract**

This brief advances a theory-first argument that the industrialization of information — most recently catalyzed by AI and automation — has materially altered the cognitive foundations of warfare, producing a distinct set of dynamics I label "cognitive wars." Industrialization reshapes information production, distribution, and decision architectures, which in turn reconfigure collective beliefs, organizational heuristics, and escalation pathways. The causal link advanced is: industrialization (automation, scale, algorithmic mediation) -> transformation of military and societal cognition (attention allocation, routinized inference, throughput heuristics) -> observable changes in the conduct and outcomes of wars (tempo, escalation dynamics, attribution ambiguity).

Predicted effects include increased primacy of efficiency/throughput heuristics in operational decision-making, organizational cognitive specialization (staffing and logistics heuristics), growth of mediated attention economies that amplify misinformation and synthetic media, and novel failure modes (systematic misattribution, adversarial exploitation of delegated judgment). Empirical strategy integrates process-tracing, textual coding of doctrinal materials, and quantitative proxies of industrial capacity and information amplification.

## **Introduction: cognitive wars and the puzzle**

Define "cognitive wars": the ensemble of collective beliefs, mental models, institutional heuristics, and decision-making shortcuts that shape behavior during conflict. Cognitive wars foreground how information architecture and organizational cognition co-produce wartime conduct beyond material inputs (troop numbers, factories, firepower).

Central puzzle: Traditional materialist accounts explain how industrialization increases production capacity and logistics — but they under-explain how industrialization reconfigures the cognitive substrate of warfare: why did states with similar material capacity adopt different doctrines, why did conflicts after industrialization exhibit new patterns of escalation, and how do modern AI-driven influence operations alter belligerents' perceptions and choices in ways that materialist metrics do not capture?

This brief addresses that puzzle by building a causal chain from industrial-scale information technologies to cognitive transformations in organizations and populations, and by operationalizing cognitive variables for empirical investigation.

# Foundations

## Why these anchors?

I anchor this analysis in a deliberately multi-source strategy. Anchor sources were chosen to be peer-reviewed or canonical scholarship when available (military history, political science, cognitive psychology). These anchors provide durable theoretical commitments (e.g., persuasion theory, organizational cognition, diffusion models) that are necessary for robust first-principles reasoning. When direct peer-reviewed literature on the newest AI-driven phenomena is sparse, I supplement with high-quality technical and empirical studies (including computational and system-focused preprints) to capture the latest mechanisms (e.g., social bots, synthetic media) and to link them back to canonical theory.

## Multi-layer selection strategy

- Direct Sources (Layer 1): Empirical and technical studies that speak directly to AI-driven influence operations, social bots, and synthetic media. These capture mechanisms of scale, automation, and adversarial exploitation in present-day information ecosystems [\[2\]](#)[\[5\]](#)[\[6\]](#)[\[7\]](#)[\[8\]](#)[\[9\]](#).
- Domain Sources (Layer 2): Literature in information operations, propaganda studies, and computational propaganda that situates influence within political and military practice; these works clarify institutional actors, tactics, and historical continuity with older PSYOP methods.
- Foundational Sources (Layers 3–5): Canonical papers and books on persuasion, social influence, information theory, game theory, and organizational behavior. These provide the theoretical backbone to map low-level mechanisms (e.g., algorithmic amplification) to high-level cognitive outcomes (e.g., altered priors, shifted equilibria).

How canonical papers from broader layers ground the argument

Canonical works on persuasion and contagion (e.g., persuasion theory, threshold models) give microfoundations for how information changes beliefs and behavior; organizational and game-theoretic treatments provide a basis for modeling strategic interaction under information constraints. When modern preprints document new tools (e.g., deepfakes, botnets), the canonical works allow us to reason from first principles about how these tools should change cognition and strategic outcomes.

Key empirical-tech anchors used in mechanism discussion include studies of social bots and botnets [\[2\]](#)[\[5\]](#), analyses of synthetic visual reenactment [\[6\]](#), and surveys of malicious uses of AI [\[7\]](#); these are complemented by technical literature on adversarial ML defenses [\[3\]](#) and attention-driven misinformation supply [\[4\]](#).

# Theoretical Grounding and Conceptual Framework

Abstraction layers and their central concepts

- Layer 1 (Specific): Cognitive warfare artifacts — automated disinformation, computational propaganda, deepfakes, algorithmic amplification — are the proximate mechanisms that alter what agents perceive and how they update beliefs.
- Layer 2 (Domain): Information operations and PSYOP provide institutional frames for how states and non-state actors deploy influence; social media manipulation and propaganda ecosystems describe the operational arena where cognition is shaped.
- Layer 3 (Applied Methods): NLG/LLMs, social bots, recommender systems, network analysis, and adversarial ML are the toolset enabling industrial-scale influence and the analytical methods for detecting it [1][2][5][10].
- Layer 4 (Abstract): Persuasion and social influence models explain individual and group belief updating; attention economy and cognitive biases explain selective exposure and susceptibility.
- Layer 5 (Foundational): Information theory, game theory, behavioral economics, and control theory supply principles for signaling, strategic manipulation, bounded rationality, and system-level stability.

Reasoning chain from foundational principles to the specific topic

1. Information-theoretic and game-theoretic foundations indicate that signals shape beliefs and that strategic actors will manipulate signals when incentives align (Layer 5).
2. Behavioral models show people are subject to bounded rationality and systematic biases; social influence models indicate cascades and threshold effects (Layer 4).
3. Recommender systems and attention economies concentrate exposure, changing the distribution of signals (Layer 3), while automation multiplies the speed and volume of signals (Layer 1).
4. Institutional and doctrinal practices mediate how organizations encode, transmit, and act on information (Layer 2). The confluence creates novel cognitive equilibria — "cognitive wars" — in which industrialized information production changes priors, heuristics, and coordination norms.

## How tangential canonical papers help

Papers not directly about AI-driven influence (e.g., threshold models, classic persuasion theory) still provide necessary microfoundations: they explain why an increase in cheaply produced persuasive signals will disproportionately affect collective outcomes via cascades and shifting thresholds. The conceptual map links these layers: foundational theory sets constraints and plausible dynamics; domain literature identifies institutional loci; applied-methods research demonstrates available leverage points and failure modes.

Reference: conceptual map "Cognitive Wars: The AI Industrialization of Influence" (see conceptual mapping above).

# Theory-first Framework: cognitive processes, wars, and industrialization influence

## Causal model (high level)

Industrialization of information (I) -> Mediators: information flows, bureaucratization and routinization, professionalization of information functions, compressed time horizons, and technological framing (M) -> Cognitive transformations in organizations and populations (C) -> Changes in warfare conduct and outcomes (O).

## Mechanisms (concise)

- Information Flows: Scale and velocity increase available signals, amplify fringe narratives via algorithmic recommendation, lower cost per manipulated impression [\[4\]](#)[\[10\]](#).
- Bureaucratization & Professionalization: Creation of staffed information operations, specialized analytic cadres, and doctrine-bound decision processes that favor throughput and procedural metrics over deliberative assessment.
- Time Horizons & Routinization: Faster operational tempos incentivize heuristics and templates (e.g., checklist-based targeting, automated attribution pipelines), increasing vulnerability to adversarial manipulation.
- Technological Framing: The presence of synthetic media (deepfakes) changes attribution uncertainty and erodes epistemic trust [\[6\]](#)[\[7\]](#).

## Mediation, not accompaniment

I argue these cognitive shifts mediate the effects of industrialization on military outcomes: rather than material capacity alone producing outcomes, industrialization reshapes the information environment and organizational cognition that determine how material assets are used.

# Literature Review: cognitive, wars, and industrialization debates

## Mapping debates

- Military history/materialist tradition: Emphasizes factories, logistics, and firepower; industrialization as input-output function of war-making capacity.
- Constructivist/ideational approaches: Focus on norms, doctrines, and beliefs but often under-specified about mechanisms linking technology to cognition.
- Cognitive sociology/organizational cognition: Provides microfoundations for institutional heuristics and decision processes but less often tied to macro-level industrial transformations.

## Gaps identified

While materialist accounts delineate how industrial capacity changes resources, they under-theorize how industrial-scale information production changes belief systems and heuristics. Conversely, constructivist accounts emphasize ideas but lack mechanistic linkages to technological affordances. This brief bridges that gap by situating cognitive mechanisms between industrial-scale information technologies and wartime behavior.

## Relation to computational propaganda and AI literature

Recent empirical and technical studies demonstrate how automation (bots, NLG, synthetic media) enable scale and adversarial manipulation [\[2\]](#)[\[5\]](#)[\[6\]](#)[\[7\]](#)[\[8\]](#)[\[9\]](#). However, many of these contributions stop at detection or cataloging; connecting them to institutional cognitive change remains an open task that this brief begins to map.

## Hypotheses and Propositions

Hypothesis 1: Industrialization increases the salience of throughput and efficiency heuristics in military decision-making. Organizations will weight metrics like "targets processed per hour" or "information triaged" more heavily than deliberative counterfactual reasoning.

Hypothesis 2: Greater industrial capacity leads to organizational cognitive specialization — staff planning, logistics primacy, and dedicated information operations — which concentrates decision authority and routinizes judgment.

Proposition: Cognitive shifts mediate the relationship between industrialization and operational tempo/escalation: higher industrialization -> cognitive routinization and compressed time horizons -> higher tempo and increased risk of miscalculated escalation, conditional on adversary signaling strategies.

### Testable sub-propositions

- P1a: Periods of rapid automation adoption will correlate with doctrinal language emphasizing tempo and efficiency in doctrinal texts.
- P1b: Organizations with higher reliance on automated sourcing (e.g., algorithmic attribution) will experience higher false-positive rates in target selection under adversarial influence.

## Methodology: operationalizing cognitive change and industrialization

### Operational indicators

- Cognitive variables (empirical proxies): frequency of heuristic language in doctrinal texts ("tempo," "throughput," "efficiency"), training curricula emphasis on automated tools, decision logs showing rule-based decisions, network metrics of information centralization (e.g., degree centrality of information nodes), linguistic markers of certainty/ambiguity.
- Industrialization measures: production capacity indices (armament output, factory throughput), mobilization rates, infrastructural integration (rail/telegraph historically; broadband/compute capacity contemporarily), automation adoption rates (numbers of deployed bots, LLM-based systems).

### Mixed-methods approach

- Comparative case selection: choose paired cases with similar material capacity but divergent information-industrial profiles (early vs late industrializers; high vs low algorithmic mediation).
- Process-tracing: link industrial milestones (build-out of communications infrastructure, adoption of automation) to observed changes in doctrine, decision procedures, and battlefield outcomes.
- Textual coding: supervised/unsupervised NLP to measure shifts in doctrinal language and heuristics over time.
- Quantitative correlates: regression analyses linking industrialization indices to coded cognitive variables, with temporal lags and robustness checks.

### Data sources and ethical considerations

- Open doctrinal manuals, archived decision logs (where declassified), press releases, digital media archives, compute/infrastructure statistics. Ensure human subjects protections where interviews or internal logs are used.

# Analytical Strategy and Evidence

## Process-tracing design

- Identify industrialization milestones (e.g., telegraph/railway adoption, mass broadcasting, internet penetration, LLM deployment).
- Trace contemporaneous changes in organizational structure, doctrine, and training.
- Establish temporal precedence and rule out alternative mechanisms (economic capacity shifts, leadership changes) through counterfactual reasoning.

## Coding scheme (high level)

- Lexical categories: efficiency/throughput, attribution/verification, automated decision language, escalation/avoidance.
- Decision-heuristic markers: template usage, checklist prevalence, delegations to automated systems.

## Robustness checks

- Alternative mechanisms: control for material capacity, leadership ideology, and alliance commitments.
- Temporal ordering: Granger-style lead-lag tests where high-frequency data are available (e.g., digital-era cases).
- Counterfactuals: process-trace plausible alternative histories where industrial information changes did not occur.

## Empirical anchors (examples)

- Studies of botnets and their growth quantify the scale at which automated influence is feasible [2][5].
- Work on attention and supply of misinformation links exposure dynamics to production incentives [4].
- Technical assessments of synthetic media demonstrate feasibility for plausible causation of attribution uncertainty [6][7].

# Mechanisms: how industrialized influence reshapes cognition

Four interlocking mechanisms translate industrial information capabilities into cognitive outcomes:

## 1. Signal Multiplication and Noise Inflation

Automation multiplies persuasive signals faster than human verification scales. The result is elevated noise-to-signal ratios that change priors: decision-makers, expecting higher noise, may assign lower credence to individual signals and rely more on heuristics and aggregated indices (e.g., trending counts) rather than source-level verification [4][10].

## 1. Algorithmic Selection and Attention Concentration

Recommender systems and platform incentives concentrate attention on a subset of content. This reweights exposure distributions, making fringe narratives achieve mainstream reach and altering perceived base rates — a cognitive shift from treating fringe signals as rare to treating them as common [1][10].

## 1. Delegation and Routinized Judgment

Organizations respond to information overload by delegating to automated pipelines (attribution algorithms, watchlist filters). Delegation reduces deliberative bandwidth and creates lock-in: when automated outputs become inputs to doctrine and targeting, adversarial actors can exploit predictable pipelines [3][7].

## 1. Attribution Uncertainty and Trust Erosion

High-fidelity synthetic media and coordinated bot amplification increase uncertainty about provenance. Heightened uncertainty reduces epistemic trust, incentivizes preemptive or precautionary actions, and complicates escalation control — a strategic lever for adversaries aiming to induce either paralysis or precipitous action [6][7][8].

These mechanisms combine to produce cognitive equilibria characterized by speed-optimized heuristics, specialized information staffs, routinized automated decision-making, and heightened attribution ambiguity.

# Applications: parameterized vignettes and metrics

*Image generated with OpenAI dall-e-3*

## Vignette A — Urban Counterinsurgency with Intermittent Communications

Scenario: A mid-sized state faces urban insurgency while its communications infrastructure is intermittently degraded (signal jamming, network outages). The state has adopted automated information triage systems to manage social media monitoring and force coordination.

### Parameters

- MTTA (Mean Time To Action) under normal comms: 15 minutes; under intermittent comms: variable, modeled as  $\text{MTTA}' = \text{MTTA} (1 + p_{\text{drop}} s_{\text{delay}})$ , where  $p_{\text{drop}}$  is probability of packet loss and  $s_{\text{delay}}$  is systemic delay factor.
- Failure probability (false-positive targeting due to automated attribution): baseline 0.02 per incident; increases to 0.12 under adversarially amplified misinformation.
- Human override latency: median 45 minutes when human-in-loop is activated.

### Failure modes

- Automated cascade: fast triage marks a viral synthetic video as an insurgent propaganda node; automated prioritization routes a kinetic response before human review (delegation policy violated), leading to collateral damage.
- Attention diversion: algorithmic amplification channels public attention to a fabricated atrocity, degrading local intelligence sources and spiking false-positive rates.

### Mitigations and metrics to monitor

- Trigger: sudden spike in cross-platform amplification of novel media or matched errors across attribution pipelines. Operational policy: escalate to human review if cross-platform amplification rate exceeds 5x baseline within 30 minutes.
- Delegation policy: require two independent algorithmic signals + one human confirmation before kinetic escalation in civilian-populated zones.

## Vignette B — Maritime Denial and Strategic Deception with AI-driven Influence

Scenario: Two states contest a maritime chokepoint. One state uses AI-generated disinformation campaigns to create ambiguous narratives about ship identities and intent, seeking to raise friction and cause preemptive closures.

### Parameters

- Time-to-escalation sensitivity: probability of interdiction given ambiguous signal =  $f(\text{uncertainty}, \text{readiness})$ , where uncertainty is modeled on a scale 0–1; a 0.2 increase in uncertainty raises interdiction probability by 15%.
- Failure probability (misattribution leading to international incident): baseline 0.01 annually; can jump to 0.08 under coordinated synthetic deception.
- MTTA for diplomatic clarification: baseline 72 hours; under attention-saturated conditions, diplomatic MTTA may extend to 120+ hours.

### Failure modes

- Strategic foreshortening: decision-makers, faced with high uncertainty and pressure to act, lean on throughput heuristics, leading to disproportionate interdiction and escalation.
- Attribution race: both sides accelerate evidentiary production (forensic timelines), but automated production of counterfeit evidence undermines verification, prolonging crises.

### Mitigations and metrics to monitor

- Trigger: emergence of high-fidelity synthetic media tied to critical nodes (e.g., naval command feeds). Policy: freeze-move rule — no kinetic escalation for 48 hours when provenance confidence < 0.6.
- Delegation policy: create an independent, multilateral forensic task force with rapid-exchange agreements to compress MTTA for verification to <24 hours when incidents involve third-party commercial AIS and satellite feeds.

Vignette C — Information-Denial on a Mobilized Home Front (short)

Scenario: Rapid mobilization requires public trust and compliance. Automated rumor networks (botnets) propagate panic about shortages.

## Parameters

- Evacuation compliance drop per rumor wave: 3–7% depending on trust baseline.
- Failure probability of mass noncompliance: increases with number of coordinated bot accounts active; modeled as  $p_{\text{noncompliance}} = \text{base} + k * \text{bot\_density}$ .

## Failure mode

- Panic cascade leading to logistical gridlock and reduced mobilization efficiency.

## Mitigation

- Early detection trigger: botnet growth exceeding 2000 coordinated accounts in 24h; immediate public transparency briefings and routing of official communications through multiple redundant channels to counteract algorithmic suppression.

(Collectively these vignettes illustrate how MTTA, failure probabilities, and delegation rules interact with cognitive and informational dynamics. They can be parameterized for simulation or wargame exercises.)

# Case Studies: industrialization-driven shifts in warfare cognition

## Suggested comparative cases (research agenda)

- Early industrializer conflict: World War I — emergence of staff systems, telegraph/rail logistics, doctrine emphasizing mass fire and tempo; examine doctrinal texts for throughput language.
- Late industrializer conflict: Russo-Japanese War / Russo-Ukrainian conflicts — compare adoption of communications tech and automated information tools; trace differences in organizational cognition.
- Digital-era cases: Information operations in electoral and hybrid conflicts; analyze botnet-driven influence and synthetic media incidents using computational detection studies [2][5][6][7].

## Illustrative empirical observations

- Doctrinal change: archival analysis often shows increased formalization of staff procedures and checklists coincident with communications/infrastructure investments.
- Organizational restructuring: states with large-scale information operations create specialized directorates and routinized targeting pipelines, consistent with the specialization hypothesis.
- Counter-examples: Cases where high industrial capacity did not produce routinized cognition often feature strong norms of deliberation and institutional checks, highlighting boundary conditions.

# **Discussion: theoretical implications and influence pathways**

## **Synthesis of evidence**

Industrialized information systems alter the cognitive substrate of warfare through measurable mechanisms: amplification, delegation, reframing of uncertainty, and institutional specialization. These changes shift strategic equilibria — e.g., they can shorten the credible signaling horizon, increase the prevalence of preemptive actions under uncertainty, and reconfigure civil-military relations as public information spaces become operational battlegrounds.

### Implications for theory

- War causation: Adds an information-cognition mediation channel to materialist accounts, suggesting that industrialization affects not only capacity but also the heuristics and priors that govern the use of that capacity.
- Military innovation: Highlights the role of non-kinetic technological adoption (automation, recommender systems) as drivers of doctrinal change.
- Civil-military relations: As publics become contested cognitive terrains, civilian oversight and public trust become operational variables in wartime decision-making.

## **Policy relevance**

- Anticipatory governance: Monitor indicators of cognitive shift (doctrinal language, automation adoption) to predict changes in conduct.
- Defensive measures: Invest in forensic attribution, cross-platform verification, and human-centered delegation policies to mitigate adversarial exploitation.

# **Conclusion and Future Research**

## **Contribution recap**

This brief advances a theory-first account that industrialization — now augmented by AI and automation — reshapes warfare by transforming the cognitive substrate of actors involved. The central claim is mediation: industrial information capabilities change cognition (beliefs, heuristics, institutional routines), which mediates the effect of industrialization on military conduct and outcomes.

## **Empirical priorities and methodological refinements**

- Systematic textual codings of doctrinal archives to test shifts in heuristic language.
- Comparative process-tracing across historical and contemporary cases to identify causal pathways and boundary conditions.
- Simulation and wargame experiments parameterized by the vignettes above to test delegation rules and MTTA trade-offs.

## **Extensions**

- Study of digital industrialization and LLM-driven influence as next-stage cognitive disruptors.
- Integration of adversarial ML defense research into operational doctrines to reduce systemic vulnerability [\[3\]](#).

Final note: understanding cognitive wars requires bridging disciplines — political science, military history, cognitive psychology, and AI systems research — and moving from cataloging new tools to modeling how they rewrite the heuristics and institutions that determine life-and-death policy choices.

## Assumptions Ledger

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
The industrialization of information via AI and automation materially alters the cognitive foundations of warfare, producing distinct dynamics characterized as "cognitive wars."	AI/automation increases signal volume, speed, and algorithmic mediation; these changes plausibly reshape attention, belief updating, and organizational heuristics in ways that affect decision-making independent of material inputs.	Shifts in doctrinal language and training (mentions of information operations, automation, or AI), growth of staffed information units, metrics emphasizing throughput/speed in after-action reports, case studies where similar material forces adopt different tactics correlated with information-environment differences.	Deployment or scale-up of AI-driven information tools, visible changes in organizational structure (creation of information/AI units), or sudden changes in operational tempo that force faster decision cycles.	Introduce explicit human-in-the-loop policies, restore deliberative decision nodes (e.g., mandatory human review for high-consequence actions), and delegate verification and strategic adjudication to cross-functional oversight teams (intelligence, legal, red teams).	Applies chiefly to actors operating in digitally mediated information ecosystems with access to AI/automation (modern militaries, states, major non-state actors); less applicable in low-tech, localized conflicts or where information channels remain primarily analog.
Algorithmic recommender systems and attention economies amplify fringe narratives, misinformation, and synthetic media disproportionately, increasing their downstream influence on populations and organizations.	Recommender algorithms optimize for engagement/throughput, which favors salient, emotional, or novel content; automation lowers marginal cost of producing persuasive signals, so scaled synthetic content and botnets can exploit these incentives.	Platform-level engagement spikes for fringe or synthetic items, rapid cascades originating from low-credibility accounts, high share-to-audience ratios, botnet activity signatures (coordinated timing/replication), and repeated resurfacing of the same synthetic artifacts across networks.	Unexplained virality of contested content, coordinated account behavior detected by platform analytics, electoral/mobilization periods, or incidents where rapid belief change is reported among target audiences.	Apply platform-level mitigations (throttle recommendations, label or demote synthetic content), route suspicious items to independent fact-checkers and platform trust-and-safety teams, and escalate to regulatory or interagency contacts when necessary.	Most relevant on commercial platforms with engagement-optimizing recommenders and broad public reach; impact varies by platform architecture, moderation regimes, and audience media literacy.
Under compressed time horizons and high-tempo operations, organizations will prefer	Organizational cognition literature predicts reliance on heuristics and checklists when time and attention are constrained; automation	Use of automated attribution pipelines, templated intelligence products, standard	Crisis situations demanding real-time decisions, alerts of rapid cascading information events, or discovery of	Enforce probabilistic confidence reporting, mandate second-opinion or red-team review for high-risk automated outputs,	Applies to organizations with formalized decision pipelines and

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
throughput/efficiency heuristics and routinized inference (templates, automated attributions), increasing susceptibility to adversarial manipulation of information.	and bureaucratization institutionalize these heuristics, producing predictable operational routines that adversaries can anticipate and exploit.	operating procedures that prioritize speed over uncertainty reporting, KPIs rewarding fast responses, and documented instances where adversarial signals triggered routine automated responses.	automated decision-support tools used for targeting/attribution.	throttle automated decision systems during anomalous input patterns, and delegate final-action authority to verified human operators.	access to automated analytics (armed forces, intelligence agencies, large NGOs); less applicable to small, decentralized groups or contexts where decisions are inherently deliberative and low-tempo.
Synthetic media and adversarial tactics will substantially increase attribution ambiguity and erode epistemic trust across publics and institutions.	Advances in generative models enable high-fidelity false visual/audio/textual artifacts; as such artifacts circulate, observers face higher Type I/II errors when validating signals, producing skepticism and contested truth claims documented in empirical studies.	Increased frequency of contested media claims, longer verification times for multimedia evidence, survey measures showing declining public trust in media or institutions after high-profile synthetic incidents, and legal/political disputes over provenance of evidence.	Emergence of high-impact contested media (e.g., purported battlefield videos, political deepfakes), intelligence reports flagging likely synthetic artifacts, or litigation/official disagreement over the authenticity of proof.	Invest in cryptographic provenance systems (signed content pipelines), establish rapid verification labs and accredited third-party validators, default to conservative policy in presence of ambiguous high-impact media, and delegate public communication to trained information-relations units.	Most acute where high-quality synthetic tooling is widely available and where publics rely on distributed digital media for situational awareness; less acute where authenticated analog verification (direct witnesses, physical inspection) remains dominant.
Cognitive transformations mediated by industrialized information (beliefs, heuristics, priors) can change the use and effectiveness of material military assets such that	Decision-making frameworks and doctrines determine force employment; if industrialized information reshapes those frameworks (e.g., through perception management, altered thresholds for	Cases where states with comparable materiel show different escalation behaviors or operational choices correlated with differences in information environments or	Comparative analyses showing outcome divergence among materially similar actors, or when influence operations precede atypical tactical/strategic choices by an opponent.	Revise doctrine/training to incorporate information-environment contingencies, implement war-gaming and red-team exercises that stress-test cognitive	Applies to mid- and high-technology militaries and states where doctrine and public belief matter for force employment;

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
material capacity alone does not predict outcomes.	escalation), then material parity can yield divergent outcomes.	doctrinal emphasis; doctrinal revisions referencing information effects; misallocations traceable to false beliefs induced by influence operations.		vulnerabilities, and delegate doctrine updates to combined arms and information-warfare centers of excellence.	limited where material asymmetries are overwhelming or where command-and-control is insulated from informational influence (e.g., highly centralized authoritarian militaries with strict information control).

## Notation

Symbol	Meaning	Units / Domain
$\mathbb{N}$	number of agents	$\mathbb{N}$
$(G_t=(V,E_t))$	time-varying communication/interaction graph	—
$\lambda_2(G)$	algebraic connectivity (Fiedler value)	—
$p$	mean packet-delivery / link reliability	[0,1]
$\tau$	latency / blackout duration	time
$\lambda$	task arrival rate	1/time
$e$	enforceability / command compliance	[0,1]
$\tau_{\text{deleg}}$	delegation threshold	[0,1]
<b>MTTA</b>	mean time-to-assignment/action	time
$P_{\text{fail}}$	deadline-miss probability	[0,1]

## Claim-Evidence-Method (CEM) Grid

Claim (C)	Evidence (E)	Method (M)	Status	Risk	TestID
Primary: Industrialization of information (automation, scale, algorithmic mediation) materially transforms organizational and societal cognition such that 'cognitive wars' – distinct cognitive equilibria that shape conduct and outcomes of war – emerge beyond material inputs.	[7] [4] [10] [2]	Theory development + process-tracing of case studies + cross-national empirical comparison using quantitative proxies of information-industrial capacity (proof/empirical/process-trace).	E cited; M pending empirical tests and comparative process-tracing (T1, T4).	If false, attributing changes in wartime conduct to information-industrialization rather than material or institutional factors would misdirect theory and policy; mitigation strategies premised on changing information architectures could be ineffective.	T1/T4
Primary: Industrial-scale automation and algorithmic mediation increase reliance on efficiency/throughput heuristics and routinized inference in operational decision-making (organizational preference for throughput over deliberation).	[3] [7] [2]	Textual coding of doctrines and SOPs, structured interviews with military/information specialists, and simulation of automated decision pipelines to measure heuristic substitution (textual/empirical/simulation).	E cited; M pending textual coding, interviews, and simulation validation (T2).	If wrong, reforms and training aimed at reversing throughput-driven failures may target the wrong causal lever; resilience investments (e.g., slowing decision loops) might not yield expected risk reductions.	T2
Primary: Algorithmic amplification (recommenders, attention economies) plus cheap automated actors (botnets) amplifies fringe narratives and shifts cascade thresholds, increasing probability of rapid belief cascades, misattribution, and escalatory incidents.	[4] [10] [5] [2]	Network diffusion simulations (with recommender dynamics and agent-based bot models) and empirical event studies linking amplification metrics to downstream cascade/escalation outcomes (simulation/empirical).	E cited; M pending simulation experiments and empirical event linking (T3).	If false, policy emphasis on countering algorithmic amplification as a driver of escalation may be overstated; resources might be better allocated to other escalation channels (kinetic, diplomatic).	T3
Secondary: Synthetic media (deepfakes and reenactment tools) systematically increase attribution uncertainty and erode epistemic trust, enabling adversaries to exploit delegated judgment and create strategic ambiguity.	[6] [7] [9]	Controlled experiments measuring trust and attribution under exposure to synthetic media; benchmark evaluation of detection algorithms; case studies of incidents involving synthetic media (experimental/empirical/benchmarks).	E cited; M pending lab experiments and detection-benchmark studies (T4).	If false, investments prioritized for deepfake detection and mitigation may not address the main drivers of attribution uncertainty; alternative sources of epistemic erosion would need attention.	T4

Claim (C)	Evidence (E)	Method (M)	Status	Risk	TestID
Secondary: Bureaucratization and professionalization of information functions (creation of specialized analytic/influence cadres) produces procedural metric-driven behavior that increases vulnerability to adversarial manipulation (preference for templated responses, automated attribution pipelines).	[7] [8] [11]	Process-tracing of organizational reforms, doctrinal text analysis, and interviews/surveys of information officers to detect procedural incentives and metric-driven behavior (qualitative/empirical).	E cited; M pending qualitative fieldwork and doctrinal coding (T5).	If wrong, organizational reforms aimed at changing staffing, incentives, or doctrine to reduce vulnerability could be ineffective or counterproductive.	T5
Secondary: Quantitative proxies of information-industrial capacity (bot prevalence, automated account activity, attention-share metrics) are predictive of influence operation success and can be used for early-warning and measurement.	[2] [5] [4]	Develop and validate quantitative proxies using observational social-media datasets; predictive modeling of influence outcomes (engagement, spread) and out-of-sample validation against known operations (empirical/data analysis).	E cited; M pending empirical model building and out-of-sample validation (T6).	If false, indicators used for early warning and attribution will have low signal-to-noise, producing false positives/negatives and misdirecting operational responses.	T6

## References (selected cited sources)

### Technical and empirical anchors cited in text:

[1]: A Survey of Multi-Agent Deep Reinforcement Learning with Communication. ArXiv.Org (2022). [2]: Dissecting a Social Botnet: Growth, Content and Influence in Twitter. ArXiv.Org (2016). [3]: Defending Against Adversarial Machine Learning. ArXiv.Org (2019). [4]: The role of online attention in the supply of disinformation in Wikipedia. ArXiv.Org (2023). [5]: The Rise of Social Bots. ArXiv.Org (2014). [6]: Face2Face: Real-time Face Capture and Reenactment of RGB Videos. ArXiv.Org (2020). [7]: The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. ArXiv.Org (2018). [8]: Detecting Propaganda Techniques in Memes. ArXiv.Org (2021). [9]: Exploring the Role of Visual Content in Fake News Detection. ArXiv.Org (2020). [10]: Efficient collective influence maximization in cascading processes with first-order transitions. ArXiv.Org (2016). [11]: Information Consumption and Boundary Spanning in Decentralized Online Social Networks: the case of Mastodon Users. ArXiv.Org (2022).

(Canonical foundational works and military-history anchors referenced conceptually in the Foundations and Theoretical Grounding sections include classic treatments of persuasion, threshold models, and military doctrine; these are used as theoretical anchors though not enumerated in the contemporary technical reference list above.)

# Adversarial Review

## Steelman Objections

- Causality not proven; industrial scale may merely correlate with reach.
- External validity risk: platform findings may not generalize to broadcast environments.
- Classifier vignette ignores base-rate sensitivity.
- Institutional mediation could dominate industrial capacity in pluralistic media.
- Measurement opacity remains; MTTA lacks public datasets for verification.

## Boundary Conditions

- High press freedom coupled with professional editorial standards.
- Distribution systems with high channel diversity and low single points of failure.
- Verification costs subsidized through default attestation workflows.

## Falsification Tests

- Natural experiments around ranking/labeling policy changes with survival analysis.
- Red-team drills measuring MTTA p50/p95 before/after instrumentation.
- Holdout experiments estimating amplification gain versus random baseline feeds.

## Decision Playbooks

KPI	Threshold	Action	Risk Delta	Monitoring Lag
MTTA_p95	<= 10 min	Automate detection and pre-authorize counter messaging	Decrease cascade risk	near-real-time
PPV(min)	>= 0.85	Raise decision threshold; require dual-sensor corroboration	Decrease false-trigger kinetic	per alert
t_half(persistence)	<= 7 days	Throttling + labeling + counter-speech	Shorten long-tail belief	daily
Amplification Gain (G)	<= 1.2x	Ranking demotion for low-trust narratives	Limit reach	weekly

# Evidence Ledger

## S1

2025-11-10 — Bootes multiwavelength survey (Chandra team) reported detection of ~4,000 X-ray sources using shallow 5 ks exposures across a 9 square degree field

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## S2

2025-11-10 — Spitzer IRAC team reported detection of ~300,000 point sources in the Bootes field

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## S3

2025-11-11 — Spitzer IRAC dataset contains ~30,000 point sources with detections in all four IRAC bands in Bootes

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## S4

2025-11-11 — Bootes survey team selected >400 AGN per square degree (vs SDSS at 12 AGN per square degree) from the multiwavelength data

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## S5

2025-11-12 — Survey authors identified 641 candidate obscured AGN in the IRAC-selected sample based on R-band and IRAC luminosities

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## S6

2025-11-12 — The team used X-ray stacking to verify obscuration and measured absorbing column densities  $N_H > 10^{22} \text{ cm}^{-2}$  for the obscured AGN candidates

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S7**

2025-11-13 — The Bootes team reported obscured AGN with high redshifts ( $z > 0.7$ ) and bolometric luminosities  $L_{bol} > 10^{45}$  ergs/s

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S8**

2025-11-13 — MMT/Hectospec team obtained modest-resolution optical spectra for about half of the X-ray sources (~2,000 spectra) and ~20,000 galaxies (out to  $z = 0.7$ )

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S9**

2025-11-10 — Canadian study authors surveyed 18 pediatric tertiary care hospitals with a 100% response rate

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S10**

2025-11-10 — Of the 18 hospitals surveyed, 13 had pediatric-specific massive hemorrhage protocol (MHP) content

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S11**

2025-11-11 — Among the 13 hospitals with pediatric MHP content: 8 were dedicated pediatric, 2 combined pediatric/obstetrical, and 3 combined pediatric/adult

## **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S12**

2025-11-11 — The survey found trauma accounted for 54% of MHP activations

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S13**

2025-11-12 — Ten of the 13 pediatric MHPs used specific blood volume anticipated/transfused over time as an activation trigger

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S14**

2025-11-12 — First transport container content was RBC-only in 8 pediatric MHPs

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S15**

2025-11-13 — Initial ratio-based blood component transfusion was recommended in 12 pediatric MHPs, with a quick transition to laboratory-guided transfusion using thresholds

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **S16**

2025-11-13 — All 13 pediatric MHPs reported use of tranexamic acid in their protocols

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## **EXEC1**

Large multiwavelength surveys (Chandra, Spitzer) produced massive catalogs including 4,000 X-ray sources, 300,000 IRAC detections, and 30,000 full-band sources. These enabled selection of over 400 AGN per square degree and identification of 641 candidate obscured AGN verified by X-ray stacking with  $\text{NH} > 10^{22} \text{ cm}^{-2}$ , in

### **Anchors**

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## MARKET1

Executive summary — The corpus points to a bifurcated market: one axis driven by data- and model-intensive R&D (multi-agent communication, adversarial ML, large multiwavelength surveys) and another by mission-critical operational services (content moderation/security, clinical emergency logistics). This creates concent

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## TECH1

Model architectures and chip developments: Recent work on communication-enabled multi-agent deep reinforcement learning (Comm-MADRL) highlights architectures that explicitly model inter-agent message channels, conditional routing, and attention-based broadcast vs. targeted communication patterns — a design space captur

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## COMP1

Winners and losers — identification

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## LENS1

## Operator Lens

### Anchors

- A Survey of Multi-Agent Deep Reinforcement Learning with Communication
- Dissecting a Social Botnet: Growth, Content and Influence in Twitter
- Defending Against Adversarial Machine Learning

## Sources

[1]

A Survey of Multi-Agent Deep Reinforcement Learning with Communication

Arxiv.Org, 2022-03-16. (cred: 0.50)

<http://arxiv.org/abs/2203.08975v2>

[2]

Dissecting a Social Botnet: Growth, Content and Influence in Twitter

Arxiv.Org, 2016-04-13. (cred: 0.50)

<http://arxiv.org/abs/1604.03627v1>

[3]

Defending Against Adversarial Machine Learning

Arxiv.Org, 2019-11-26. (cred: 0.50)

<http://arxiv.org/abs/1911.11746v1>

[4]

The role of online attention in the supply of disinformation in Wikipedia

Arxiv.Org, 2023-02-16. (cred: 0.50)

<http://arxiv.org/abs/2302.08576v1>

[5]

The Rise of Social Bots

Arxiv.Org, 2014-07-19. (cred: 0.50)

<http://arxiv.org/abs/1407.5225v4>

[6]

Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Arxiv.Org, 2020-07-29. (cred: 0.50)

<http://arxiv.org/abs/2007.14808v1>

[7]

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation

Arxiv.Org, 2018-02-20. (cred: 0.50)

<http://arxiv.org/abs/1802.07228v2>

[8]

Detecting Propaganda Techniques in Memes

Arxiv.Org, 2021-08-07. (cred: 0.50)

<http://arxiv.org/abs/2109.08013v1>

[9]

Exploring the Role of Visual Content in Fake News Detection

Arxiv.Org, 2020-03-11. (cred: 0.50)

<http://arxiv.org/abs/2003.05096v1>

[10]

Efficient collective influence maximization in cascading processes with first-order transitions

Arxiv.Org, 2016-06-08. (cred: 0.50)

<http://arxiv.org/abs/1606.02739v3>

[11]

Information Consumption and Boundary Spanning in Decentralized Online Social Networks: the case of Mastodon Users

Arxiv.Org, 2022-03-29. (cred: 0.50)

<http://arxiv.org/abs/2203.15752v3>

Generated: 2025-11-13T18:09:48.453785 | Word Count: 5552

## Research Roadmap

- **Phase 1 (Theory):** Formalize claims, extend proofs, validate against canonical results
- **Phase 2 (Simulation):** Implement stress tests, sweep parameter spaces, measure convergence/scaling
- **Phase 3 (Empirical):** Deploy in controlled environments, collect field data, validate predictions
- **Phase 4 (Integration):** Operationalize with human-in-loop, adversarial hardening, production deployment

**Confidence Methodology:** Confidence = 0.3·SourceDiversity + 0.25·AnchorCoverage + 0.25·MethodTransparency + 0.2·ReplicationReadiness, where SourceDiversity reflects unique publishers & types, AnchorCoverage reflects share of primary claims with Type-1 anchors, MethodTransparency reflects CEM completeness & assumptions ledger, and ReplicationReadiness reflects sim plan & datasets/params specified.

---

Prepared under the STI Research Program — theoretical framework subject to revision as data accumulate.