Article

Not peer-reviewed version

# Weaponization of the Social Sciences and Strategic Reconstruction of the National Academic System: A Policy Proposal

Wei Meng *

Posted Date: 8 September 2025

doi: 10.20944/preprints202509.0604.v1

Keywords: weaponisation of social sciences; cognitive warfare; national security; institutional resilience

*Article*

# Weaponization of the Social Sciences and Strategic Reconstruction of the National Academic System: A Policy Proposal

**Wei Meng** [1,2]

Dhurakij Pundit University, Thailand; wei.men@dpu.ac.th
Fellow, Royal Anthropological Institute, UK

**Abstract**

This study aims to explore the potential role and pathways for transforming social sciences within the national security and strategic competition landscape. Its objectives address pressing challenges including the inadequate alignment between social sciences and national strategy, insufficient conversion of academic findings into policy and practical tools, intensifying disinformation offensives and countermeasures, and uneven allocation of research resources. Methodologically, this study combines strategic intelligence analysis with policy evaluation, employing comparative case studies, interdisciplinary literature reviews, and institutional design frameworks. It draws upon the DARPA 'Heilmeier Eight Questions' and the EU's mission-oriented innovation model for its arguments. Findings indicate that by establishing a national-level social science research platform, optimising research funding mechanisms, fostering closer alignment between academic pursuits and strategic imperatives, and developing deepfake forensics alongside cognitive immunity systems, social sciences may evolve from traditional knowledge accumulation into resources serving strategic decision-making and cognitive competition. This transformative trajectory enhances political narrative power and cognitive dominance, bolsters institutional adaptability, and furnishes nations with combined soft and hard competitive advantages on the international stage. The conclusion posits that the strategic utilisation of social sciences is not a singular pathway but a multi-layered, multi-stage exploratory process, offering significant reference value for enhancing national security and international influence.

**Keywords:** weaponisation of social sciences; cognitive warfare; national security; institutional resilience

---

## Glossary

1. Weaponisation of Social Sciences: The process of converting social science research findings into tools for national strategy, intelligence operations, and cognitive warfare.
2. Cognitive Immunity: Enhancing the public's ability to identify and resist disinformation through education and social communication.
3. Firehose Model: An attack pattern employing massive volumes of disinformation to disrupt public cognition.
4. Deepfake Technology: Artificial intelligence-generated false audio-visual content utilised for information manipulation.

## I. Strategic Context

The rapid advancement of globalisation and digitalisation has given rise to a new paradigm of competition centred on information warfare and cognitive warfare. Within this framework, international standing is no longer determined directly by military or economic might; rather, the capacity to control intelligence and information dissemination mechanisms has become pivotal to

 doi:10.20944/preprints202509.0604.v1



national stability and strategic advantage. As Putin stated in his 2017 Knowledge Day address: 'Artificial intelligence is the future... Whoever becomes the leader in artificial intelligence will become the ruler of the world' (Putin, 2017). This not only underscores the strategic importance of artificial intelligence but also reveals a deeper understanding of cognitive and information control. It profoundly illuminates the strategic logic that 'whoever controls information and cognitive tools holds the strategic initiative'. This heralds a new era where, in the 21st century, the social sciences are no longer confined to ivory-tower academic pursuits but constitute vital components of national strategic assets. During the Trump administration, the United States weakened its intelligence apparatus, compromising its defences in the information and cognitive domains. Renowned CIA historian Tim Weiner asserted that Trump 'placed national security in the hands of incompetents and puppets,' inflicting severe detrimental effects on intelligence decision-making and international cooperation (The Guardian). This illustrates how nations become vulnerable to disinformation and cognitive warfare attacks when social sciences—including psychology, cognitive science, and communication studies—are not effectively integrated into intelligence systems. Concurrently, the historical failure of Chiang Kai-shek's regime in narrative control and 'psychological warfare' eroded the foundational legitimacy of its rule (Paul & Matthews, 2016; Kavanagh & Rich, 2018). Such cases demonstrate the strategic value of social sciences in shaping political legitimacy and controlling narratives, while underscoring their importance for national stability. In summary, unless social science research transforms from static academic resources into dynamic national strategic tools, states will be unable to effectively address the novel security challenges of the information age.

## II. Strategic Issues

### *2.1. Disconnect Between Academia and the State*

A significant volume of social science research with strategic potential remains excluded from national security or strategic decision-making processes. This structural disconnect constrains the influence of social sciences on core national interests. Research indicates that the early development of social sciences in the United States relied heavily on defence funding, which both propelled disciplinary growth and introduced ethical dilemmas concerning the alignment of academic pursuits with national interests – a contradiction that remains unresolved to this day (Stimson Center, 2023).

### *2.2. Lack of Scientific Support for Intelligence Systems*

Research findings from fields such as psychology, communication studies, and cognitive science have yet to be systematically integrated into intelligence practices, resulting in inadequate strategic responsiveness by the state in cognitive warfare and information warfare. Take the US Department of Defence's Minerva programme as an example: originally designed to bolster national security by funding social science research, it faced substantial funding cuts in 2025. This shift underscores the disconnect between social science and national strategy (Kupferschmidt, 2025).

### *2.3. Imbalance in Disinformation Attack and Defence*

Hostile actors persistently launch disinformation campaigns through the 'Firehose model' and deepfake technology, while national investment in defensive research remains inadequate. Existing research indicates that deepfake technology significantly amplifies the destructive impact of disinformation on public perception. Without effective countermeasures, democratic institutions and national security face grave threats (Chesney & Citron, 2019).

### *2.4. Imbalanced Allocation of Research Funds*

Research funding allocation exhibits significant bias, with substantial resources directed towards studies unrelated to strategic applications, while high-value research pertinent to national security and cognitive warfare receives inadequate support. The academic community has

 doi:10.20944/preprints202509.0604.v1



highlighted that the research funding system suffers from “path dependency”, tending to perpetuate existing models rather than stimulating interdisciplinary and strategically significant research (Laird, 2020; Issues in Science and Technology, 2023).

## III. Strategic Objectives

**G1: Establishment of "weaponisation of social sciences" as a national strategic priority area**

To transform the social sciences from a “stock of knowledge” into a “strategic variable”, it is imperative to first establish them as a priority domain at the national level, institutionally integrating them into the primary battleground of national security and comprehensive national power competition. In an era where “fact decline” coexists with the narrative battlefield, the systematic contestation of the cognitive domain has become the foremost challenge for governance and security (Kavanagh & Rich, 2018). Transnational security communities are likewise reshaping operational concepts around ‘cognitive warfare,’ emphasising the sustained shaping of individual and collective mental spaces (du Cluzel, 2021). The latest US military doctrine explicitly designates ‘information superiority’ as a foundational capability for joint operations, signifying that social science knowledge (psychology, communication, social networks, behavioural economics) constitutes both the prerequisite investment and the amplifier for generating such advantages (U.S. Army, 2023; U.S. Air Force, 2023). Accordingly, at the national level, weaponising the social sciences should be established as a hard constraint priority. Strategic resources and routine assessments should be allocated at the highest level of design, placing it on a par with cyber, electronic warfare, space, and artificial intelligence.

**Key Initiative Pathways:**

Incorporate the “weaponisation of social sciences” into national strategic planning as part of five- to ten-year medium-to-long-term frameworks for national security and scientific innovation (at the same level as cybersecurity and AI). Establish dedicated budgetary provisions for a “Cognitive and Narrative Testing Ground” initiative and a “Behavioural Science Operations Unit” development programme. Assess annual Key Performance Indicators (KPIs) to measure external narrative influence, domestic resilience metrics, and cross-departmental conversion rates.

**G2: Embedding Social Science Results in Intelligence and Cognitive Warfare Systems**

The core of embedded transformation lies in integrating peer-reviewed and empirically validated methodologies into the intelligence cycle, information operations, psychological operations, and integrated civil-military narrative frameworks. Classic evidence in intelligence analysis has long emphasised that managing cognitive biases, employing analytical reasoning, and conducting structured analysis constitute fundamental pathways to reducing judgmental errors (Heuer, 1999). When countering the deluge of false narratives from ‘Firehoses,’ relying solely on post-hoc debunking proves inadequate; proactive design must be implemented in dissemination cadence, channel positioning, and psychological cues (Paul & Matthews, 2016). Joint operations doctrine and air force regulations now routinely integrate information/influence activities as core joint warfare elements, mandating joint deployment, training, and evaluation across operational and research lines (U.S. Air Force, 2023; U.S. Army, 2023).

**Key Initiative Pathways:**

Embed the psychological-communication-behavioural-network science team within the intelligence cycle nodes (collection-analysis-output-evaluation), establishing a closed-loop ‘research-action-evaluation’ metric system (e.g., audience segmentation reach rates, belief update resilience, narrative diffusion R-values). Establish ‘red-blue adversarial’ and ‘multi-armed bandit’ narrative experimentation platforms to iteratively refine intervention strategies online (A/B/n testing, causal inference, Bayesian adaptation). Incorporate SOCINT (Social Intelligence) units within task formations to coordinate with OSINT, HUMINT, and CYBINT.

**G3: Restructuring of research funding mechanisms and establishment of a "Strategic Value Coefficient (SVC)"**

 doi:10.20944/preprints202509.0604.v1



Traditional academic funding is susceptible to 'path dependency' and 'disciplinary self-reinforcement', making it difficult to provide timely support for high-risk, high-externalities, cross-domain missions (Mazzucato, 2018). An evaluation framework centred on 'mission-oriented' principles and DARPA's Heilmeier criteria should be established to develop a Strategic Value Coefficient (SVC) tailored to national security and operational narratives. This should primarily quantify five dimensions—'problem significance, potential operational gains, verifiable milestones, transferable chains, and time-cost-risk'—supplemented by peer review, shifting focus from 'paper output' to 'operational applicability and strategic leverage'. For areas with significant strategic potential but where short-term paper output is challenging, accelerate progress through a 'sprint funding + milestone-based formalisation' mechanism (DARPA, n.d.; European Commission, 2018).

**Key Initiative Pathways:**

Formulate a one-page mission statement based on Heilmeier's eight questions as a prerequisite for initiating all social science security projects. Establish 'application verification milestones' and 'mission accomplished metrics' with mandatory funding disbursement linkage. Create a 'rapid transformation fund' and 'joint testing consortium' to interface task forces with frontline departments.

**G4: Building a "Social Science Weaponisation Platform": Integrating Psychological Warfare, Narrative Competition, Historical Defence and "Cognitive Immunity"**

The platform's objective is to integrate attack, defence and resilience into a unified framework: upstream, leveraging threat intelligence and narrative mapping for proactive identification; midstream, employing psychological/behavioural interventions and multi-channel delivery for precise shaping; downstream, establishing foundational public resilience and cognitive immunity. The NATO framework and multiple research institutions have identified deepfakes and the "cognitive dimension" within hybrid warfare as structural risk points (NATO StratCom COE, 2020). The WHO, meanwhile, has provided a governance framework for infodemics and public health resilience solutions, emphasising a full-chain system of "monitoring-identification-intervention-evaluation-empowerment" (Tangcharoensathien et al., 2020; WHO, 2021). Drawing on countermeasure insights from the 'Firehose' model, social science platforms may be structured across four tiers: technology, methodology, organisation, and evaluation. Technology enables multimodal detection and attribution; methodology underpins narrative design and audience modelling; organisational frameworks ensure cross-departmental coordination; and evaluation establishes auditable operational KPIs/KPAs (Key Performance/Activity Indicators) (Paul & Matthews, 2016; NATO StratCom COE, 2020; WHO, 2021).

**Key Initiatives:**

Establish a closed-loop platform integrating 'narrative mapping—impact assessment—iterative intervention' (including deepfake forensics, source tracing, and cross-platform diffusion monitoring). Transplant the four modules of the WHO infodemic framework—'continuous monitoring, targeted interventions, effectiveness evaluation, and social empowerment'—into a national cognitive immunity engineering programme. Develop a 'historical defence' mechanism: construct baseline narratives, evidence repositories, and visualised educational materials for critical historical issues, integrating these into education and public communication systems.

## IV. Programme of Action

### *4.1. National Institute for the Weaponisation of Social Sciences*

Positioning and Organisation: Establish a national-level research institute that is interdisciplinary, cross-departmental, and cross-domain. Centred on the triad of 'intelligence-information-cognition', it will routinely bring together expertise in political science, psychology, sociology, communication studies, and artificial intelligence/data science. This will create a closed-loop system of 'research-validation-deployment-evaluation' (du Cluzel & Claverie, 2021; U.S. Army, 2023; U.S. Air Force, 2023). The institute shall comprise four mission centres: ① Cognitive Warfare

 doi:10.20944/preprints202509.0604.v1



and Information Superiority (operational concepts and metric systems); ② Narrative Countermeasures and Audience Modelling (cross-platform dissemination, public sentiment/mind mapping, causal inference); ③Deepfake Forensics and Multimodal Authentication (text/voice/image/video); ④ Evaluation and Ethical Governance (Auditable Chains, Red Line Lists, Ethical/Legal Safeguards). (du Cluzel & Claverie, 2021; U.S. Army, 2023; U.S. Air Force, 2023).

**Key Deliverables:**

Cognitive Warfare Algorithms and Operational Workflows: Targeting information superiority by integrating structured analysis, bias correction, adversarial narrative generation, and channel pacing optimisation in parallel. This forms measurable, reproducible, and joint-training-capable workflows (Heuer, 1999; U.S. Army, 2023; U.S. Air Force, 2023).

Narrative Counter-Framework: Integrating anticipatory countermeasures from the 'firehose' model to establish an intervention sequence: 'pre-emptive narratives – credible anchors – multi-source reuse – effect attribution' (Paul & Matthews, 2016).

Public Sentiment/Influence Operations Tools: Within legal and ethical boundaries, employ micro-segmentation of target audiences and multi-armed bandit online experiments to optimise reach and persuasion efficiency (U.S. Air Force, 2023; U.S. Army, 2023).

### *4.2. Weaponisation Orientation of Research Funding*

Mechanism Reconstruction: Centred on a 'Mission-oriented + Heilmeier Principles' framework, establish a Strategic Value Coefficient (SVC). This involves a five-dimensional quantitative assessment of 'problem significance, potential operational gains, verifiable milestones, transferable chains, and time-cost-risk', tightly coupled with funding allocation rhythms. This rewrites the singular 'paper output-oriented' incentive (Mazzucato, 2018; DARPA, n.d.). The funding structure adopts a three-stage pathway: 'Frontier Sprint + Milestone Validation + Joint Testing and Transfer', enabling rapid initiation, swift damage control, and expedited transfer for high-risk interdisciplinary research (Mazzucato, 2018; DARPA, n.d.).

Implementation Key Points:

Project Brief: All projects are initiated using Heilmeier's Eight Questions, explicitly stating 'Why it matters', 'How to validate', and 'Success metrics and exit criteria' (DARPA, n.d.).

Balancing External Impact and Operational Readiness: For directions where short-term publication is challenging but significant operational applicability exists, establish 'mission accomplished metrics' (application validation milestones).

Joint Testing and Collaborative Evaluation: Establish a tripartite evaluation mechanism integrating 'research-intelligence-action' to prioritise real-world effectiveness. (DARPA, n.d.; Mazzucato, 2018)

### *4.3. Intelligence and Social Science Integration Project*

System Embedding: Integrate psychological, communication and behavioural science methodologies throughout the intelligence cycle and influence operations, establishing unified operational metrics spanning collection, analysis, output and evaluation (Heuer, 1999; U.S. Army, 2023; U.S. Air Force, 2023). At the doctrine level, with 'information superiority' as the objective, clarify the division of labour and coordination among military branches and civilian departments, alongside joint training, joint evaluation, and routine task force organisation (U.S. Army, 2023; U.S. Air Force, 2023). (Heuer, 1999; U.S. Army, 2023; U.S. Air Force, 2023)

Red-Blue Engagement Range: Establishing a cognitive and narrative red-blue adversarial experimentation framework:

Red Teams conduct adversarial narrative experiments based on MITRE ATT&CK® adversarial patterns and the Adversary Emulation Programme (AEP); Blue Team implements combined interventions of 'proactive prompts—trusted anchors—alternative narratives', evaluating effectiveness via KPP/KPA (Key Performance/Activity Metrics) (UFMCS, 2019; MITRE, 2017; Paul & Matthews, 2016). Simultaneously, multi-source telemetry and causal inference frameworks (e.g.,

 doi:10.20944/preprints202509.0604.v1



synthetic control, difference-in-differences) generate auditable evidence, integrated into procurement and funding closed-loop systems (U.S. Army, 2023). (UFMCS, 2019; MITRE, 2017; Paul & Matthews, 2016; U.S. Army, 2023)

Civil-Military Synergy and External Engagement: For foreign-related matters, implement a rolling iteration of the 'research-action-evaluation' approach, directly applying research outcomes to intelligence products, public diplomacy, and negotiation scenarios. This yields outputs comprising 'narrative operations plans + audience matrices + evaluation dashboards' (U.S. Air Force, 2023; U.S. Army, 2023).

### *4.4. Deep False Forensics and the Cognitive Immune System*

Technical Capabilities: Construct a multimodal deepfake forensics platform integrating semantic forensics with cross-modal consistency verification, enabling a comprehensive chain of 'detection-attribution-traceability-response'. Interface with content platforms, judicial appraisal, and emergency response systems (NATO StratCom COE, 2020; DARPA, 2025; DARPA, n.d.). The platform shall integrate academic frontiers (e.g., semantic forensics) with engineering standards, establishing comparable testing benchmarks and red team sample repositories (NATO StratCom COE, 2020; DARPA, 2025). (NATO StratCom COE, 2020; DARPA, 2025; DARPA, n.d.)

Public Resilience/ Cognitive Immunity: Integrate the WHO's five action streams for infodemic governance and UNESCO's Media and Information Literacy (MIL) curriculum framework into national education and civil service training. Establish a routine mechanism of 'monitoring—identification—intervention—evaluation—empowerment' synchronised with national cyber/public health systems (Tangcharoensathien et al., 2020; WHO, 2021; UNESCO, 2021/2022). For high-risk topics, establish a 'baseline narrative–evidence repository–visualised teaching materials' framework to conduct historical defence and scenario simulations, enhancing societal resilience against manipulation (Tangcharoensathien et al., 2020; WHO, 2021; UNESCO, 2021/2022).

## V. Expected results

### *5.1. Enhanced Political Control: Direct Conversion of Social Science Findings into Strategic Manipulation Tools*

Embedding methodologies from psychology, communication studies, and cognitive science within intelligence and information operations can systematically reduce analytical bias, strengthen structured analysis and evidence chain assessment, thereby enhancing the verifiability and consistency of decision-making (Heuer, 1999). At the regulatory level, establishing 'information superiority' as a core capability for joint operations and national security signifies that influence and counter-influence designs underpinned by social sciences can proactively shape agendas and audience perceptions, thereby enhancing political control and policy implementation synergy (U.S. Army, 2023; U.S. Air Force, 2023). Pre-emptive intervention and pacing control against high-frequency false narratives—akin to a 'Firehose' approach—have also demonstrated superior cost-effectiveness and sustainability compared to post-hoc clarification (Paul & Matthews, 2016).

### *5.2. Establishing Cognitive Dominance: Taking the Initiative in International Narrative Competition*

When social science outputs are transformed into deployable tools such as 'narrative counter-frame frameworks,' 'audience modelling and channel orchestration,' and 'contextualised evidence anchors,' nations can establish pioneering narratives and credible anchors across multi-platform opinion arenas. This creates measurable 'impact trajectories' and 'diffusion R-values,' thereby securing initiative in narrative warfare (Paul & Matthews, 2016). . Joint operational doctrine institutionalises the organisation of information/influence activities and mandates 'information superiority' as an operational requirement, effectively establishing the organisational and procedural foundations for cognitive dominance (U.S. Army, 2023; U.S. Air Force, 2023). Concurrently,

 doi:10.20944/preprints202509.0604.v1



systematic capabilities for forensics and governance against deepfakes and hybrid warfare can diminish adversaries' ability to achieve 'perceptual surprise' during critical periods, thereby stabilising our narrative ecosystem (NATO StratCom COE, 2020).

*5.3. Increased Institutional Resilience: Scientific Research-Intelligence-Strategy Trinity to Enhance National Security*

The closed-loop system of "research-action-evaluation", centred on social sciences, enables the implementation of end-to-end governance encompassing "monitoring-identification-intervention-evaluation-empowerment" within infodemics and cross-platform rumour-mongering environments. This enhances the adaptability, learning capacity, and feedback velocity of governmental and societal systems (Tangcharoensathien et al., 2020; WHO, 2021). This mechanism, which deeply couples knowledge production with action deployment, helps maintain institutional coherence and public trust in complex environments. It reduces societal fragmentation in scenarios of 'fact decay,' thereby strengthening the structural resilience of national security (Kavanagh & Rich, 2018).

*5.4. International Status Enhancement: Weaponisation of Social Sciences as a Combination of Hard and Soft Power*

The weaponisation of social sciences is not merely a narrow extension of "information warfare", but rather the integration of hard power's deterrent foundation with soft power's appeal, thereby forming a quantifiable and assessable "smart power" combination (Nye, 2004; Armitage & Nye, 2007). When narrative dominance, evidence-gathering capabilities, and public resilience operate in concert, nations find it easier to secure a leading position within international institutions and public opinion structures. This enhances their capacity to shape rules and build international trust, thereby translating into sustained gains in status and interests (Armitage & Nye, 2007; U.S. Air Force, 2023).

## VI. Risks and Mitigation

*6.1. Academic Freedom and Ethical Risks: Towards an Ethical + Strategic Values (ESV) Mechanism*

The weaponisation of social sciences, lacking standardised ethical governance, readily leads to curtailed academic freedom, distorted researcher incentives, and irreversible harm to subjects or society. To this end, existing research ethics frameworks—such as the Belmont Report's three principles for human subjects and the Menlo Report's extensions for ICT/security research—should be augmented with auditable assessments incorporating national security and strategic value dimensions: Establish a dual-review system ('ESV dual review') combining IRB/ethical review with a Strategic Value Committee to jointly assess risk-benefit, rights-security, and research necessity-substitutability trade-offs. This must operate with traceability and governability as foundational principles (OHRP, 1979; DHS, 2012; DoD, 2020). To align compliance thresholds for AI research institutions and funders with first-principles governance, establish a single, auditable line of sight from ethics to evidence: anchor high-level norms in the OECD AI Principles and UNESCO's Recommendation to risk-proportionate, testable controls in NIST SP 800-53/53A, such that each principle resolves into enforceable procedure, telemetry, and artefacts. Human rights protections become ex-ante impact assessments, contextual risk registers, and mitigation plans traceable to RA and governance (GV, PM/PL) controls; transparency crystallizes as public system/model cards, decision registries, data-lineage records, and disclosure change-logs evidenced via AU plus documentation planning (PL); accountability is operationalized through RACI-clear role design, escalation and sanction regimes, exception tracking, and immutable audit trails (again PM/PL with AU as proof); human oversight is guaranteed by documented on-/in-the-loop checkpoints, privilege boundaries, dual-control for sensitive actions, and reversal/kill-switch runbooks mapped to AC, PM/PL, and incident response (IR). Thresholds are tiered by impact rather than form: at baseline every project produces a DPIA/AI risk assessment, access matrix, enabled audit logging, system card,

 doi:10.20944/preprints202509.0604.v1



and an oversight plan; elevated-risk work adds bias/robustness test batteries, red-team protocols, model change control with rollback criteria, and pre-deployment sign-off; rights-critical deployments trigger independent review, continuous monitoring with alert KPIs, tamper-evident logs, and grievance and remedy channels external to the project team. Evidence is gated across the lifecycle to compress ambiguity: design yields a principles-to-controls traceability matrix (OECD/UNESCO → control IDs), build yields configured AC roles, AU schemas, test reports and MCC records, deploy yields oversight SOPs and IR playbooks plus public-facing artefacts, and operate yields quarterly audit excerpts, exceptions backlogs with remediation SLAs, drift and incident metrics, and stakeholder-feedback ledgers. The minimal KPIs that matter are similarly few and falsifiable—coverage of principle→control traceability, audit-log completeness and tamper-evidence, mean time to risk remediation, proportion of privileged actions requiring human approval, and transparency artefact coverage—so that funders can discriminate maturity without prescribing method. This architecture turns ethics into control logic and log evidence (OECD, 2019; UNESCO, 2021; NIST, 2020/2022), reduces governance to verifiable claims, and couples accountability to telemetry, enabling proportionate oversight that protects rights, sustains scientific integrity, and scales with risk rather than rhetoric.Concurrently, for classified or sensitive tasks, a closed-loop approach is adopted: 'pre-research registration – phased milestones – public disclosure of results (without compromising security or privacy) – independent review', balancing academic freedom with national security (DoD, 2020; NIST, 2023).

**Key points for phased implementation:**

Dual-track review: IRB/ethical review + strategic value assessment (ESV); introduction of external ethics observers and security/confidentiality officers for high-sensitivity tasks (DoD, 2020). Align compliance with rights-based governance by making ethics empirically testable and proportional to risk: require end-to-end traceability of the model–data–decision chain with immutable, time-synchronized logs; versioned datasets, code, and model cards bound to change-management and incident-response runbooks that define ex-ante rollback criteria and dual-control gates for privileged actions, all mapped to verifiable NIST SP 800-53/53A control families (AU for auditability, AC for access, RA for risk registers and impact assessments, PM/PL for governance and documentation, IR for kill-switch/rollback, CM for configuration evidence) (NIST, 2020; 2022). Institutionalize audit as a statistical act, not a ritual—pre-register sampling frames and ratios justified by power analyses and publish coverage against plan—so that logs, tests, and exceptions accumulate as decision-grade evidence rather than post-hoc narrative. In parallel, embed human rights and oversight by design: publish actionable transparency artifacts (system/model cards, data lineage, release notes); execute fairness and robustness evaluations against preset thresholds with corrective action plans and re-test criteria; operate accessible redress channels with time-bound remedies; and codify human-in/on-the-loop checkpoints tied to escalation paths and reversal authority—each element explicitly benchmarked to the OECD AI Principles and UNESCO's Recommendation on AI Ethics (OECD, 2019; UNESCO, 2021). Funders and institutions should evaluate with a compact, falsifiable KPI set—log completeness and tamper-evidence rates, proportion of privileged actions requiring human approval, mean time to risk remediation, fairness-drift and exception-closure rates, and audit adherence to the declared sampling plan—creating a single, auditable line of sight from normative commitments to operational controls and measurable outcomes, and ensuring governance scales with impact rather than rhetoric.

### *6.2. International Criticism and Diplomatic Pressure: External Representation in a "Cognitive Defence/Social Resilience" Framework*

The external environment may interpret the "weaponisation of social sciences" as an offensive information operation, triggering public opinion and diplomatic pressure. To mitigate sensitivity, external narratives should be positioned as a 'cognitive defence—social resilience—public health governance initiative,' aligning with the WHO's infodemic governance framework and its four-pronged approach of 'listening—understanding—building resilience—empowering.'

 doi:10.20944/preprints202509.0604.v1



Make the core duties—compliance, transparency, accountability, and human oversight—empirically testable and proportionate to risk: obligations should be observable (decision logs, evaluation traces, release notes), auditable (independent review against pre-registered sampling frames), and enforceable (role-bound approvals, reversal authority, redress with time-bound remedies), so governance is evidenced rather than asserted (Tangcharoensathien et al., 2020; WHO, 2021; UNESCO, 2021).Calibrate oversight to the EU AI Act's taxonomy and conformity logic by first fixing intent and scope—declare the system's purpose, map affected populations, and classify the use case—then, for high-risk applications, require an open risk dossier that is versioned, machine-readable, and lifecycle-updated: include hazard analysis and risk appetite; data-governance controls with lineage and access matrices; performance, robustness, and fairness metrics disaggregated across cohorts and conditions; stress tests and adversarial red-team results; a residual-risk rationale tied to mitigation plans; human-oversight design with explicit in/on-the-loop authority, escalation paths, and reversal criteria; continuous monitoring and incident-response playbooks; ex-ante rollback triggers; and third-party assessment outcomes aligned to the Act's conformity regime (European Union, 2024). For China, institutionalizing routine publication of these dossiers—paired with a compact, falsifiable KPI set (audit coverage versus plan, fairness drift, mean time to remediation, grievance-resolution rates)—would demonstrate regulatory capability and goodwill, align domestic practice with international norms, and create a verifiable chain from principle to control to evidence to outcome, ensuring that oversight scales with impact rather than rhetoric.

**Key points for gradual implementation:**

External transparency: Annual white paper on 'Cognitive Defence and Social Resilience'; publication of high-risk scenario impact assessments and grievance mechanisms (European Union, 2024; UNESCO, 2021).

Multilateral alignment: Sharing harm reduction indicators and evaluation frameworks with WHO/UN mechanisms to establish international public goods attributes (WHO, 2021).

Terminology Desensitisation: Equivalently replacing 'weaponisation' in diplomatic texts with 'cognitive defence/social resilience building' to align with public health governance discourse (Tangcharoensathien et al., 2020).

### *6.3. Risks of Technology Misuse: Hierarchical Permissions + Auditable Chain + Dual-Use (DURC) Governance*

Tools emerging from the convergence of social sciences, intelligence, and AI exhibit pronounced dual-use characteristics. Throughout their lifecycle, they require implementation of a five-ring control framework encompassing 'minimum privilege, tiered governance, real-time auditing, red team simulation, and post-incident accountability.' This must be operationalised across requirements, design, training, evaluation, and deployment by referencing the four-function closed-loop of governance, measurement, management, and mapping outlined in NIST AI RMF 1.0 (NIST, 2023). Anchor security and privacy in NIST SP 800-53/53A by turning requirements into verifiable control–evidence pairs: enforce least-privilege and dual control for privileged actions (AC), generate immutable, time-synchronized audit trails (AU), pre-plan containment and rollback through incident runbooks (IR), maintain ex-ante risk registers and impact analyses (RA), attest secure development and supply-chain provenance (SBOM, vendor assurances) via SA and SR, and bind program accountability to named owners and review cadences (PM)—thereby creating end-to-end traceability across access, logging, incident response, and supply-chain assurance (NIST, 2020; 2022). Where research or data carry significant diffusion risk, elevate to DURC/PPPR/P3CO gates: require dual institutional and national-level review, restrict dissemination of sensitive methodological details and data through controlled access and embargoes, maintain continuous oversight with revocation authority, and codify explicit penalties and civil liabilities for breach—aligning incentives with biosecurity while preserving legitimate inquiry (NSABB, 2014; U.S. Government, 2024).

In defence contexts, adhere to the 'traceable and governable' requirements within the DoD AI Ethics Principles, conducting regular impact assessments and outage/rollback drills (DoD, 2020).

 doi:10.20944/preprints202509.0604.v1



Key points for mitigating risks:

Classify and segregate high-sensitivity models, data, and tools into physically and logically controlled domains with sealed network paths, escrowed keys, and just-in-time least privilege; release only redacted, de-weaponized artifacts externally to decouple capability from misuse (NIST, 2020). Make the pipeline auditable end-to-end via four time-synchronized, hash-chained ledgers—data, models, instructions, outputs—cross-indexed for provenance reconstruction and subjected to independent annual assurance and adversarial red-team penetration testing with tracked remediation SLAs (NIST, 2022; NIST, 2023). Embed dual-use governance from inception: route proposals through DURC/PPPR (P3CO) reviews; pre-authorize misuse scenarios, cut-outs, and disclosure limits; bind containment and rollback playbooks to explicit escalation authority; and codify restricted dissemination with penalties and civil liabilities—preserving legitimate inquiry while aligning incentives with biosecurity (U.S. Government, 2024; NSABB, 2014).

## VII. Conclusions and Recommendations

Within the global competitive landscape of information and cognitive warfare, nations must prioritise the comprehensive weaponisation of social sciences as a strategic imperative. This transcends mere academic transformation, fundamentally concerning the reshaping of national security, institutional resilience, and international discourse. Firstly, establishing a national-level institute for weaponising social sciences is imperative to integrate interdisciplinary capabilities. This involves embedding political science, psychology, sociology, communication studies, and artificial intelligence systems within strategic research and policy implementation, thereby constructing a verifiable, deployable 'cognitive warfare toolkit' (du Cluzel & Claverie, 2021; Heuer, 1999).

Secondly, research funding mechanisms must be restructured, adopting a 'strategic value coefficient' as the core evaluation criterion. This ensures resources flow towards research yielding strategic offensive and defensive gains, minimising waste while drawing upon the EU's institutional experience in mission-oriented innovation (Mazzucato, 2018).

Thirdly, deep integration between intelligence and social sciences must be advanced. Through establishing red-blue adversarial testing grounds and narrative counter-frame frameworks, social science research should be translated into actionable intelligence products and cognitive intervention measures, securing the nation's proactive position in global information warfare and narrative combat (Paul & Matthews, 2016; U.S. Army, 2023; U.S. Air Force, 2023).

Finally, a deepfake forensics and cognitive immunity system must be established. By creating a national-level deepfake detection platform and cognitive immunity education framework, we can effectively counter the infiltration of deepfakes and disinformation, enhancing public resilience and institutional trust. This approach aligns not only with the WHO's framework for managing infodemic crises but also with UNESCO's initiatives on AI ethics and societal resilience (Tangcharoensathien et al., 2020; WHO, 2021; UNESCO, 2021).

In summary, through the integrated approach of research institute development, scientific funding mechanism adjustments, intelligence-academic convergence, and cognitive immunity systems, the nation stands poised to secure cognitive dominance in global strategic competition, enhance institutional resilience, and establish its status as a major power in the 21st century.

## References

1. Armitage, R. L., & Nye, J. S. (2007). CSIS Commission on Smart Power: A smarter, more secure America. Center for Strategic and International Studies. https://csis-website-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/media/csis/pubs/071106_csissmartpowerreport.pdf
2. Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753–1819. https://scholarship.law.bu.edu/faculty_scholarship/640
3. Clifford, C. (2017, September 4). Putin says the nation that leads in AI 'will be the ruler of the world'. CNBC. https://www.cnbc.com/2017/09/04/putin-ai-leader-ruler-of-world.html

 doi:10.20944/preprints202509.0604.v1




4. DARPA. (n.d.). The Heilmeier Catechism. https://www.darpa.mil/about/heilmeier-catechism
5. Department of Defense. (2020, February 24). DOD adopts ethical principles for artificial intelligence. https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/
6. du Cluzel, F. (2021). Cognitive warfare. NATO Innovation Hub. https://innovationhub-act.org/wp-content/uploads/2023/12/20210113_CW-Final-v2-.pdf
7. European Commission. (2018). Mission-oriented research & innovation in the European Union: A problem-solving approach to fuel innovation-led growth (M. Mazzucato). https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/mission-oriented-research-innovation-eu-problem-solving-approach-fuel-innovation-led-growth_en
8. European Union. (2024). Artificial Intelligence Act (Regulation (EU) 2024/1689). https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
9. Heuer, R. J., Jr. (1999). Psychology of intelligence analysis. Central Intelligence Agency. https://www.cia.gov/resources/csi/books-monographs/psychology-of-intelligence-analysis-2/
10. Issues in Science and Technology. (2023). The next 75 years of science policy. https://issues.org/wp-content/uploads/2023/02/The-Next-75-Years-of-Science-Policy.pdf
11. Kavanagh, J., & Rich, M. D. (2018). Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life. RAND Corporation. https://doi.org/10.7249/RR2314
12. Kupferschmidt, K. (2025, March). Pentagon guts national security program that harnessed social science. Science. https://www.science.org/content/article/pentagon-guts-national-security-program-harnessed-social-science
13. Laird, F. N. (2020). Path dependency and the problems of government funding. American Journal of Public Health, 110(6), 723–729. https://doi.org/10.2105/AJPH.2020.305648
14. National Institute of Standards and Technology. (2020). SP 800-53 Rev.5: Security and privacy controls for information systems and organizations. https://doi.org/10.6028/NIST.SP.800-53r5
15. National Institute of Standards and Technology. (2022). SP 800-53A Rev.5: Assessing security and privacy controls in information systems and organizations. https://doi.org/10.6028/NIST.SP.800-53Ar5
16. National Institute of Standards and Technology. (2023). AI Risk Management Framework (AI RMF 1.0). https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf
17. National Science Advisory Board for Biosecurity. (2014– ). About NSABB. https://osp.od.nih.gov/policies/national-science-advisory-board-for-biosecurity-nsabb/
18. NATO Strategic Communications Centre of Excellence. (2020). Deepfakes – Primer and forecast. https://stratcomcoe.org/publications/deepfakes-primer-and-forecast/42
19. Nye, J. S. (2004). Soft power: The means to success in world politics. PublicAffairs. https://www.publicaffairsbooks.com/titles/joseph-s-nye-jr/soft-power/9781610390699/
20. Office for Human Research Protections. (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf
21. Organisation for Economic Co-operation and Development. (2019). OECD AI principles. https://oecd.ai/en/ai-principles
22. Paul, C., & Matthews, M. (2016). The Russian "firehose of falsehood" propaganda model: Why it might work and options to counter it. RAND Corporation. https://doi.org/10.7249/PE198
23. Putin, V. (2017, September 1). Address on Knowledge Day to Russian students. Reported in The Verge. https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world
24. Putin. (2017, September 4). Whoever becomes the leader in [AI] sphere will become the ruler of the world. The Verge. https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world
25. Putin. (2024, February 8). Vladimir Putin said that the country that led the development of artificial intelligence (AI) would become the "ruler of the world". The Economist. https://www.economist.com/business/2024/02/08/vladimir-putin-wants-to-catch-up-with-the-west-in-ai
26. Stimson Center. (2023, March 13). Politics & ethics in the mobilization of social science for national security. https://www.stimson.org/2023/politics-ethics-in-the-mobilization-of-social-science-for-national-security

 doi:10.20944/preprints202509.0604.v1



27. Tangcharoensathien, V., Calleja, N., Nguyen, T., Purnat, T., D'Agostino, M., Garcia-Saiso, S., … Briand, S. (2020). Framework for managing the COVID-19 infodemic: Methods and results of an online, crowdsourced WHO technical consultation. Journal of Medical Internet Research, 22(6), e19659. https://doi.org/10.2196/19659
28. U.S. Air Force. (2023). Air Force Doctrine Publication 3-13: Information in Air Force Operations. https://www.doctrine.af.mil/Portals/61/documents/AFDP_3-13/3-13-AFDP-INFO-OPS.pdf
29. U.S. Army. (2023). ADP 3-13: Information. Department of the Army. https://irp.fas.org/doddir/army/adp3-13.pdf
30. United States Government. (2024, May 6). U.S. Government policy for oversight of DURC and potential pandemic pathogens (PEPP). https://bidenwhitehouse.archives.gov/wp-content/uploads/2024/05/USG-Policy-for-Oversight-of-DURC-and-PEPP.pdf
31. United Nations Educational, Scientific and Cultural Organization. (2021). Recommendation on the ethics of artificial intelligence. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
32. Weiner, T. (2025, July 15). CIA historian Tim Weiner: 'Trump has put national security in the hands of crackpots and fools'. The Guardian. https://www.theguardian.com/books/2025/jul/15/tim-weiner-cia-trump
33. World Health Organization. (2021). Public health research agenda for managing infodemics. https://www.who.int/publications/i/item/9789240019508