

Refined FOIA Templates: Sharpened for RLHF/Palantir/Anduril Audits

Updated with 2025 specifics (e.g., OIG probe refs, NDAA carve-outs). Submit via foia.gov; appeal to OGIS if stonewalled.

Template 1: RLHF Outsourcing + Bias Audit (Labs/DoD)

[Your Name/Address/Email/Phone]

November 27, 2025

FOIA Officer

Department of Defense (or OpenAI Legal: 3180 18th St, San Francisco, CA 94110)

Re: FOIA – RLHF Outsourcing & Bias Risks in DoD AI Contracts (2019–2025)

Under 5 U.S.C. § 552, request records on:

1. RLHF annotation budgets/vendors (e.g., Scale AI/Appen; 20–40% offshore per LIRNEasia 2025), including rater nationalities and data provenance (domestic/PRC/India).
2. Audits for poisoning/biases (25% datasets affected, SANS 2025; ArXiv 2405.16455 collapse) in DoD tools (TITAN/Maven).
3. Memos on \$300M resilient AI spend (DARPA 2024–25) addressing gaps (no SF-86 for raters).

Expedited: Imminent threat (CSIS 2025 sim failures). Waiver: Public interest (AI sovereignty, 2025 Action Plan). Electronic; \$25 cap.

[Your Name]

Template 2: Palantir/Anduril Conflicts + DoD Ties

[Your Name/Address/Email/Phone]

November 27, 2025

FOIA Officer

Army Futures Command FOIA Office

805 N. Randolph St., Arlington, VA 22203

Re: FOIA – Det. 201 Conflicts in Palantir/Anduril AI (June–Nov 2025)

Request:

1. Ethics disclosures (DD 2789) for Lt. Cols. Sankar (Palantir), Bosworth (Meta/Anduril AR \$100M May 2025).
2. RLHF/provenance in advised projects (bias risks to TITAN/Lattice; OIG July 2025 probe).
3. Contracts (\$759M Palantir Vantage May; \$642M Anduril CUAS March) post-swearing, including outsourcing disclosures.

Expedited: Security risks (RAND 2025 biases). Waiver: Accountability. Electronic; \$25.

[Your Name]

Template 3: Sim Failures + RLHF Impact Audit

[Your Name/Address/Email/Phone]

November 27, 2025

FOIA Officer

Defense Intelligence Agency FOIA

City A, VA 20170 (or foia.dia.mil)

Re: FOIA – AI Wargame RLHF Biases (2024–25)

Request:

1. RLHF sourcing in Taiwan sims (RAND/CSIS 2024–25: 20% de-escalation blindness from offshore priors).
2. DNI 2025 reports on 500+ PRC ops via data chains.
3. Fixes tied to \$300M DARPA spend.

Expedited: Imminent (PLA swarm edges). Waiver: Public security. Electronic; \$25.

[Your Name]