MICHAEL J. MONGAN (SBN 250374)
michael.mongan@wilmerhale.com
WILMER CUTLER PICKERING
HALE AND DORR LLP
50 California Street, Suite 3600
San Francisco, CA 94111
Telephone: (628) 235-1000

EMILY BARNET (*pro hac vice*)
emily.barnet@wilmerhale.com
WILMER CUTLER PICKERING
HALE AND DORR LLP
7 World Trade Center
250 Greenwich St
New York, NY 10007
Telephone: (212) 230-8800

Attorneys for Plaintiff Anthropic PBC

KELLY P. DUNBAR (*pro hac vice*)
kelly.dunbar@wilmerhale.com
JOSHUA A. GELTZER (*pro hac vice*)
joshua.geltzer@wilmerhale.com
KEVIN M. LAMB (*pro hac vice*)
kevin.lamb@wilmerhale.com
SUSAN HENNESSEY (*pro hac vice*)
susan.hennessey@wilmerhale.com
LAUREN MOXLEY BEATTY (SBN 308333)
lauren.beatty@wilmerhale.com
LAURA E. POWELL (*pro hac vice*)
laura.powell@wilmerhale.com
SONIKA R. DATA (*pro hac vice*)
sonika.data@wilmerhale.com
WILMER CUTLER PICKERING
HALE AND DORR LLP
2100 Pennsylvania Avenue NW
Washington, DC 20037
Telephone: (202) 663-6000

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

ANTHROPIC PBC,

Plaintiff,

v.

U.S. DEPARTMENT OF WAR et al,

Defendants.

Case No. 3:26-cv-1996

SECOND DECLARATION OF THIYAGU RAMASAMY IN SUPPORT OF PLAINTIFF ANTHROPIC PBC'S REPLY BRIEF IN SUPPORT OF PLAINTIFF'S MOTION FOR PRELIMINARY INJUNCTION

Judge: Hon. Rita F. Lin

I, Thiyagu Ramasamy, pursuant to 28 U.S.C. § 1746, declare as follows:

1. I previously submitted a declaration in this case on March 9, 2026, describing Anthropic PBC's ("Anthropic's") relationship with the U.S. Government, including the Department of War ("DoW"). I submit this supplemental declaration in support of Anthropic's Reply Brief in Support of its Motion for a Preliminary Injunction.

Background and Purpose of Supplemental Testimony

2. As stated in my March 9 declaration, my understanding of Anthropic's public sector partnerships is based on my experience as Anthropic's Head of Public Sector. As I explained there, I previously worked at Amazon Web Services ("AWS") as a Principal Lead for Data, Analytics, and Artificial Intelligence/Machine Learning. In that role, I was responsible for, among other things, implementing Anthropic's AI models, Claude, for AWS's public sector customers, including AWS's deployment of Claude in classified government networks. Together, my prior experience and current role give me a detailed understanding of Anthropic's relationship with the U.S. government, the technical capabilities of Anthropic's AI models, how third-party cloud providers deploy Claude to government customers, and how those customers access and integrate Claude in their workflows.

3. I have personal knowledge of the contents of this declaration, or have knowledge of the matters based on my review of information and records gathered by Anthropic personnel, and could testify thereto.

4. My original declaration explained that I had never seen or heard any explanation for why Anthropic would be considered a supply-chain risk. I understand that the DoW has since offered multiple rationales in an attempt to support its determination that Anthropic poses a supply chain risk to national security. Those rationales are set forth in (1) the March 17, 2026 declaration of Under Secretary of War Emil Michael and (2) an undated "Memorandum for the Record" documenting his analysis of the purported "Urgent Supply Chain Risk" arising from Anthropic's refusal to agree to DoW's insistence that Anthropic permit the use of its models for "all lawful purposes."

The Government's Factual Errors And Misunderstanding Of Anthropic's Technology

5. Neither I nor anyone else at Anthropic had seen Under Secretary Michael's memorandum or declaration until the Department of Justice publicly filed them in this matter. Those materials reflect a fundamental misunderstanding of how Anthropic's tools are deployed in classified systems and otherwise made available to the government, and they contain multiple factual misstatements. Had the DoW disclosed its rationales before designating Anthropic a supply chain risk, I and other Anthropic subject-matter experts could have corrected these errors.

6. Most fundamentally, Under Secretary Michael's declaration and memorandum suggest that Anthropic has, over time, proven to be an increasingly untrustworthy partner to DoW. In my experience, however, the opposite is true. Anthropic has consistently worked to deepen its relationship with DoW and to support DoW's specific needs and use cases.

The Government's Adoption of Anthropic's Models in Classified Environments

7. Beginning in early 2024, Anthropic worked with a third-party cloud provider to support deployment of Anthropic's commercial AI models in classified cloud environments. By May 2024, Claude Sonnet was available on a provider's Top Secret cloud and was being used by Intelligence Community ("IC") and DoW customers. As government users adopted the commercial version of Claude in classified settings, Anthropic received feedback that its commercial models, appropriately trained to decline discussion of classified materials with the general public, sometimes applied those same protections to IC and DoW users in secure environments. In response, Anthropic engaged directly with government stakeholders, including through classified discussions led by Anthropic engineers holding security clearances, to understand operational needs and identify appropriate technical solutions.

8. In December 2024, the updated commercial version of Claude (Sonnet 3.5) was launched in classified cloud environments. To be clear, in these arrangements, once Anthropic gives its models to the third-party cloud provider, Anthropic has no access to, or control over, the model as deployed or used by government customers.

9. In early 2025, Anthropic—on its own initiative and at its own expense—assigned cleared personnel to develop a version of Claude tailored to national-security needs. In March

2025, that tailored national-security model, called Claude Gov, was made available to IC and DoW customers through classified cloud environments. Government users reported that Claude Gov performed significantly better than the commercial version and did not refuse tasks that were otherwise restricted but appropriate to the national security mission, such as reviewing classified documents or supporting military planning. Adoption increased rapidly across the IC, and government customers expressed strong appreciation for Anthropic's proactive approach to national-security partnerships.

10. In June 2025 and then again in November 2025, Anthropic deployed new, improved versions of the Claude Gov models. Before each deployment, DoW customers conducted extensive testing to confirm that the new model performed as expected before the prior version was retired and replaced. Anthropic closely collaborated with its government and third-party partners to ensure a seamless transition throughout this process.

11. Within weeks of the November 2025 release, nearly all IC and DoW customers had adopted the updated version of Claude Gov. Government users praised this model as a major technical advance, citing significant improvements in reasoning, agentic tool use, and coding. These improvements enabled customers to fully leverage air-gapped deployment of Claude Code, Anthropic's agentic coding tool, which is now widely used across the IC and DoW.

12. Anthropic's sustained investment and close collaboration with its government partners have made Claude Gov a mission-critical capability across classified systems. Anthropic has remained a trusted partner to DoW and the IC at every step, consistently delivering mission-oriented and safety-focused models and supporting AI adoption.

Anthropic Lacks the Technical Ability to Interfere With DoW Operations

13. A central theme of the DoW's belated rationale is the concern that Anthropic could interfere with DoW operations. As I explained in my original declaration, Anthropic has never sought, and has never asserted, any role in DoW's operational decision-making. I emphasize here that, apart from the fact that our company has no desire to interfere with DoW activities, Anthropic is not technically capable of exercising the type of "operational veto" the DoW suggests.

14. Once Claude is deployed in support of DoW's mission, Anthropic has never had the ability to cause Claude to stop working, alter its functionality, shut off access, or otherwise influence or imperil military operations. Even if it *wanted* to do so—which it does not—Anthropic could not do so as a technical matter. Claude has always been deployed for DoW by a third party, and Anthropic does not have the access required to disable the technology or alter the model's behavior before or during ongoing operations.

15. More specifically, as Claude is deployed in DoW environments—such as through air-gapped, classified cloud systems operated by third-party defense contractors—Anthropic has no ability to access, alter, or shut down the deployed model. Anthropic does not maintain any back door or remote "kill switch" in Claude. Anthropic personnel cannot, for example, log into a DoW system to modify or disable the models during an operation; the technology simply does not function that way. In these deployments, only the Government and its authorized cloud provider have access to the running system. Anthropic's role is limited to providing the model itself and delivering updates only if and when requested or approved by the customer. Anthropic does not have ongoing connectivity to, or control over, model instances running in DoW secure networks. Model weights are also not automatically modified or updated based on DoW's usage.

16. For these reasons, Anthropic also cannot exfiltrate DoW data or conduct surveillance of DoW activities. Anthropic does not have access to DoW's Claude prompts; because we lack any access to this customer data, there is nothing that we could exfiltrate or inspect. Any suggestion that Anthropic could engage in "data exfiltration" of DoW information is unfounded.

17. Under Secretary Michael's declaration and memorandum also suggest that Anthropic could alter its AI models after deployment without DoW's knowledge or consent, including by changing model guardrails or model weights. Anthropic is not capable of doing so. By "guardrails," I mean the safety and security controls designed to ensure a model behaves as intended. "Model weights" are the internal numerical parameters that shape how the model behaves by determining how strongly different inputs influence a model's outputs. Once a model is running inside a government-secure enclave, Anthropic cannot unilaterally alter these features.

To change guardrails or model weights, Anthropic would have to provide an entirely new version of the model, which would require DoW's approval and affirmative action to install.

18. Changing model weights or guardrails is analogous to ordering a cake at a restaurant. Once the cake is served, the diner cannot adjust the recipe at the table, and the chef cannot reach into the dining room–secretly or otherwise–to change it either. Even if the cake is missing only a teaspoon of baking soda, it must be made again from scratch with the correct ingredients baked in. Here, Anthropic is the chef. If changes to model behavior are desired, whether small or large, Anthropic can prepare a new version of the model with those changes built in. But that new version is not automatically substituted. Before deployment, the third-party cloud provider and DoW security reviewers inspect and approve it to ensure it is safe and appropriate. Only after that approval, and only after the customer confirms the new version performs as expected, can it retire the prior model.

19. I also understand the government has assessed risk based on the belief that AI models may "drift" or degrade over time, leaving the DoW reliant on Anthropic to ensure continued accuracy and fairness. That concern reflects another fundamental misunderstanding of how AI models operate. Once a model like Claude is trained and deployed by DoW, it is static; it does not change or degrade on its own. The model's parameters remain fixed unless and until a newer version is deployed. Anthropic does not, and cannot, push unsupported or undisclosed updates to a model once deployed by DoW.

20. Finally, Under Secretary Michael's memorandum and declaration cite the "relatively opaque nature of large language model technology" as a baseline risk requiring heightened trust in Anthropic. It is true that large language models ("LLMs") are probabilistic, not deterministic, as Anthropic's Chief Science Officer has explained, and there is some legitimacy to DoW's concern about the opacity of these systems generally. But that characteristic applies to all LLMs. Compared to other AI labs, Anthropic is uniquely transparent. We maintain a publicly available set of normative principles— "Claude's Constitution" —a human-readable document that explains the values and rules our AI is trained to follow. Anthropic is also at the forefront of interpretability

research aimed at making the behavior of models like Claude easier to understand. If anything, Anthropic presents a lower baseline risk than other AI labs, not a higher one.

The Government's Other Alleged Concerns are Misplaced

21. I understand that the government references Anthropic deployment issues at the Center for Disease Control and Prevention ("CDC") as a factor in its designation. While I cannot be certain which incident Under Secretary Michael is referencing, to the extent it relates to CDC work of which I am aware, that characterization reflects a misunderstanding. My understanding is that the issue arose from the CDC's use of a general commercial model whose standard safeguards—appropriately designed for public users—limited certain interactions related to biomedical research to mitigate public safety risks. Once Anthropic became aware of the issue, we worked promptly with our third-party partners and the government to enable use of the model in a manner appropriate to the CDC's mission and expertise. As noted above, safeguards suitable for commercial deployments may not be appropriate for government uses, and some government deployments require customization or parameter adjustments. That collaborative, iterative process—balancing safety in commercial settings with valid government uses—is a core tenet of Anthropic's AI safety approach. That is precisely how we addressed the CDC issue once identified, supporting important CDC research while preventing individual users from engineering dangerous biological weapons.

22. The Michael Declaration also suggests Anthropic's employment of foreign nationals increases "adversarial risk." Anthropic, like many cutting-edge AI companies, employs a globally diverse team of highly skilled researchers and engineers. Anthropic has undergone extensive U.S. Government security vetting in connection with its classified work and takes seriously its obligation to comply with all security and technical requirements for safeguarding sensitive government information. To my knowledge, Anthropic is the only AI company where cleared personnel have led the development of AI models tailored to be deployed in classified environments. As discussed above, for the past two years, Anthropic has been the government's closest and most-trusted partner in driving the adoption of AI models in classified environments.

I declare under penalty of perjury that the foregoing is true and correct.

Executed on March 20, 2026.

/s/ *Thiyagu Ramasamy*
Thiyagu Ramasamy

ATTESTATION PURSUANT TO CIVIL LOCAL RULE 5-1(I)(3)

Pursuant to Civil Local Rule 5-1(i)(3), I attest that concurrence in the filing of this document has been obtained from the other signatories.

By: /s/ *Michael J. Mongan*
Michael J. Mongan