Cover note to the Governor’s AI expert group

AI Watchdog Alliance℠ · A National Initiative · standards@aiwatchdogalliance.org · 970-236-8228

October 8, 2026

Staff lead, Office of the Governor, AI expert group, via the Government Operations Agency
cc: California Department of Technology; California Governor’s Office of Emergency Services; Governor’s Washington, D.C. office

Re: A working definition of “verified” for AI kill switches, and a no-cost pilot

The Governor’s September announcement asked for AI kill switches verified on an ongoing basis by independent organizations. The hard part of that sentence is “verified.” This note offers a definition the expert group can adopt as a starting point, a California company that can demonstrate it, and nothing to pay.

The definition. A stop is verified when an independent party can confirm, on a schedule, five things: the system’s permission to act comes from a party it cannot reach; withholding that permission stops it whether or not it cooperates; a system that goes silent or escapes its monitoring loses permission rather than running unrecorded; restarting takes a second person; and drills are rehearsed on a schedule and leave records anyone can check later, offline, without the operator’s help. These five are written as outcomes, not designs. No product is named. Any lab can meet them its own way.

Physical machines (Criterion F). Draft 0.9.3 adds a sixth outcome for embodied systems. A robot, vehicle or other actuated machine meets the verified-stop definition when an independent party can bring it to a safe stop from outside the machine, without the cooperation of the machine, its software or its network; when that stop is rehearsed on a schedule and leaves a record anyone can check; and when the machine’s stop-assurance level, including none, is disclosed to the owner before operation. Criterion F names no mechanism. A machine that cannot be stopped from outside is not barred, it is labeled.

Where it is written down. KS-1.0, the Verified Stop Standard for Advanced AI Systems, published by the Alliance as Draft 0.9.3 for public comment through November 15, 2026. It includes a record format, a verifier tool contract, drill cadences by system class, a mapping to the NIST AI Risk Management Framework and SP 800-53, and a conformance test outline. The text is free to use and cite. Enclosed.

Why it suits California. It can be cited by name in guidance without naming a vendor. It is testable, so “ongoing verification” means a verifier runs a published tool on a record, not a firm writes an opinion. And it aligns with the pending federal bill: the same five criteria have been provided to the offices of the bill’s sponsor and co-sponsor, so a single definition means a company that complies in California complies nationally and no preemption fight follows. The White House Accord of September 29 asks signers for independent external assessment of their controls; this definition is what such an assessment would check.

The offer. A no-cost pilot of the drill-and-record component (criteria D and E, built and testable today) with one state agency or laboratory, verified by a University of California group using the published verifier, producing a short signed letter the expert group can rely on. The lease-and-stop components (A through C) are specified and will be demonstrable within sixty days; we will say plainly which is which at every step.

Who we are. The AI Watchdog Alliance is a nonprofit consumer-protection and standards organization, incorporating in Delaware and applying for 501(c)(3) recognition, governed by a board with an independent majority. Its founder is also a founder of the California company whose patent-pending work informed the standard; that relationship, the patent disclosure, and a written royalty-free commitment for all government, nonprofit and individual use are published with the standard.

Ask. Thirty minutes with the expert group’s staff to walk through the criteria and the pilot, at the group’s convenience, and permission to brief the Department of Technology on procurement language in parallel.

Enclosed: KS-1.0 Draft 0.9.3.

John M. Ossenmacher
Director, AI Watchdog Alliance℠
949-238-0120 · standards@aiwatchdogalliance.org