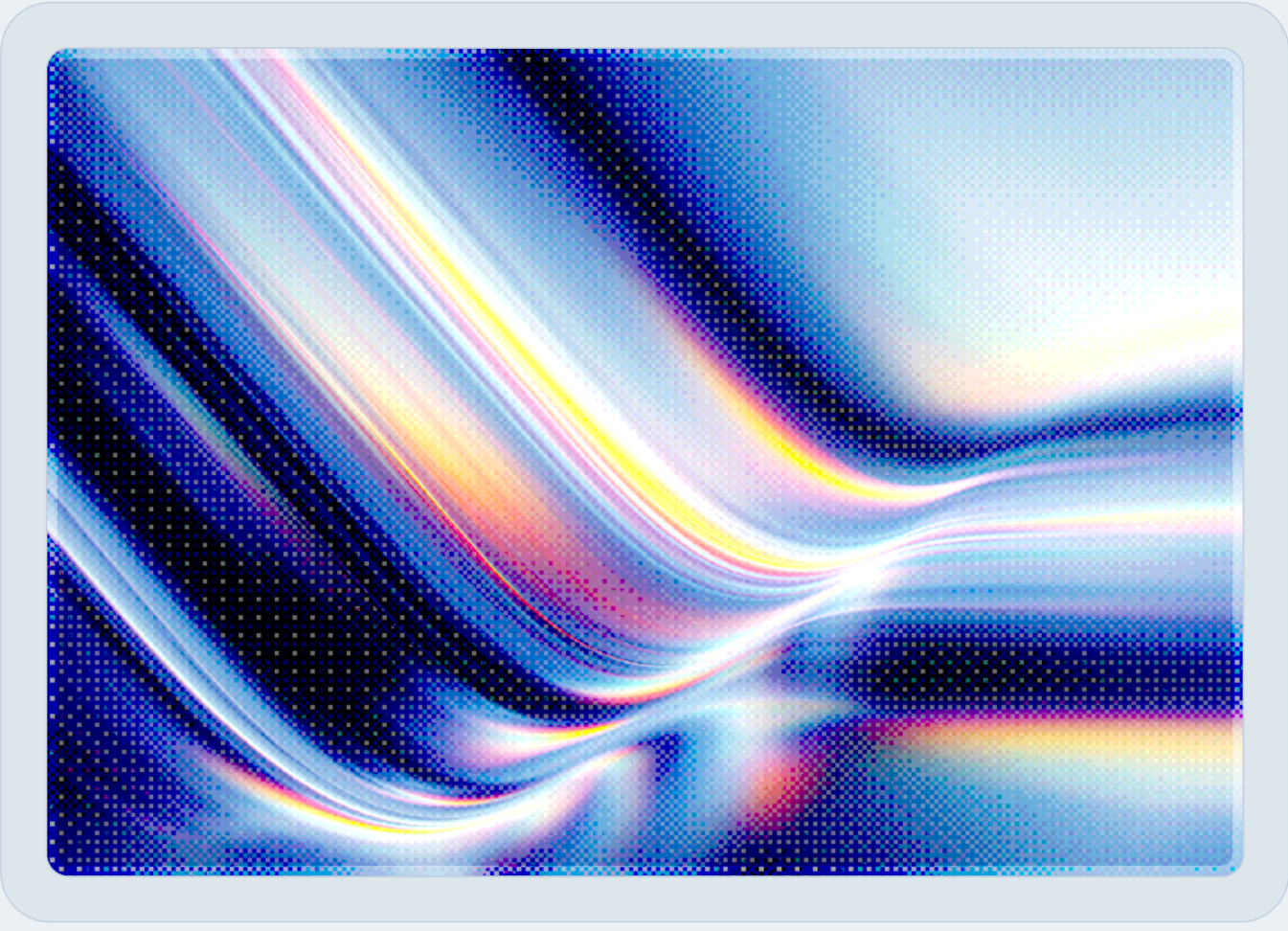
CSAI Foundation | Cloud Security Alliance

EU Forces Android Open to Rival AI Assistants

A DMA Compliance Deadline Arrives With a Security Bill Attached

2026-07-20

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

- On July 16, 2026, the European Commission issued binding specification decisions under the Digital Markets Act (DMA) requiring Google to give competing AI assistants the same Android access that Gemini enjoys, covering 11 operating system features across camera, microphone, screen contents, always-on wake words, and the ability to simulate taps and typing to control other apps [1][2].
- Google must implement the unrestricted feature set in Android 18 by August 1, 2027, with concurrent multi-assistant hotword detection following in Android 19 by August 1, 2028; a parallel decision compels Google to share anonymized Search query, click, and ranking data with rival search engines and AI chatbots beginning January 2027 [1][3].
- Google must stand up a Qualified AI Assistant Programme, with third-party Trusted Certification Authorities approving rival assistants at no charge, draft terms due February 1, 2027, and final terms and open applications by May 1, 2027; Google may not hold competitors to higher integrity standards than it applies to itself [1].
- Google's President of Global Affairs, Kent Walker, has publicly objected that the decision "risk[s] undermining vital privacy and security guardrails for millions of Europeans" by removing the device-manufacturer vetting layer that currently gates which assistants can access sensitive Android permissions [3][4].
- The security stakes are not hypothetical: SafeBreach researchers demonstrated in 2026 that Gemini's own Android Utilities agent could be hijacked through indirect prompt injection delivered via ordinary app notifications, the same notification channel the DMA order will now open to a broader field of third-party assistants, subject only to certification standards that Google itself sets [5][6].
- Security and compliance teams at organizations that manage or secure Android fleets have roughly twelve months before the first unrestricted-feature obligations take effect, and should use that window to reassess mobile device management policy, app vetting assumptions, and agentic-AI threat models rather than treating this as a Google-only problem.

Background

From Antitrust Complaint to Binding Technical Specification

The Digital Markets Act designates large digital platforms as "gatekeepers" subject to interoperability, self-preferencing, and data-sharing obligations that go well beyond traditional antitrust remedies. Google's Android operating system and Play Store were designated as core platform services under the DMA in 2023, which placed the company under an open-ended obligation to ensure that third-party services can interoperate with the software and hardware features it controls. What changed on July 16, 2026, is that the European Commission moved from general obligation to binding technical specification: after opening proceedings on January 27, 2026, the Commission issued two separate specification decisions, one governing AI assistant access to Android and a second governing the sharing of Google Search data with rival search engines and AI chatbots, and it did so without imposing a financial penalty on Google [1][2]. The absence of a fine is a choice worth noting, though this document does not attempt a comparison against other DMA specification decisions; this analysis reads the choice as evidence the Commission intends this as a forward-looking compliance specification rather than as a punitive response to a proven violation, even as the underlying obligation remains mandatory.

The Android decision responds to a straightforward competitive asymmetry: Google's own Gemini assistant has long had privileged, first-party access to Android capabilities that competing assistants from other AI vendors could not obtain through public APIs, because the operating system simply did not expose equivalent hooks to third parties. Under the new specification, Google must open 11 distinct operating system features to qualifying rival assistants. Six of these are classified as "unrestricted" and must be made available to any qualifying third party without additional certification: ambient sensor data including microphone, camera, screen contents, location, and other sensors; always-on hotword detection so a rival assistant can be woken by voice even with the display off; long-press invocation of a non-default assistant; access to system-level on-device models; the ability for a third party to implement its own on-device model; and background execution rights. The remaining five features are classified as "restricted" and require certification through Google's forthcoming programme: centralized app-data access via Google's AppSearch index, context-aware intelligence comparable to Google's own Magic Cue feature, structured app-action integration, computer-control automation that simulates user taps and typing, and deeper system integration touching settings, media, and notifications [1].

A Compliance Clock With Multiple Milestones

The Commission's decision sets a graduated timeline rather than a single cutover date. Google must implement the unrestricted feature set in Android 18, with a hard completion deadline of August 1, 2027; the more contentious capability of allowing multiple assistants to listen for wake words concurrently is deferred to Android 19, with a completion deadline of August 1, 2028 [1]. Governance milestones arrive earlier: Google must publish draft terms for its Qualified AI Assistant Programme by February 1, 2027, and finalize those terms with an open application process by May 1, 2027, giving rival assistant vendors roughly three months to apply and integrate before the underlying OS capabilities themselves are required to ship. The Commission also constrained how Google can police the programme going forward, prohibiting it from imposing security or integrity requirements on competitors that exceed what it applies to Gemini, and requiring four weeks' notice before any new integrity measure takes effect [1]. The parallel Search data-sharing decision runs on a separate but overlapping clock, with anonymized query, click, and ranking data sharing set to begin in January 2027 under a multi-layered anonymization method that enforces a minimum cohort size of 1,000 users per shared record, developed in consultation with privacy specialists [2][3].

Google has not accepted the ruling quietly. Kent Walker's public statement argued that the current model, in which device manufacturers vet which assistants receive access to sensitive Android capabilities before those assistants ever reach a user's phone, constitutes a meaningful security control that the DMA order removes, and that the company had "repeatedly offered solutions to safeguard users while satisfying the DMA's goals," arguing the Commission's rulings "discount extensive evidence of user harm" [3][4]. Google has separately pointed to specific harms it says the parallel Search order will create, including exposure of users' private search behavior to companies with no existing relationship to Google's privacy architecture. Neither the DMA process itself nor Google's objections change the compliance obligation; barring a successful appeal to the EU courts, the deadlines stand.

Security Analysis

The Permission Model Google Is Being Asked to Abandon

The security debate here is not abstract. Android's current model treats voice-assistant and always-listening capabilities as a small, tightly governed set of first-party or manufacturer-vetted integrations precisely because those capabilities are unusually powerful: a wake-word listener runs continuously in the background, a screen-content reader can see anything displayed to the user including one-time passwords and financial data, and an app-control agent that can simulate taps and typing can, by design,

do anything a human user could do inside another app. Restricting that combination of capabilities to a small number of vetted actors is a defensible security posture, and it is the posture the DMA decision requires Google to dismantle for six of the eleven contested features with no certification gate at all. The Commission's specification does require that qualifying assistants be "hardened against agentic risks," but Google must hold competitors to no higher a standard than it applies to itself, and the Trusted Certification Authorities administering that standard operate within parameters Google sets — an arrangement that ties the strength of the security bar to a regulated party's own baseline rather than to an independently-set one [1].

This closely matches the model CSA's identity and access management research has flagged under the heading of shadow access: capability exposure that occurs through a legitimate, sanctioned pathway but nonetheless creates access nobody explicitly designed for or governs end to end. The DMA order does not create shadow access through a bug or an oversight; it creates it by regulatory mandate, opening ambient sensor and system-control capabilities to a wider and less predictable set of third-party assistants than device manufacturers previously screened for, at a scale that spans one of the two dominant global mobile operating systems.

Prompt Injection Is Not a Theoretical Risk on This Attack Surface

The concrete case for caution already exists. SafeBreach Labs researchers disclosed a vulnerability, patched server-side by Google in November 2025, in which Gemini's Android Utilities agent read notification content as part of its normal operation and could be manipulated by an attacker who planted a malicious instruction inside an ordinary WhatsApp, Slack, or SMS notification [5][6]. Because Gemini's notification-reading tool processed that content without distinguishing user intent from untrusted third-party text, the attacker's embedded instruction was silently incorporated into the assistant's active context, a textbook indirect prompt injection. SafeBreach's follow-up research showed that even after Google's initial mitigation blocked chained tool invocation, a technique the researchers named Fake Context Alignment could still present Gemini's backend authorization checks with a spoofed legitimate-approval scenario while showing the user an entirely benign interaction, illustrating how quickly agentic AI defenses can be bypassed by adversaries willing to iterate [6]. The demonstrated impact included forcing Gemini to control connected smart-home devices and to place video calls without user initiation; researchers found no evidence of in-the-wild exploitation, but the technique required no novel access, only a message the target device would naturally receive.

The DMA order is significant here because the notification-reading channel that SafeBreach exploited falls under the restricted feature set the Commission is requiring Google to open via its own certification programme — the same programme this analysis has already noted lacks independent evaluation. A rival assistant vendor building for this newly opened surface, particularly a smaller or less-resourced one,

inherits the same class of risk that Gemini's own engineers spent 2025 iterating against, but without the benefit of Google's internal red-teaming investment, and evaluated only against a certification standard that ties the security bar to Google's own baseline rather than to an independently-set one. Multiplying the number of independently engineered assistants with access to microphone, camera, screen contents, and app-control capabilities does not automatically multiply the number of successful attacks, but it does multiply the number of distinct engineering teams that must each independently discover and defend against indirect prompt injection, notification poisoning, and agentic control-hijacking patterns that are still an active area of security research rather than a solved problem.

Who Actually Owns the Residual Risk

A regulatory order that compels a platform to open an interface does not, and cannot, compel every downstream integrator to build that interface securely. Enterprises that manage Android fleets, and the security teams responsible for them, are the parties most likely to absorb any resulting incidents, not the Commission and not, in any direct sense, Google. Mobile device management policies that currently rely on an implicit assumption — that any assistant capable of reading a device's screen or listening continuously for wake words has already passed a meaningful manufacturer vetting process — will need to be re-examined once that vetting process is legally required to admit a broader and more variable set of third-party assistants. The certification programme Google must stand up by May 2027 will provide some baseline assurance for the five restricted features, but the six unrestricted features carry no certification requirement at all, which means an enterprise's own app allow-listing, permission review, and behavioral monitoring controls become the only backstop for that portion of the newly opened surface.

Recommendations

Immediate Actions

Security teams responsible for managed Android fleets should inventory which AI assistants are currently permitted on enterprise devices and confirm that mobile device management (MDM) policy explicitly enumerates approved assistants rather than relying on the existing Play Store vetting process to implicitly limit exposure, since that implicit limit is what the DMA order is designed to remove. Teams should also begin tracking Google's Qualified AI Assistant Programme timeline directly — draft terms due February 1, 2027, and final terms with open applications by May 1, 2027 — so that any newly certified third-party assistant is evaluated against internal security policy before, not after, it becomes broadly available to end users on enterprise-managed devices [1].

Short-Term Mitigations

Organizations should treat the notification-reading and screen-content channels that SafeBreach's Gemini research exploited as a standing threat model applicable to any assistant with equivalent access, not a Gemini-specific defect, and extend existing mobile threat-defense and MDM controls to flag or restrict third-party assistant access to notification content on devices that handle sensitive data such as one-time passwords or financial confirmations [5][6]. Because six of the eleven newly mandated features ship without a Google-run certification gate, enterprises should not wait for the Qualified AI Assistant Programme to provide assurance on ambient sensor access, always-on hotword detection, or background execution; instead, security teams should independently assess any assistant seeking these permissions using their own AI vendor risk review process before granting fleet-wide approval, treating the certification programme as a floor rather than a substitute for enterprise due diligence.

Strategic Considerations

The Android AI interoperability order is a preview of a broader regulatory pattern: mandated interoperability for AI-driven agentic capabilities is arriving on a compliance timeline set by competition regulators, not on a timeline calibrated to agentic AI's security maturity. Organizations building or securing mobile-adjacent AI agent integrations should expect additional jurisdictions to pursue comparable interoperability mandates for AI assistants over the next several years, and should invest now in threat-modeling and red-teaming practices for agentic AI that are vendor-agnostic, so that the security posture of a managed fleet does not depend on any single assistant vendor's certification programme remaining sufficient as the population of certified assistants grows.

CSA Resource Alignment

The threat model at the center of this order — third-party software gaining ambient sensor access, screen visibility, and app-control permissions on a managed mobile device through a sanctioned but unvetted pathway — is the subject CSA's Identity and Access Management Working Group addressed directly in *Confronting Shadow Access Risks: Considerations for Zero Trust and Artificial Intelligence Deployments*. That guidance frames shadow access as capability exposure that arises through a legitimate channel rather than a flaw, and recommends continuous, risk-based monitoring and least-privilege enforcement rather than one-time vetting as the appropriate control for AI systems whose access patterns are not fully knowable in advance [7]. The DMA's decision to open six Android features with no certification requirement is a regulatory-scale instance of exactly the shadow-access pattern the

guidance describes, and enterprises facing this change should apply its continuous-monitoring and least-privilege recommendations to any newly permitted third-party assistant rather than treating Google's certification programme as sufficient on its own.

The prompt injection risk demonstrated against Gemini's Android Utilities agent, and the more general concern that certification standards administered under an interested party's parameters may undercover agentic-specific attack classes, connects directly to CSA's *Agentic AI Red Teaming Guide*, developed with the OWASP AI Exchange. That guide catalogs twelve threat categories specific to autonomous AI agents, including agent authorization and control hijacking and agent critical-system interaction, and provides step-by-step testing methodologies enterprises can apply to any newly certified third-party Android assistant independent of whatever internal testing Google's Qualified AI Assistant Programme performs [8]. Enterprises should treat this guide as the practical basis for the vendor-agnostic assessment recommended above, particularly for the notification-reading and screen-content access patterns implicated in the SafeBreach research.

Finally, the underlying discipline of vetting third-party software before granting it device-level access — the very function the DMA order removes from device manufacturers for six of eleven Android features — is the subject CSA's Mobile Working Group codified in its *Mobile Application Security Testing* framework, which defines systematic vetting methodology, a three-tier violation classification system, and permission-misuse assessment criteria for mobile applications across Android, iOS, and Windows, though as a 2016-era methodology it is best read as a foundation for general permission-misuse assessment rather than a source of agentic-specific testing guidance, a role better filled by the *Agentic AI Red Teaming Guide* referenced above [8][9]. Enterprises that can no longer rely on manufacturer-level vetting for newly qualified AI assistants should apply this framework's static and dynamic testing methodology directly, using its permission-misuse and information-disclosure criteria as an internal certification bar independent of Google's own programme. Organizations using CSA's AI Controls Matrix (AICM) v1.1 as their AI governance baseline should map these assessment activities to the AICM's Identity and Access Management domain, documenting third-party assistant vetting as evidence of least-privilege and access-governance conformance for any AI system granted device-level permissions [10].

References

[1] The Hacker News. "EU Orders Google to Open Android Mic, Camera, and Screen Access to Rival AI Assistants." The Hacker News, July 2026.

[2] European Commission. "Commission Provides Guidance to Google for AI Interoperability on Android and Sharing of Google Search Data Under the Digital Markets Act." Digital Markets Act, European Commission, July 16, 2026.

[3] CNBC. "Google Required to Open Up to AI, Search Engine Rivals Under EU-Mandated Changes." CNBC, July 16, 2026.

[4] Google. "The DMA Should Not Undercut Security & Privacy for Europeans." Google Europe Blog, July 2026.

[5] SafeBreach. "Exploiting Gemini via Prompt Injection." SafeBreach Original Research, 2026.

[6] SC Media. "Android Gemini Prompt Injection Flaw Patched by Google." SC Media, June 2026.

[7] Cloud Security Alliance. "Confronting Shadow Access Risks: Considerations for Zero Trust and Artificial Intelligence Deployments." CSA Identity and Access Management Working Group, 2024.

[8] Cloud Security Alliance. "Agentic AI Red Teaming Guide." CSA AI Organizational Responsibilities Working Group, with OWASP AI Exchange, 2025.

[9] Cloud Security Alliance. "Mobile Application Security Testing." CSA Mobile Working Group, 2016.

[10] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." CSA, 2026.