The shared AI training control in robots.txt: what Google plans and how IETF AIPREF fits

22 November 2026 · 6 min read · blogs.iictp.org/robots-txt-ai-training-control-ietf-aipref

Short answer: on 30 October 2026 Google announced plans to support a shared AI training control in robots.txt, based on work by the AIPREF working group at the Internet Engineering Task Force (IETF). A site would set one policy token to express an AI-training preference across participating providers, and Google says it would read an opt-out as if Google-Extended had been disallowed. Google aims to support it with other AI providers in 2027, so it is a plan, not a live feature.

TL;DR The short version

- Google plans a shared robots.txt control for AI training, based on the IETF AIPREF working group, so one token can express a preference to participating providers.
- Google says it would treat an opt-out as if Google-Extended were disallowed, and aims to support it in 2027, so it is a plan, not a live feature.
- Prepare by listing the AI crawlers your robots.txt addresses and deciding per content type whether to allow training; visibility in AI answers is a separate lever.

Read online, listen to it, or change the text size: https://blogs.iictp.org/robots-txt-ai-training-control-ietf-aipref

What is the shared AI training control?

Today a publisher who wants to opt out of AI training has to name each provider’s crawler or token separately. The proposal adds a shared “policy token” in robots.txt that states a broad preference, so a site can opt out of training across all participating providers in one step or choose provider by provider.

How it works, as described by Google

- If a robots.txt file uses the shared token to opt out of training **and** Google-Extended is not explicitly specified, Google plans to treat it as if the publisher had disallowed Google-Extended.
- Other AI providers can handle it the same way.
- Existing company-specific controls keep working alongside it, so a site can set a broad choice and still tailor instructions for one developer.
- Google says sites that specifically allow Google-Extended are eligible to appear in responses and receive clicks via grounding in Gemini Apps.

What is new compared with earlier approaches

Google-Extended has been available since 2023 as Google's own AI-training control. The new element is a cross-industry vocabulary term for "AI training", taken from the IETF AIPREF working group, while discussion of the rest of the AIPREF vocabulary continues. Google says the draft reflects feedback from publishers, civil society and regulators, and is consistent with the EU Code of Practice for general-purpose AI model providers, which recognises robots.txt as a primary way for publishers to express machine-readable preferences.

Limitations and open questions

- Not yet supported: Google gives 2027 as its aim and says it will post updates.
- It is a draft proposal. The exact token syntax was not stated in the post, so check the IETF working group documents before you write rules.
- A robots.txt preference is a signal that compliant crawlers honour. It does not technically prevent a crawler that ignores it.
- The post does not say how this interacts with search crawling; training controls and Search indexing are separate decisions.

How to prepare

1. List the AI crawlers and tokens your robots.txt already addresses, including Google-Extended.
2. Decide, per content type, whether you want to allow or refuse AI training. Marketing pages and licensed content often differ.
3. Keep Search crawling rules separate from AI training rules, and test both.
4. Watch the IETF AIPREF working group and Google's Search Central blog for the final syntax before changing production files.

What it means for marketers and site owners

Visibility in AI answers and permission for training are different levers. This proposal is about the second. For how AI systems treat your content, see personal agents and analytics: site owner controls.

Frequently asked questions

- **Is the shared control live?** No. Google says it aims to support it, with other AI providers, in 2027.
- **Does it replace Google-Extended?** No. Google says existing company-specific controls continue to work alongside it.
- **Who defines it?** The IETF AIPREF working group, as described in Google's post.

Related reading

- Personal agents, analytics and site owner controls
- Responsible AI in practice
- What is prompt injection?

Sources and currency

Sources were read on 22 November 2026. Products, rules and standards change, so check the linked originals before you rely on any detail. This announcement is from 30 October 2026, older than the two-week window, and is included because it is still only a plan.

Related credential: AI Practitioner. Explore the exam facts and syllabus.

Sources and further reading

1. John Mueller (Google). Supporting a shared AI training control for websites. 30 October 2026. developers.google.com/search/blog/2026/10/ai-training-control.

Published by IICTP, an independent technical credential body, on 22 November 2026. The current version of this post is at https://blogs.iictp.org/robots-txt-ai-training-control-ietf-aipref. More guides: blogs.iictp.org.