Google Updates Gemini/Vertex AI User Agent Documentation via @sejournal, @martinibuster

Google updated the documentation for the Google-Extended user agent, which publishers can use to control whether Google Gemini and Vertex use their data for training purposes or for grounding AI answers.

Updated Guidance

Google updated their guidance on Google-Extended based on publisher feedback for the purpose of improving clarity and adding more specific details.

Previous Documentation:

> *"Google-Extended is a standalone product token that web publishers can use to manage whether their sites help improve Gemini Apps and Vertex AI generative APIs, including future generations of models that power those products. Grounding with Google Search on Vertex AI does not use web pages for grounding that have disallowed Google-Extended."*

Updated Version

The updated documentation provides more detail and is easier to understand explanation of what the user agent is for and what blocking it accomplishes.

> *"Google-Extended is a standalone product token that web publishers can use to manage whether content Google crawls from their sites may be used for training future generations of Gemini models that power Gemini Apps and Vertex AI API for Gemini and for grounding (providing content from the Google Search index to the model at prompt time to improve factuality and relevancy) in Gemini Apps and Grounding with Google Search on Vertex AI."*

Google-Extended Is Not A Ranking Signal

Google also updated one sentence to make it clear that Google-Extended isn't used as a ranking signal for Google Search. That means that allowing Google-Extended to use the data for grounding Gemini AI answers won't be counted as a ranking signal.

Grounding is a reference to using web data (and knowledge base data) to improve answers provided by a large language model with up to date and factual information, helping to avoid fabrications (also known as hallucinations).

The previous version omitted mention of ranking signals:

> *"Google-Extended does not impact a site's inclusion or ranking in Google Search."*

The newly updated version specifically mentions Google-Extended in the context of a ranking signas:

> *"Google-Extended does not impact a site's inclusion in Google Search nor is it used as a ranking signal in Google Search."*

Documentation Matches Other Guidance

The updated documentation matches a short passage about Google-Extended that's elsewhere in Google Search Central. The other longstanding guidance explains that Google-Extended is not a way to control how website information is shown in Google Search, demonstrating how Google-Extended is separated from Google Search.

Here's the other guidance that's found on a page about preventing content from appearing in Google AI Overviews:

> *"Google-Extended is not a method for managing how your content appears in Google Search. Instead, use other methods to manage your content in Search, such as robots.txt or other robot controls."*

See also: Google Confirms Robots.txt Can't Prevent Unauthorized Access

Takeaways

- **Google-Extended Documentation Update:**
  The Google-Extended documentation was clarified and expanded to make its purpose and effects easier to understand.
- **Separation From Ranking Signals:**
  The updated guidance explicitly states that Google-Extended does not affect Google Search inclusion nor is it a ranking signal.
- **Internal Use By AI Models:**
  Clarified that Google-Extended controls whether site content is used for training and grounding Gemini models.
- **Consistency Across Documentation:**
  The updated language now matches longstanding guidance elsewhere in Google's documentation, reinforcing its separation from search visibility controls.

Google updated its Google-Extended documentation to explain that publishers can block their content from being used for AI training and grounding without affecting Google Search rankings. The update also matches longstanding guidance that explains Google-Extended has no effect on how sites are indexed or ranked in Search.

Featured Image by Shutterstock/JHVEPhoto

https://www.searchenginejournal.com/google-updates-gemini-vertex-ai-user-agent-documentation/545409/