

Qwen3-Reranker-4B arrives on Regolo.ai: the new multilingual Reranking model for smarter, more precise search

The powerful AI model for text ranking is now available on Regolo.ai, delivering higher precision in over 100 languages.

MILANO, MILANO, ITALY, November 11, 2025 /EINPresswire.com/ -- Regolo.ai announces the introduction of Qwen3-Reranker-4B among the models now available on its European inference platform.

This is an advanced multilingual Reranking model, designed to enhance the quality of search and retrieval systems through a deeper understanding between queries and documents.

With support for over 100 languages, long-context handling, and an 8-billion-parameter architecture, Qwen3-Reranker-4B enables developers and enterprises to achieve more relevant, accurate, and context-aware search results - empowering any RAG (Retrieval-Augmented Generation) pipeline, chatbot, or internal search engine.

What is a Reranking model

Reranking is a technology that reorders search results by assessing the semantic relevance of each document with respect to the user's query.

Qwen3-Reranker-4B acts as a "second critical brain" within the retrieval process: it takes the initial results from a retriever, re-evaluates them with a cross-encoder, and returns only the most useful and coherent content aligned with the query's intent.

"We are excited to make Qwen3-Reranker-4B available on our platform," said the Regolo.ai team. "This model represents a major step forward for anyone seeking smarter, more accurate searches -reducing information noise and improving response quality in AI-based systems."

Multilingual, accessible, and easy to integrate

The new model is now available through Regolo.ai's OpenAI-compatible APIs, at a predictable

cost of €0.01 per query.

It can be quickly integrated into leading industry frameworks - including LangChain, LlamaIndex, Flowise, n8n, and Open WebUI - adding advanced reranking capabilities to any AI application.

Thanks to its multilingual support (100+ languages) and ability to handle long-form text, Qwen3-Reranker-4B is ideal for:

- Multilingual FAQs and corporate knowledge bases;
- Internal search engines;
- RAG applications with large document sets;
- E-discovery and intelligent information management systems.

With the addition of Qwen3-Reranker-4B, Regolo.ai further expands its model offerings, confirming its role as a reference platform for building multilingual, high-performance, and high-quality AI applications.

Learn more and try Qwen3-Reranker-4B at www.regolo.ai

What is Regolo.ai

Regolo.ai is the European platform developed within DHH (Dominion Hosting Holding) for building private, sustainable, and high-performance AI applications.

Designed for those seeking a reliable and secure AI inference solution, Regolo.ai is an “LLMs as a Service” platform that combines innovation, performance, and privacy protection.

With Regolo.ai, users can access and run AI models without worrying about computational power, while enjoying complete peace of mind regarding data protection, thanks to its zero data retention policy.

Chiara Grande
Seeweb Srl
+39 0775 880041

[email us here](#)

Visit us on social media:

[LinkedIn](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/866298708>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable

in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.