

What to Consider When Choosing a Web Categorization API Vendor

Posted on October 27, 2023

A fundamental requirement of website categorization is accurately classifying websites into relevant categories. But that can be challenging, as the task typically requires targeted web expertise and a multitude of system resources. Therefore, to improve their web categorizing capabilities, large companies and solution providers often need the help of a reliable third-party API.

In this post, we listed the essential considerations when evaluating potential website categorization API providers.

6 Features to Consider When Choosing a Website Categorization API

Website Coverage

Coverage is an essential quality indicator when evaluating a website categorization API vendor. Comprehensive coverage means that a provider has the technology to categorize even the most recently launched or unpopular websites. An API that isn't expansive enough won't be able to serve as a good data source.

Number of Categories Supported

The more relevant categories an API supports, the higher the value it provides to users. On the

one hand, an API with many categories is more likely to categorize websites accurately since it has more specific categories to choose from. However, having too many categories displayed may also confuse users and delay business and security decisions.

Accuracy

Accuracy is another crucial indicator that separates effective website categorization technology from the rest. While validating an API's accuracy with manual verification is a good practice, it is also time-consuming. Companies should choose an API that assigns confidence level scores to each category, giving users the freedom to look into websites further. The higher the score, the more accurate the category assignment is.

Update Frequency

Web filtering and categorization APIs must be updated regularly to reflect the fast-paced nature of the Internet. Websites can change their content and possibly their category anytime. An outdated web categorization tool may not be able to identify changes to existing websites and update their categories accurately. Therefore, looking for an API that does live or near-real-time queries makes sense to ensure the results are up-to-date and relevant.

Performance and Speed

Performance and speed are also crucial to a [website categorization API](#). Users should be able to run categorization queries on multiple sites simultaneously and receive results quickly. A fast and responsive web categorization tool will provide a better user experience and be more effective at preventing users from engaging with undesirable websites and enforcing Internet usage policies.

Database Download Availability

Another important feature is the availability of a downloadable website categorization database as

an alternative data consumption model, allowing users to directly integrate data into their own systems and applications. This capability can help security teams customize information to meet their specific needs.

What Our Website Categorization Product Line Offers

WhoisXML API provides a [Website Categorization API](#) that relies on machine learning (ML) and natural language processing (NLP) techniques to retrieve website content and assign categories effectively. The API does live queries, and all of the information it provides is normalized and follows a standard format.

Alternatively, users can also download the entire [web categorization database](#), which covers hundreds of millions of websites with millions added daily, in CSV format. You can download file samples [here](#).

If you're ready to see how you can use our Web Categorization API, [sign up](#) for free or [contact us](#) to request database access.