

# Know What's Really Behind A Website With Website Categorization API!

Posted on April 18, 2019



Whois API, LLC is constantly on a quest to provide accurate & reliable information about domain activities on the web. Our solid foundation, constant innovations & unparalleled databases of Domain, Whois, DNS, IP, OSINT and Threat Intelligence helps professionals from across industries to get the Intel they need. But with over a billion websites, the internet is already quite massive and is constantly growing. With the plethora of websites & their various data-sets, analyzing them can be cumbersome & time-consuming especially in cases of filtering online traffic or finding potentially dangerous websites. In such cases sometimes just knowing ‘what the website is about?’ or ‘the website category’ can help you decide your further course of action.

Website categorization information is becoming a crucial tool for individuals & businesses to know what they are dealing with. To help them precisely identify websites we have now launched [Website Categorization API](#). Our advanced system provides global categorization coverage & performs real-time analysis of each website. We leverage machine learning and artificial intelligence, alongside human-verification techniques to provide the highest quality and accuracy for website categorizations.

Our premier web categorization solution conducts 3 layers of filtering & categorization to provide the most accurate information to our users:

### **1) Website Response**

Determines if the website was active during the crawling. Especially when looking to determine whether or not a domain is malicious, the domain’s status or response can be an important factor.

### **2) Machine Learning and Rules**

This combines the application of machine learning with the versatility of rules defined by experts. Components included:

- **Text Extraction**

Our advanced systems extract important information from a given website and analyzes the content based on natural language processing. This is done with software that simulates human Web surfing to collect specified bits of information from different websites.

- **Keyword Extraction**

Our system crawl through the website's code & extracts the Meta tags present to help categorize the website's content as well.

### **3) Human Supervised Categorization**

By utilizing machine learning categorization we can get the scale necessary to deal with the incredible volumes of new websites being published at an increasing rate, yet we also involve human assistance for authentication in order to maintain the highest levels of accuracy.

The categorization algorithm is able to accurately classify websites into 25 categories, at present. Queries for a domain name will return the main categories associated with that site and its content. Every website may belong to up to 3 categories based on the above categorization technique. For eg. Instagram.com falls under Internet and Telecom, People and Society, Arts and Entertainment.

## **Website Categorization API Highlights**

- 

### **Big Data**

By leveraging advanced machine learning, artificial intelligence and human verification we categorize & update millions of websites in a day, keeping our database as one of the most accurate in the industry.

-

## **Advanced Algorithms**

Our technology categorizes content found on the website, as well as, its code to ensure accurate results.

- 

## **Always Up-to-date**

Our web crawlers are constantly visiting and classifying new and existing websites, providing real-time results and keeping the database updated.

- 

## **Developer-Friendly Responses**

Every API response is returned as JSON, which can be easily read and implemented into your systems/ applications/ tools.

- 

## **Wide Range of Applications**

Website Categorization can be used for lead generation, marketing segmentation, targeted advertising, brand safety, subscriber analytics, web content filtering, online traffic filtering, parental control, fraud prevention, malware detection, administrators who want to make smarter firewall decisions, organizations who want to keep stats about the web content their employees access, detecting & preventing insider threats, complimenting existing classification or even building a new index of the entire web.

For crucial threat detection, you can combine Website Categorization API with our [Domain Reputation API](#) to get a comprehensive picture of a domain & identify sites that contain potential security risks for networks and users.

Also, you can also combine Website Classification API with our [Domain Research Suite](#) so as to get enhanced Whois records or registration details including contact information, registrant details for any/all of the 25 categories.

Website Classification API is an ideal solution for vendors and service providers looking to protect consumers or business users from malicious or inappropriate internet sites, and is of particular use to Managed Security Service Providers (MSSP), Firewall and Unified Threat Management (UTM) providers, Security Operation Centres (SOCs), Value-added and Mobile Security providers. We are committed to providing a one-stop domain solution for our users and with the right database coverage and utilizing advanced technology, our Website Categorization API can be a great asset to your business. So what are you waiting for?

Get real-time content analysis and categorization of a website: <https://website-categorization-api.whoisxmlapi.com>