Symantec RuleSpace™ Data Sheet

OEM URL Categorization Database and Real-time Web Categorization Technology

Data Sheet: Security Intelligence

Symantec RuleSpace™

OVERVIEW

A major challenge today is ensuring a safe web environment for users and companies without impacting their web experience. Symantec RuleSpace offers an advanced web intelligence solution to enable vendors to address these key concerns. The following are key use cases for URL categorization.

Parental Controls	Web Analytics
• Keep children safe online • Protect against web based malware attacks (Phishing, Adware, Spyware) • Empower parents and educational institutions to supervise web browsing • Achieve compliance to regulations	• Gain valuable insights into consumer web browsing trends by ISPs and Mobile Operators. • Keep my brand safe
Web Threat Protection	**Acceptable Usage Policy (AUP) Enforcement**
• Enable Clean Pipe Solution at ISP and Mobile Operators • Protect against web based malware attacks like Phishing, Adware/Spyware, etc • Detect Botnet and C&C traffic	• Enforce acceptable browsing policies at Enterprises, Wifi Hot Spots, and traffic enforcement/choke points • Report against web browsing activity

FOUR CATEGORIZATION DATABASES

Symantec offers 4 categorization databases:

- Web Categorization Database
- Malware Categorization Database
- Phishing Categorization Database
- Internet Watch Foundation (IWF) Child Abuse Image Content Database

Symantec RuleSpace incorporates proven and globally respected URL Web categorization databases with real-time analysis from one of the pioneers in the industry with 18+ years of expertise delivering highly effective categorization solutions to multiple markets. RuleSpace offers a unique combination of database lookups and real-time analysis for accurate categorization of unknown or dynamic web sites, especially useful in dealing with sites containing user generated content like Facebook, Twitter and YouTube. Our business model is designed around providing superior service quality with a dedicated engineering and technical support team. Symantec's RuleSpace is powering dozens of 3rd party web filtering solutions serving over 350 million end-users worldwide, from subscribers at ISPs/mobile operators, to businesses and individual consumers.

KEY FEATURES	MARKET SEGMENTS	PROVEN TECHNOLOGY
• Unrivaled web, Malware and Phishing Categorization Databases • Sensitive categories manually curated • Real-time categorization engines • Carrier grade performance • Built for virtual environments • Flexible architecture • Category override support • Enhanced with Symantec DeepSight • Advanced machine learning	• Powers URL or Web content solutions • Internet security products or services • Web gateway and unified threat management (UTM) appliances • Parental Control / Child protection solutions • Contextual advertising services • Data analysis and mining products • Brand protection solutions	• 18 years in the market • 63+ million sites categorized • Serving over 350 million users • Malware feed from Symantec analysis of billions of events from the Global Intelligence Network • Symantec Phishing feed

Symantec

Web Categorization Database (WCD)

The WCD is one of the most accurate and relevant collections of general purpose Website category information available on the market. It has 100+ categories ranging from "parental control" categories of inappropriate sites for younger Web users (as well as a special Kids category containing sites specifically designed for children) to important businesses categories like streaming media, social networking or Webmail sites.

There are also categories specifically for sites formatted for mobile devices, sites used for serving advertising, and sites used as anonymous proxies, often used to try and circumvent Web filtering solutions.

Our main Web Categorization Database has been under constant development for 18 years and contains millions of unique Web sites in dozens of languages, broken down into 100+ categories:

CATEGORIES

- 46 parental control categories
- 40 enterprise categories
- 12 wireless categories
- 6 Malware categories
- 1 Phishing category

CATEGORY EXAMPLES

- ALCOHOL
- BLOG
- ENTERTAINMENT
- GAMBLING
- HATE
- PORNOGRAPHY
- SELF HARM
- TRAVEL

Human analysts maintain the most sensitive categories such as Violence and Suicide

Closely monitoring changes in Global Legislation and reflecting those changes as appropriate.

Symantec uses a combination of web crawling technologies, data from domain registrars, human content analysts and actual web browsing data from our large client base to help ensure new Web sites are discovered and categorized quickly. On average, 10+ million new sites are added to the databases each year and nearly as many annulled sites are purged.

The majority of newly discovered web sites are categorized by automated machine-learning categorization engines in the Symantec RuleSpace Categorization Labs. This technology uses multiple techniques including contextual pattern recognition and analysis of page links to accurately assign languages and categories. Human analysts supplement the automated technology on very complex or specialized sites, to review sites submitted for review by customers, and to help ensure the continued accuracy of the automated categorization technology.

The RuleSpace databases are updated constantly and incremental updates are published daily. These updates are downloaded automatically to our OEM partners and seamlessly applied in the background to all local or remote copies of the databases.

Malware Categorization Database (MCD)

The MCD, powered by Symantec DeepSight, contains hundreds of thousands of malicious URLs and malware locations on the Internet. These sites are organized into 6 different categories. The database is updated frequently and when integrated into a Web filtering solution to provide an effective protection against ever changing Internet-based malware threats.

DeepSight™ INTELLIGENCE

Symantec.

Symantec Malware data feeds are derived from deep, proprietary analysis of billions of events from the Symantec™ Global Intelligence Network. The Global Intelligence Network provides global visibility into the threat landscape, including:

- More than 41.5 million attack sensors in 157 countries
- An extensive anti-fraud community of enterprises, security vendors and more than 50 million end users
- More than 8 billion emails scanned per day - 1/3 of all enterprise email
- Monitoring of over 5 million decoy accounts
- More than 13 billion web requests handled per day

By combining this visibility with our automated analysis and human research efforts, the MCD provides unique, timely and actionable protection from Internet-based threats.

Phishing Categorization Database (PCD)

Powered by Symantec's unrivaled insight into Phishing URLs

The PCD is a comprehensive database of confirmed phishing URLs on the web. Phishing attacks use both social engineering and technical subterfuge to steal private information from users. Clever social-engineering schemes use 'spoofed' e-mails from well-known banks and other companies to lead users to counterfeit Websites designed to trick them into divulging secret financial and personal information like credit card numbers, pins, usernames and passwords.

Integrated into Web filtering solutions, the PCD provides an effective solution to prevent the threat to users from fast moving phishing and pharming attacks.

To provide an effective anti-phishing service, it is critical that the core data remains current in near real-time. Updates to the PCD can be downloaded and applied as frequently as every 10 minutes.

Internet Watch Foundation's Child Abuse Image Content Database (IWF)

The IWF database is the IWF feed of illegal child abuse image Web sites. It is highly recommended and often a regulatory requirement that these sites be blocked for all users by Internet service providers or other web filtering solutions.

RULESPACE WEB CATEGORIZATION REAL-TIME ENGINES

With the broad adoption of Web2.0 technologies, anyone can easily post un-moderated content directly to the Internet via Twitter, blogging or wiki sites. Likewise, social networking sites like Facebook and others have become hugely successful with hundreds of millions of active users constantly posting new content.

1. For pages with dynamic content (e.g. search, social networking or to blogging sites). Real-time Web page analysis is the best way to ensure that these ever changing pages do not contain inappropriate content at the time they are being accessed.
2. For new or "unknown" sites not yet in a categorization database, real-time web page analysis is the best way to ensure that these pages do not contain in appropriate content at the time they are being accessed.

Symantec's experts in real-time content analysis and categorization have over 10 years in delivering extremely high levels of performance and accuracy. Our real-time analysis engines can identify content in many languages in the 15 most common "inappropriate" categories (e.g. Anonymous Proxies, Gambling, Drugs, Pornography, Weapons, etc.).

OEM SOFTWARE DEVELOPMENT KIT

All of the RuleSpace URL category database lookup functionality and the real-time Web content analysis functions are accessed via Symantec's RuleSpace Web Categorization Service SDK. It is available on a wide range of platforms including Windows, OSX, Linux and Solaris; and is designed with flexibility, stability and performance in mind.

Categorization Filtering Interface (CFI)

CFI Platforms Supported:

- Linux (32/64Bit)
- Solaris (32/64Bit)
- Windows
- OSX
- Java

CFI is the single lightweight programming interface used to query all of the licensed URL category databases and real-time categorization engines. It can also be used to create and use custom categorization lists. The requesting application passes URLs or Web content to CFI which then, based on the configuration and the content being passed, determines which categorization services to invoke (database lookup or real-time analysis) and the category responses are returned to the requesting application. CFI can be configured to access the databases hosted locally or remotely on one or more CSRVs (see below) or can be configured to use local "in memory" databases for situations where lookup performance is paramount.

As part of the lookup process, URLs passed to CFI go through a normalizing and parsing process to ensure the most accurate database category match at either domain, host, directory or even page level. CFI will also extract the 'real' target URLs embedded in other longer URLs from common sites like Google Translate or Google Cache where the target URL is embedded in a larger host URL.

Categorization Server (CSRV)

CSRV Platforms Supported:

- Linux (32/64Bit)
- Solaris (32/64Bit)

CSRV is an application server that can be used to host the URL category database(s) and the Update and Feedback Service (UFS). CSRV can be run on a single small appliance/server all the way up to deployments on high-end hardware in large and distributed data centers with high levels of scalability, load balancing and redundancy. Whether running in the same physical location as CFI or remotely hosted in a data center or "cloud" service on the internet, CSRV is multi-threaded and will provide extremely fast category responses to URL lookups via CFI clients.

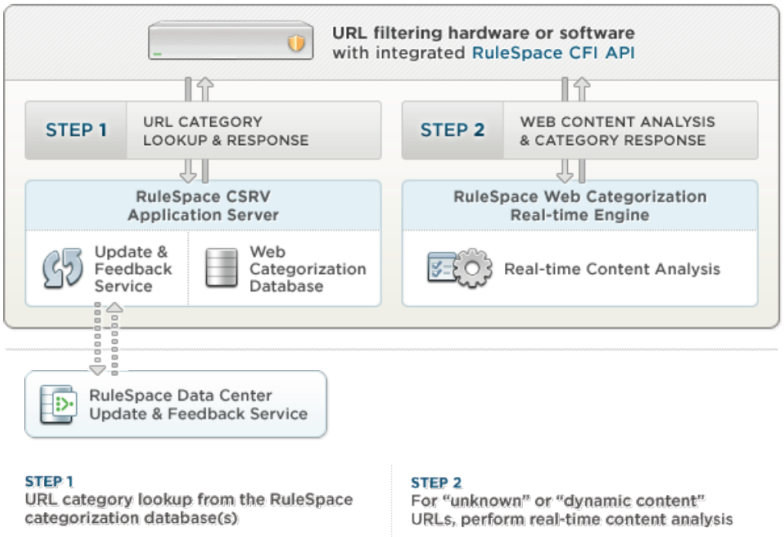
Update and Feedback Service (UFS)

UFS is the communication mechanism between local URL category databases deployed at our customers and the global Symantec update and feedback servers. The UFS fetches and applies the baseline and subsequent delta URL category database updates as a throttled background process. It is also used to stage and transmit "unknown site" feedback records back to Symantec for analysis and categorization. In an environment with multiple CSRVs, the UFS can be configured to make only one download and make the update available to other local CSRVs.

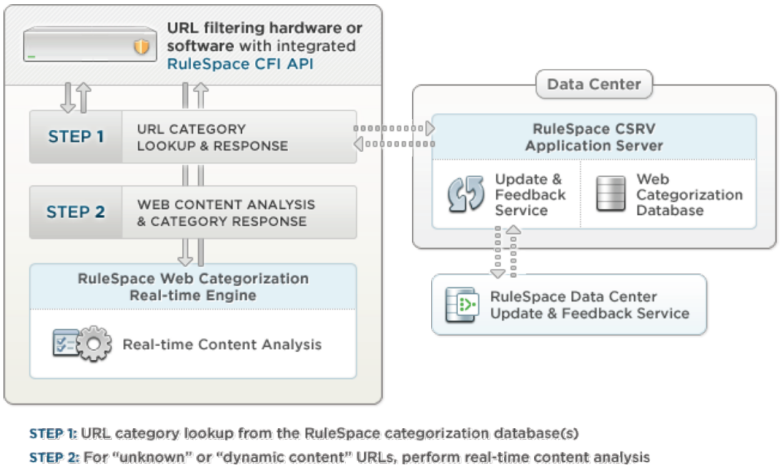
ARCHITECTURE FLEXIBILITY

The solution is engineered to be very flexible and provide for numerous architecture choices. The most common architectures range from a monolithic integration into a single client application to various client/server architectures with varying levels of database hosting, from a local database to a global cloud-based SaaS lookup solution. Our technical engineering support team will consult with you to determine the most appropriate architecture and configuration to best meet your unique requirements.

Powering solutions serving thousands of end users to solutions serving tens of millions of end users.



Example 1: Typical Web Security Appliance Architecture



Example 2: Typical Hosted Filtering Solution Architecture

More Information

Visit our website

go.symantec.com/rulespace

To contact a Product Specialist

email: Ask_RuleSpace@symantec.com

To speak with a Product Specialist outside the U.S.

For specific country offices and contact numbers, please visit our website.

About Symantec

Symantec Corporation (NASDAQ: SYMC) is a security company that helps people, businesses and governments seeking the freedom to safely unlock the opportunities technology brings – anytime, anywhere. Founded in April 1982, Symantec, a Fortune 500 company, operating the world's largest civil global data-intelligence network, has provided leading security solutions for identity, threat protection, information protection and security analytics. The company's more than 19,000 employees reside in more than 50 countries. Ninety-nine percent of Fortune 500 companies are Symantec customers. In fiscal 2015, it recorded revenues of $6.5 billion. To learn more go to www.symantec.com or connect with Symantec at: go.symantec.com/socialmedia.

Symantec World Headquarters

350 Ellis St., Mountain View, CA 94043 USA
+1 (650) 527 8000
1 (800) 721 3934
www.symantec.com

21356887-01 09/15