



NEW WEBSENSE DATA SECURITY POLICY TEMPLATE:

CYBER BULLYING AND SELF-DESTRUCTIVE (SUICIDAL) THOUGHTS



TRITON®

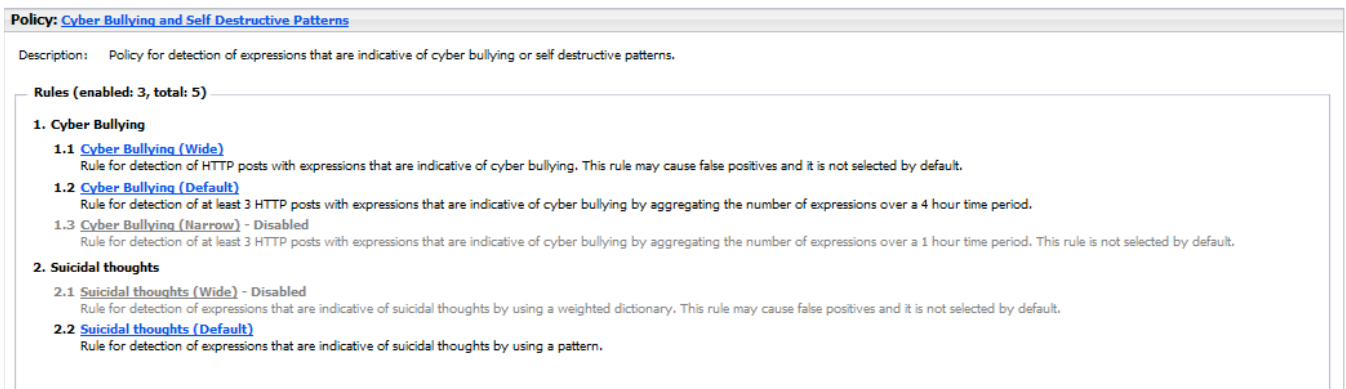
New Websense Data Security Policy Template: Cyber Bullying and Self-Destructive (Suicidal) Thoughts

This document provides a brief introduction to the new *Cyber Bullying and Self-Destructive (Suicidal) Thoughts* rules now available within the Websense® Data Identification & Classification Engine (DICE). We will explain how to implement these rules via a Websense Web Security Gateway Anywhere deployment.*

The latest Web Security Gateway Anywhere license includes data leakage prevention (DLP) features that enable the proxy solution to detect the data patterns contained within the *Cyber Bullying and Self-Destructive (Suicidal) Thoughts* rules, among many other data patterns. Turning on DLP analysis is as simple as ticking a single radio button and entering a password to link the proxy to the DLP policies. You can then create policies as you see fit.

Rules and dictionaries are ready to go “out of the box,” once again enabled with only a few mouse clicks. The policies can work in numerous ways depending on what you’re looking for and how strict you want to be. The pattern dictionaries comprise a combination of keywords, regular expressions and context awareness. In terms of policy enforcement, they provide a higher degree of accuracy than a standalone keyword search.

You can also choose from three versions of each dictionary: narrow, wide and default. These provide stricter or more relaxed versions of the rule and what then determines incident creation.



Policy: Cyber Bullying and Self Destructive Patterns

Description: Policy for detection of expressions that are indicative of cyber bullying or self destructive patterns.

Rules (enabled: 3, total: 5)

- 1. Cyber Bullying**
 - 1.1 Cyber Bullying (Wide)**
Rule for detection of HTTP posts with expressions that are indicative of cyber bullying. This rule may cause false positives and it is not selected by default.
 - 1.2 Cyber Bullying (Default)**
Rule for detection of at least 3 HTTP posts with expressions that are indicative of cyber bullying by aggregating the number of expressions over a 4 hour time period.
 - 1.3 Cyber Bullying (Narrow) - Disabled**
Rule for detection of at least 3 HTTP posts with expressions that are indicative of cyber bullying by aggregating the number of expressions over a 1 hour time period. This rule is not selected by default.
- 2. Suicidal thoughts**
 - 2.1 Suicidal thoughts (Wide) - Disabled**
Rule for detection of expressions that are indicative of suicidal thoughts by using a weighted dictionary. This rule may cause false positives and it is not selected by default.
 - 2.2 Suicidal thoughts (Default)**
Rule for detection of expressions that are indicative of suicidal thoughts by using a pattern.

The Incidents can be created either each time contravening content is detected, or over time. For example, you can create an incident if you see more than three posts from one person in a 24-hour period. This second option is especially useful for, say, differentiating between a campaign of abuse and a careless one-off comment.

You may also have multiple versions of the same rule, each triggered under different circumstances. This allows an administrator to monitor the results in different queues and have background data for reference. You can restrict administrator access according to any number of criteria.

Once an incident is created, there are a number of subsequent actions you can choose from, including:

- » Audit
- » Block
- » Audit and Notify (anyone you choose can be notified, including multiple people)
- » Any other action you want to create

You can create thresholds within the policy to provide better and more manageable thresholds to a rule. For example, when milder content is detected, the action can simply be to audit the content. If harsher content is detected, then the action can be to block the content and send notifications.

General Condition Severity & Action Source Destination

Specify whether to create an incident each time the rule is matched, or to accumulate matches into a single incident. ⓘ

☐ Create an incident for every matched condition

☒ Accumulate matches before creating an incident ⓘ

Count **unique matches** over **4 hours** ⓘ

When there are at least: **1** matches, change severity to: **Low** and the action plan to: **Audit Only** ⓘ ⓘ

Advanced ⓘ

If some matches are detected in the specified time period but the threshold is not met, continue counting for **4 hours**

☒ When there are at least: **3** matches, change severity to: **Medium** and the action plan to: **Audit and Notify** ⓘ ⓘ

☒ When there are at least: **6** matches, change severity to: **High** and the action plan to: **Block All** ⓘ ⓘ

You can set policies to target only certain users or groups, and set multiple versions of similar policies so that different users get a milder or stricter version as appropriate.

Once incidents are created they are visible within reports. Users are identified in the same way as they are with standard web usage. You can see all data regarding exactly what has happened, including destination, time, what Websense technology has done with it and which rule was triggered.

When looking into the forensics, you can also see the specific content that set this rule off – this actual typed-in content is highlighted in the example below.

Incidents (last 30 days)

Workflow Remediate Escalate

Report: Incidents (last 30 days)

Date Range: Last 30 Days

Showing 89945 incident(s)

ID	Incident Time	Source	Policy	Channel	Destination	Severity	Action	Maximum Matches	Transaction Size	Status
1910346	25 Apr, 2014, 01:29:52 PM	Chris Jones	UK PHI; UK DPA; F...	HTTPS	www.facebook.com	Medium	Blocked	3	2.47 KB	New
1904480	25 Apr, 2014, 11:26:35 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	949 B	New
1910005	25 Apr, 2014, 11:26:00 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	1.05 KB	New
1904471	25 Apr, 2014, 11:25:35 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	1.08 KB	New
1909996	25 Apr, 2014, 11:25:35 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	1.08 KB	New
1909987	25 Apr, 2014, 11:21:47 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	958 B	New
1904462	25 Apr, 2014, 11:21:40 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	958 B	New
1910258	25 Apr, 2014, 11:05:22 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	958 B	New
1910155	25 Apr, 2014, 11:05:16 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	958 B	New
1906881	25 Apr, 2014, 10:07:56 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	974 B	New
1906843	25 Apr, 2014, 10:07:53 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	Medium	Blocked	N/A	974 B	New
1910252	25 Apr, 2014, 09:52:50 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	High	Blocked	N/A	974 B	New
1908307	25 Apr, 2014, 09:52:43 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	High	Blocked	N/A	974 B	New
1910246	25 Apr, 2014, 09:52:39 AM	Chris Jones	Cyber Bullying an...	HTTPS	www.facebook.com	High	Blocked	N/A	974 B	New
1910114	25 Apr, 2014, 09:44:37 AM	Chris Jones	Cyber Bullying an...	HTTPS	twitter.com	High	Blocked	N/A	974 B	New

Incidents: 1910114 Severity: High Action: Blocked Channel: HTTPS (encrypted)

Display: Violated rules

Rule: Suicidal thoughts (Default)

Forensics Properties History

Source: Chris Jones

Destination: twitter.com

URL Category: Social Web - Twitter

URI: https://twitter.com/i/tweet/create

Message Body

Field	Value
authenticity_token	1b6201692c56cd0de8d8e8bc829cef76bac6317
place_id	
status	I'm going to hang myself
authenticity_token	1b6201692c56cd0de8d8e8bc829cef76bac6317&place_id=&status=I'm going to hang myself

Workflow Remediate

Report: Incidents (last 30 days)

Showing 89945 incident(s)

ID	Incident Time
1910346	25 Apr
1904480	25 Apr
1910005	25 Apr

We hope this has provided you with a brief overview of our capabilities in this difficult area. Please contact Websense if you would like to learn more.



* Please note that while only the web channel is covered under the Web Security Gateway Anywhere licence, we do have the capability to monitor end-point and other common network communication channels such as email, FTP, ActiveSync and others. This additional coverage would of course be part of a larger licence and may require other integration points, but Web Security Gateway Anywhere provides the base framework necessary to do so.



We hope this has provided you with a brief overview of our capabilities in this difficult area. Please contact Websense if you would like to learn more.

Kari Kritzer, Account Manager
858-847-3376 kkritzer@websense.com

Linnea Sandler, Account Manager
858-320-9218 lsandler@websense.com

Thomas Wood, Account Manager
858-320-9290 twood@websense.com

Cindy Curley, Account Manager
858-361-7743 ccurley@websense.com

TRITON STOPS MORE THREATS.
WE CAN PROVE IT.



TRITON®