What's New

This section includes the following topics:

What's New in Release 5.4.0

Release 5.4.0 provides following enhancements:

- Variable Support

Variable Support makes writing and managing text conditions easier and more flexible. It allows users to define a variable with multiple values and reuse it across text conditions or phrase-based patterns. Instead of manually creating conditions for every possible variation such as different names of the same organization, you can define the values once and reference the variable using . Helpful suggestions appear as you type, making it quick to select and insert variables. During classification, the system automatically expands the variable into all valid combinations, eliminating the need for manual duplication or complex*REGEX* and helping you create cleaner, more consistent policies with less effort.

- Support has been added for pre-2008 South Korea passport numbers that begin with SR. This ensures broader coverage and improved accuracy when processing older passport formats.
- A set of United States state-specific policies now use the North American Telephone Number - Personal pattern instead of the previously used North American Telephone Number pattern. This change impacts 23 policies and improves the accuracy of telephone number detection.

Policies

Following policies are updated:

- Client Concerns - Employee Error Policy
- Client Concerns - General Policy
- Client Concerns - Negative Language Policy
- Client Concerns - Operational Policy

- Client Concerns - Performance and Losses Policy
- Client Concerns - Promises and Guarantees Policy
- Client Concerns - Trade Execution Policy
- Client Concerns Policy
- Gifts and Entertainment Policy
- Market Abuse Policy
- Off Channel Signaling Policy
- Secrecy Policy

Patterns

Following new patterns are added:

- South Korea National ID (Resident Registration) Number (RNN) Beginning Oct 2020
- South Korea National ID (Resident Registration) Number (RNN) Before Oct 2020

Following patterns are updated:

- Client Concerns
- Client Concerns (Sentiment)
- Client Concerns - Employee Error
- Client Concerns - Employee Error (Sentiment)
- Client Concerns - General
- Client Concerns - General (Sentiment)
- Client Concerns - Negative Language (Sentiment)
- Client Concerns - Operational
- Client Concerns - Performance and Losses (sentiment)
- Client Concerns - Promises and Guarantees
- Client Concerns - Trade Execution
- Gifts Donations Sponsorships
- Insider Trading
- Items and Diversions of Value
- Market Manipulation
- Off-Channel Baidu

- Off-Channel Baidu Phrases
- Off-Channel Blink Phrases
- Off-Channel Dust Phrases
- Off-Channel Element Phrases
- Off-Channel Facebook Phrases
- Off-Channel IMO phrases
- Off-Channel Instagram Phrases
- Off-Channel Line Phrases
- Off-Channel LinkedIn Phrases
- Off-Channel Low Frequency Channels
- Off-Channel Matrix Phrases
- Off-Channel Pulse Phrases
- Off-Channel QQ Phrases
- Off-Channel Reddit Phrases
- Off-Channel Signal Phrases
- Off-Channel Silence Phrases
- Off-Channel Slack Phrases
- Off-Channel Snapchat Phrases
- Off-Channel Threads Phrases
- Off-Channel Twitter Phrases
- Off-Channel WeChat
- Off-Channel WeChat Phrases
- Off-Channel Weibo
- Off-Channel Weibo Phrases
- Off-Channel Wire Phrases
- Off-Channel X - Low Confidence
- Secrecy
- Ticket, Tickets, Tix, Tkt, or Tkts

What's New in Release 5.3.0

Release 5.3.0 provides following enhancements:

- Auto-Update for Department-Based Policies

Arctera Insight Classification now supports Auto-Update for policy conditions based on Author and Recipient departments in Arctera Insight Surveillance. If auto update is enabled for a parent department, any newly created child departments under this parent are automatically added to the policy, keeping policy in sync with organizational changes and reducing manual updates.

Key features:

- Automatic Sync: New child departments under a parent with auto update enabled, are added to the policy daily at 1 AM (server time).
- Granular Control: Enable auto update for individual parent department within each policy.
- Less Admin Work: No need to manually track or update policies after department changes.
- Consistent Coverage: Prevents policy gaps during restructuring or growth.
- Enhanced Analyze Feature for Policy Testing

The Analyze feature now supports exporting the Risk & Compliance Analysis summary in a new, more detailed CSV format. This update makes it easier to filter results in Excel, review classification accuracy, and fine-tune policies before applying them to live data.

The following new columns are included in the newer CSV format:

- Highlighted Text \- The exact string that triggered the match.
- Matched Text \- The surrounding text or metadata for better understanding.
- Policy Name \- The specific policy that flagged the content.
- Improved Excel File Extraction Reliability

Arctera Insight Classification continues to use Apache Tika for all file extraction, including text, embedded documents, and OCR from images. With this update, if Tika fails on Microsoft Excel files, a custom parser automatically tries the extraction - provided time remains within the classification timeout. The custom parser extracts text only and does not support embedded documents, OCR, or hyperlink extraction, but ensures better handling of Excel files that previously failed to process.

Policies

Following new policies are added:

- Disclaimers Policy
- United States Minnesota Consumer Data Privacy Act (CDPA / "Minnesota Act") Policy
- United States Maryland Online Data Privacy Act (MODPA) and Maryland Personal Information Protection Act (PIPA) Policy
- United States Nebraska Data Privacy Act (NDPA) (LB 1074) Policy

Following Transparent policies are updated with patterns:

- Client Concerns Policy
- Client Concerns - Communication Policy
- Secrecy Policy
- Off-Channel Signaling Policy

Patterns

Following new patterns are added:

- Disclaimer - Generic
- U.S. Minnesota Postal Address - Form
- U.S. Minnesota Drivers License Number
- U.S. Maryland Postal Address - Form
- U.S. Maryland Drivers License Number
- U.S. Maryland Drivers License Number
- U.S. Nebraska Postal Address - Form
- U.S. Nebraska Drivers License Number

Following new Transparent patterns are added:

- Off-Channel Brax
- Off-Channel Brax Phrases
- Off-Channel Line Phrases
- Off-Channel Pinterest
- Off-Channel Pinterest Phrases
- Off-Channel Wire Phrases

Following Transparent patterns are updated:

- Off-Channel Weibo Phrases, Off-Channel IMO phrases, Off-Channel Phone, Off-Channel Element Phrases
- Off-Channel Snapchat Phrases, Off-Channel Silence Phrases, Secrecy, Off-Channel Substack Phrases
- Off-Channel WhatsApp Phrases, Off-Channel Tinder Phrases, Off-Channel Signal Phrases, Off-Channel Baidu Phrases
- Off-Channel Telegram Phrases, Off-Channel Tango Phrases, Off-Channel Instagram Phrases, Off-Channel Blink Phrases
- Off-Channel Chomp, Off-Channel Grindr Phrases, Off-Channel X - High Confidence, Off-Channel Amino Phrases, Off-Channel Matrix Phrases
- Off-Channel LinkedIn Phrases, Client Concerns (Sentiment), Off-Channel Twitter Phrases, Off-Channel WeChat Phrases
- Off-Channel Chomp Phrases, Off-Channel Kik Phrases, Off-Channel Slack Phrases, Off-Channel Skype Phrases, Off-Channel Low Frequency Channels
- Off-Channel BBM phrases, Off-Channel Dust Phrases, Off-Channel BBM, Off-Channel Pulse Phrases, Client Concerns - Communication (Sentiment)
- Off-Channel Discord Phrases, Off-Channel Facebook Phrases, Off-Channel Reddit Phrases, Off-Channel QQ Phrases, Off-Channel Threads Phrases, Off-Channel X - Low Confidence

What's New in Release 5.2.0

Release 5.2.0 provides following enhancement:

- Introducing New Language Detection Engine: FastText
    - A new language detection engine, FastText, has been introduced to enhance both the speed and accuracy of identifying languages within documents. Users now have the flexibility to choose between the existing Apache Tika engine and FastText based on their requirements. FastText offers improved accuracy for complex or noisy text, processes large volumes of data efficiently, and performs well with mixed-language content. It also supports over 170 languages, covering a wide range of global and regional dialects. Support for additional languages will be added in future updates.
- Audit Diff View

    - The Audit Diff view, previously available only for patterns, has now been extended to include policies and tags. It has also been enhanced to highlight all types of changes, helping the users to understand and interpret modifications by referencing the color-coded badges displayed in the interface.

- Preset Tags
    - Preset tags used to identify specific failure types in classification, such as SENTIMENT NOT DETECTED, PROTECTED DOCUMENT, and PARTIALLY CLASSIFIED, are now visible in the UI. This enhancement improves the usability of the classification Test functionality on the Patterns page by providing clearer insights into why certain documents may not have been fully classified. Users can now review and interpret classification outcomes more effectively, which helps in faster issue resolution and improves the accuracy of pattern configuration.

Policies

- Following new policies are added along with the patterns:
    - Māori Te Mana Raraunga Policy
    - United States Florida Senate Bill 264 (SB 264) Policy
    - United States Kentucky Consumer Data Protection Act (CDPA) (HB 15) Policy
    - United States New Hampshire Consumer Data Privacy Act (CDPA) Policy
    - United States New Jersey Data Privacy Law (DPL) Policy

Patterns

- Following new patterns are added:
    - Te Reo Māori
    - Person's Name - Japan
    - Arctera Keys
    - Authentication
    - Email Credentials
    - Generic Database Credentials
    - MySQL Credentials
    - OpenAI Keys
    - PostgreSQL Credentials

- Secret Key

Known Issues

This section includes the following topics:

Known Issues and Limitations

Refer to the following table of known issues and limitations for the current version of Arctera Classification.

ISSUE ID	DESCRIPTION
VIC-8232	When multiple conditions are removed from a single block, the Audit Diff should highlight each removed condition individually. Currently, the change is not properly itemized, making it difficult to track what was removed.
VIC-8227	Resetting a built-in policy does not display the correct Audit Diff . Multiple attributes are missing in the audit diff after the reset.
VIC-6179	Delay is observed after clicking selected department link at the bottom of the Department tree view.
VIC-4905	If enableExactDataMatch is set to False and VIC is restarted, the feature remains disabled and the EDM GetRulepack API returns a 404 Not Found error.

Getting Started

This section includes the following topics:

Introducing Arctera Insight Classification

Arctera Insight Classification lets you classify items based on their content and metadata. The main areas of the Arctera Classification are as follows:

> “ ”

Note: Arctera Insight Classification supports Google Chrome, Mozilla Firefox, Microsoft Edge, and Internet Explorer 11. Please note that the Analyze feature does not work when Internet Explorer 11 is used. **Note:** Arctera Insight Classification does not support classification of encrypted files.

> “ ”

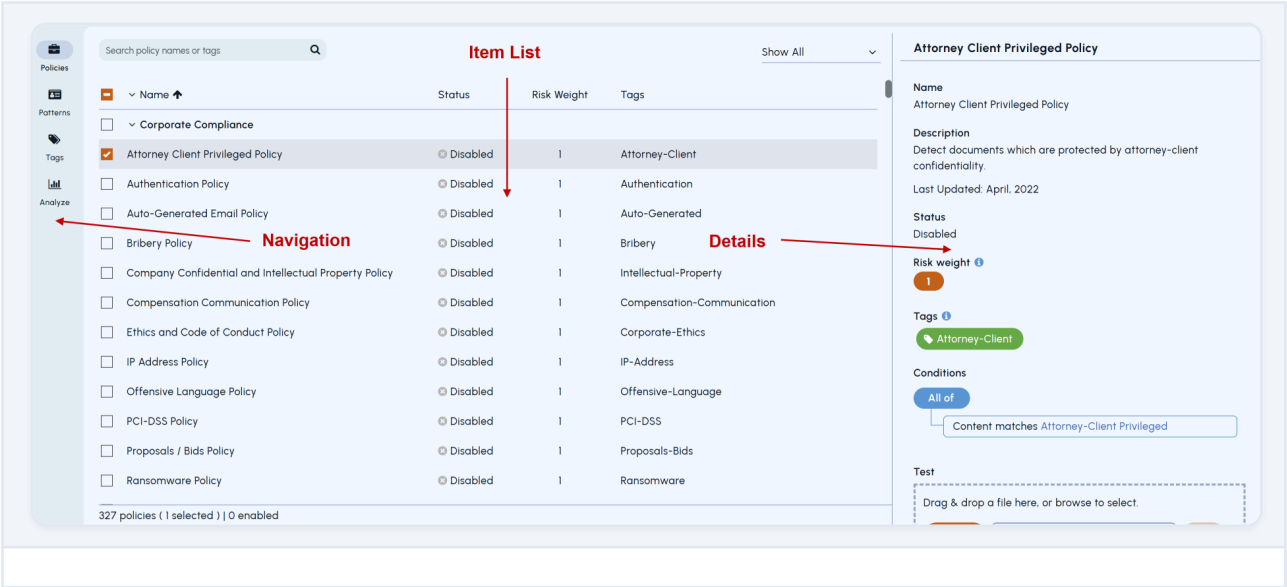
	**POLICIES . THE ARCTERA CLASSIFICATION EVALUATES THE ITEMS THAT YOU SUBMIT FOR**
	CLASSIFICATION AGAINST A SET OF POLICIES. EACH POLICY SPECIFIES THE CONDITIONS THAT AN
	ITEM MUST MEET TO BE ASSIGNED A SPECIFIC CLASSIFICATION TAG. THE NUMEROUS BUILT-IN POLICIES COVER MANY OF THE REGULATIONS AND CORPORATE STANDARDS FOR WHICH YOU MAY WANT TO CLASSIFY ITEMS, AND YOU CAN CREATE CUSTOM POLICIES IF YOU HAVE ADDITIONAL REQUIREMENTS.
	SEE ABOUT POLICIES **.**
	Patterns . Each of the built-in policies checks the items that you submit for classification for one or more patterns. These patterns use sophisticated algorithms to look for matches that meet the required confidence level. You can incorporate the built-in patterns in any custom policies that you create, and also create custom patterns of your own.
	See About patterns .
	Tags . When an
	item that you have submitted for classification meets the conditions of a policy,
	the Arctera Insight Classification assigns the associated tags to the item. You can create custom tags to add to the large number of built-in ones.
	See About tags .
-	Analyze . Use this feature to look for policy matches in a sample folder of your organization's file content. You can evaluate

	POLICIES . THE ARCTERA CLASSIFICATION EVALUATES THE ITEMS THAT YOU SUBMIT FOR
	CLASSIFICATION AGAINST A SET OF POLICIES. EACH POLICY SPECIFIES THE CONDITIONS THAT AN
	ITEM MUST MEET TO BE ASSIGNED A SPECIFIC CLASSIFICATION TAG. THE NUMEROUS BUILT-IN POLICIES COVER MANY OF THE REGULATIONS AND CORPORATE STANDARDS FOR WHICH YOU MAY WANT TO CLASSIFY ITEMS, AND YOU CAN CREATE CUSTOM POLICIES IF YOU HAVE ADDITIONAL REQUIREMENTS.
	SEE ABOUT POLICIES .
	the content against all the available policies, or limit the analysis to individual policies or groups of policies.
	While the analysis proceeds, the Arctera Insight Classification displays real-time statistics on the percentage of files that may contain sensitive data and their estimated risk level. Then, when the analysis has finished, you can download a report of its findings in comma-separated value (CSV) formats.
	See About analyze .

For details on policies and patterns, refer to the *Arctera Insight Classification Policies and Patterns Guide*.

Finding your way around

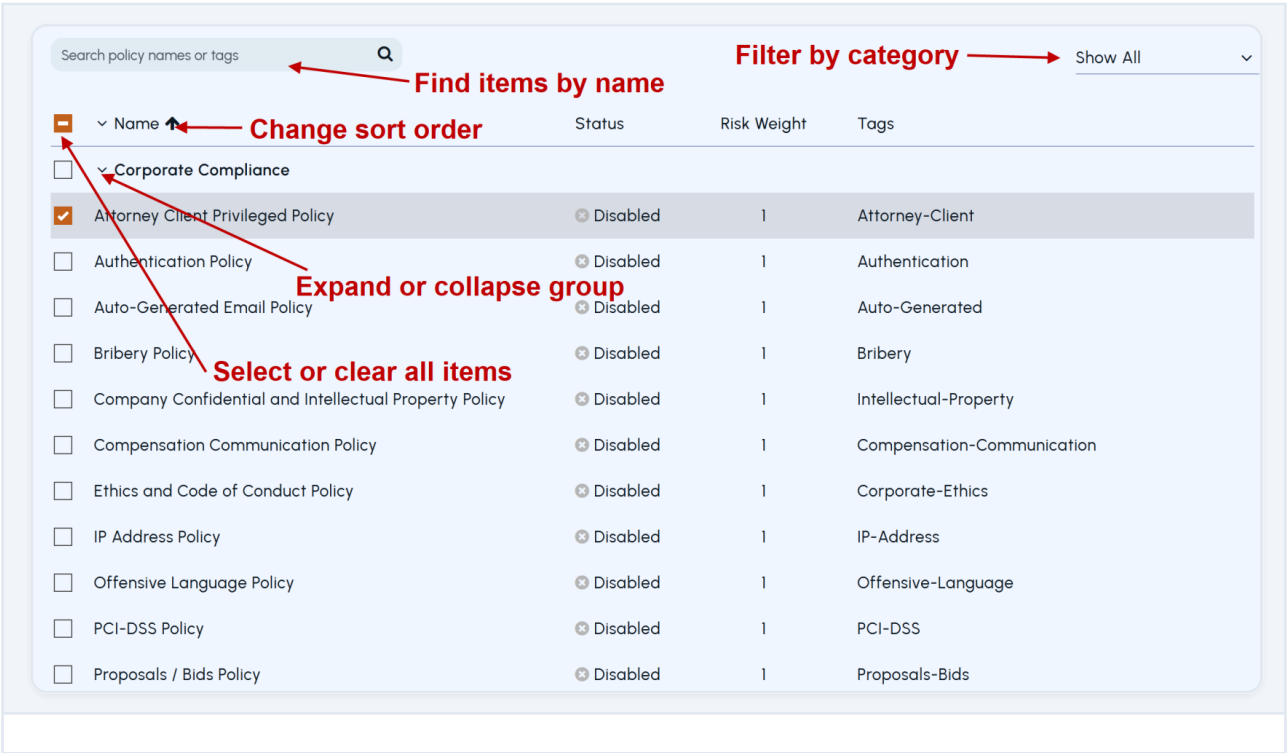
The Arctera Insight Classification window is divided into three main areas: the navigation bar, item list, and details pane.



Item list

The item list provides a list of the available items, together with basic information about them. Click an item to view more information in the details pane.

The controls at the top of each list let you search for items by name, filter the items according to various criteria, expand and collapse the list, and change the sort order.



Details pane

The details pane provides extensive information on the selected item. You also use this pane to edit an item or create a new one to add to the list.

> “ ”

Note: Arctera Insight Classification has option to opt for light or dark theme. For ease og understanding, we will be referring to the light theme in documentation.

> “ ”

Tenant Level Preferences

Tenant-level preferences allow for the customization of application behavior to meet the unique requirements of individual tenants. These preferences are dynamically stored, managed, and accessed to control the application's behavior based on tenant-specific settings.

Please note that you need to reach out to the Arctera ops team for updating these settings. You will not be able to update settings either from the platform or through YAML configuration file.

Tenant Level Preferences Properties

The tenant-level preferences currently include the following five configurable properties:

- Sentiment Analysis Feature
  - Enable or disable the use of sentiment analysis features.
  - Default Value: false
- Partial Classification Feature
  - Enable or disable the use of partial classification features.
  - Default Value: false
- Maximum Data Size for Partial Classification
  - If partial classification is enabled, the classification will be performed on a specified data size.
  - Default Value: 10 MB
- Maximum Content Matches per Rule
  - The maximum number of content matches included per rule.
  - Default Value: 100

- Maximum Content Matches in Classification Response
  - The maximum number of content matches included in the document classification response.
  - Default Value: 3000

Supported languages

FEATURE	SUPPORTED LANGUAGES
Language detection	33 language detection policies to detect a specific primary language.
	Multiple Languages policy and Primary Language - Unknown policy.
	See About built-in classification policies .
Named Entity Recognition (NER)	English

Supported file formats

If you are using Arctera Classification to extract the content the extraction engine will classify only the file formats that are part of the list given below. Arctera Classification has also provided the configurable field to exclude the *MIME* types from the supported list. It means that the user can exclude certain file formats from classification even if the file format is included in the list below.

If the source file is not within the supported type list *UNSUPPORTED_FILE_FORMAT* error will appear.

The list of supported file formats is as follows:

message/rfc822

application/vnd.ms-htmlhelp

application/atom+xml

application/vnd.visio

image/x-xcf

image/wmf

application/x-matlab-data

image/vnd.dgn

application/deflate64

application/vnd.ms-powerpoint.slideshow.macroenabled.12

application/vnd.openxmlformats-officedocument.presentationml.slide

application/vnd.apple.keynote

application/vnd.oasis.opendocument.spreadsheet-template

application/x-xliff+xml

application/x-ibooks+zip

application/x-plist

application/vnd.ms-word.document.macroenabled.12

application/vnd.apple.unknown.13

application/pdf

application/mp4

application/x-midi

application/rss+xml text/html

application/vnd.ms-visio.template text/csv image/ vnd.microsoft.icon

application/zlib

application/x-sh

application/vnd.wordperfect

application/x-wacz

application/vnd.apple.numbers

application/x-archive application/vnd.ms-word2006ml

application/onenote application/x-tika-msoffice

application/x-font-adobe-metric

application/vnd.ms-visio.drawing

application/java-archive

application/sldworks

application/x-httpresponse

application/x-tika-ooxml-protected

application/vnd.ms-excel.sheet.macroenabled.12

image/heif

image/heic

application/vnd.ms-outlook

application/vnd.ms-word.template.macroenabled.12

application/x-compress

image/vnd.dwg

text/x-groovy

application/vnd.openxmlformats-officedocument.spreadsheetml.template

model/vnd.dwfx+xps

application/vnd.openxmlformats-officedocument.spreadsheetml.sheet

application/vnd.oasis.opendocument.chart-template

application/zip

application/vnd.oasis.opendocument.text-master

application/vnd.oasis.opendocument.tika.flat.document

image/vnd.adobe.photoshop

image/gif

application/x-sharedlib

application/java-vm

image/webp

application/x-activemime

application/vnd.adobe.indesign-idml-package

application/vnd.wap.xhtml+xml

application/applefile

application/vnd.ms-spreadsheetml

application/msword

application/vnd.apple.numbers.13

application/x-bplist-memgraph

application/x-java-pack200

application/vnd.oasis.opendocument.image-template

application/warc

application/rtf

image/bpg

application/vnd.oasis.opendocument.text

application/x-tnef

application/x-xz

application/vnd.ms-powerpoint.template.macroenabled.12

image/vnd.wap.wbmp

application/kate

application/pkcs7-mime

application/x-executable

application/x-coredump

application/x-msaccess

application/vnd.apple.numbers.18

text/plain

image/png

application/vnd.ms-outlook-pst

application/x-cpio

application/x-tika-msworks-spreadsheet

application/vnd.apple.pages

application/vnd.ms-xpsdocument

application/vnd.ms-visio.template.macroenabled.12

application/x-tar

application/vnd.oasis.opendocument.presentation-template

application/x-bplist

image/x-jbig2

application/x-dbf

application/vnd.ms-excel.template.macroenabled.12

application/mbox

application/vnd.oasis.opendocument.formula

application/chm

application/vnd.ms-excel.workspace.3

application/vnd.ms-excel.workspace.4

image/x-bpg

application/x-xliff+zip

application/vnd.openxmlformats-officedocument.wordprocessingml.template

application/vnd.apple.iwork

image/heic-sequence

application/vnd.ms-excel.sheet.2

application/vnd.ms-excel.sheet.3

application/vnd.ms-powerpoint.presentation.macroenabled.12

application/x-brotli application/dif+xml

application/vnd.ms-excel

application/vnd.ms-excel.sheet.4

application/x-tika-ole-drm-encrypted

application/x-tika-ooxml

application/vnd.ms-project

application/epub+zip

application/x-quattro-pro

application/x-sas-data

application/x-snappy

application/vnd.oasis.opendocument.text-template

application/vnd.openxmlformats-officedocument.presentationml.presentation

application/vnd.ms-visio.stencil

application/vnd.ms-visio.stencil.macroenabled.12

application/x-bplist-webarchive

application/xml

application/vnd.oasis.opendocument.flat.presentation

image/bmp

application/xhtml+xml

application/pkcs7-signature

text/x-java-source

application/vnd.sun.xml.writer

application/vnd.oasis.opendocument.formula-template

application/x-tmx

application/vnd.ms-powerpoint.addin.macroenabled.12

application/vnd.ms-powerpoint application/timestamped-data

text/x-c++src

application/vnd.openxmlformats-officedocument.presentationml.template

application/x-bplist-itunes

image/tiff

application/vnd.ms-excel.addin.macroenabled.12

application/vnd.ms-wordml

application/x-object

application/x-asp

application/x-mspublisher

application/x-hwp-v5

application/x-rar-compressed

image/heif-sequence

application/vnd.oasis.opendocument.graphics-template

application/vnd.openxmlformats-officedocument.wordprocessingml.document

application/vnd.ms-powerpoint.slide.macroenabled.12

application/vnd.ms-tnef application/ vnd.oasis.opendocument.text-web

application/x-maker

image/emf

application/x-bzip

application/vnd.oasis.opendocument.graphics

text/vnd.iptc.anpa

application/vnd.apple.keynote.18

application/x-arj

application/x-lzma

application/vnd.apple.keynote.13

application/x-lz4

application/vnd.oasis.opendocument.flat.spreadsheet

application/vnd.oasis.opendocument.presentation

application/vnd.mif

application/x-tika-msoffice-embedded application/x-7z-compressed

image/jxl

application/x-msdownload

application/vnd.oasis.opendocument.chart

image/jpeg

image/icns

application/ogg

image/svg+xml

application/vnd.ms-excel.sheet.binary.macroenabled.12

application/warc+gz

application/x-ms-owner

application/gzip

application/x-tika-unix-dump

application/vnd.oasis.opendocument.spreadsheet

application/vnd.apple.pages.18

application/vnd.oasis.opendocument.image

application/x-bzip2

application/vnd.apple.pages.13

application/x-fictionbook+xml

application/vnd.oasis.opendocument.flat.text

application/x-elf

text/tsv

application/vnd.ms-visio.drawing.macroenabled.12

application/vnd.openxmlformats-officedocument.presentationml.slideshow

application/x-chm

application/x-font-ttf

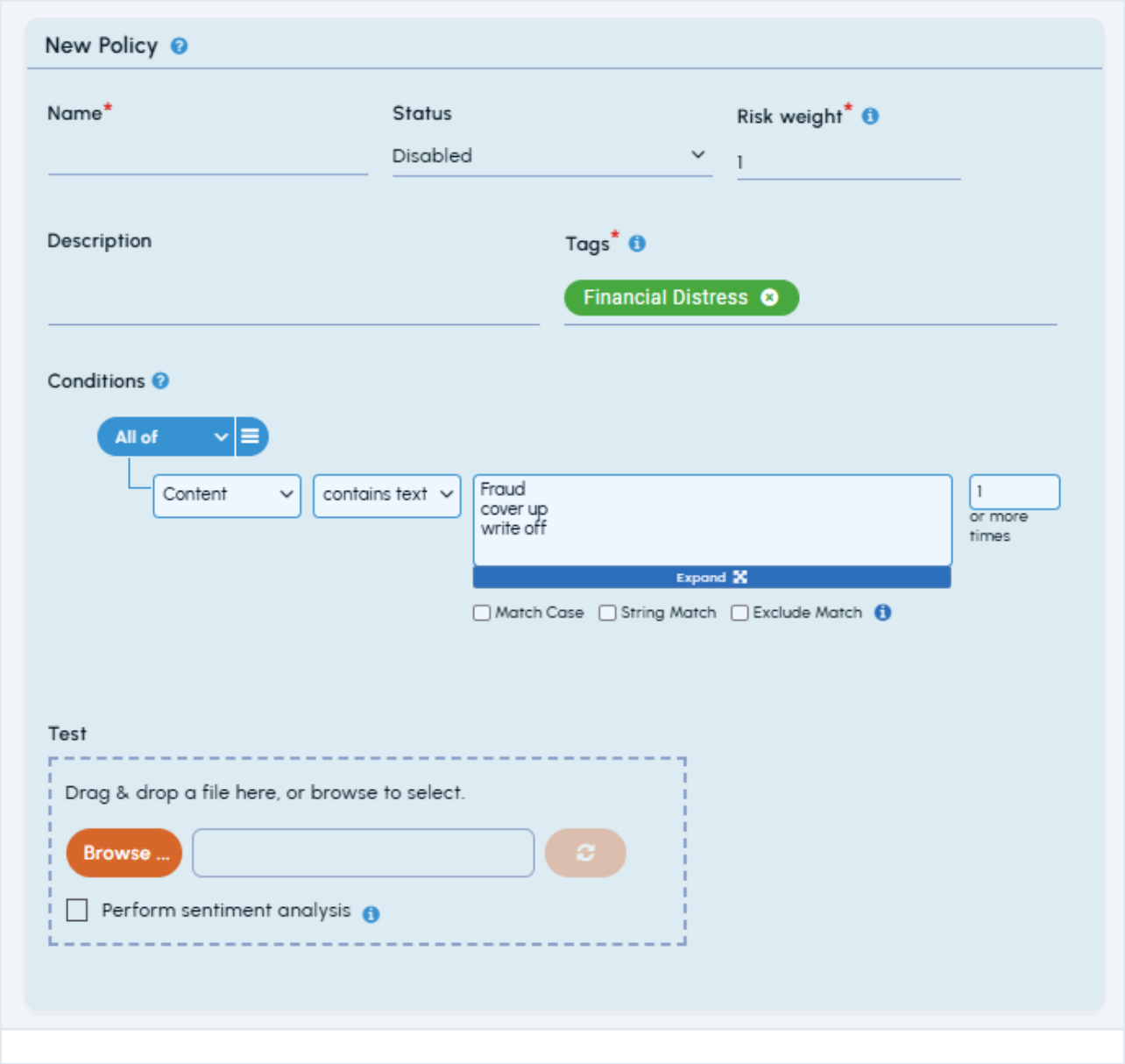
application/x-prt

Policies

This section includes the following topics:

About policies

The Arctera Insight Classification evaluates the items that you submit for classification against a set of policies. Each policy specifies the conditions that the items must meet to be assigned a specific classification tag. For example, you can create a simple policy to assign the tag *Financial Distress*to items that contain any of the terms*fraud*,*cover up*, and*write off*.



The Arctera Insight Classification comes with a large number of built-in policies, but you can create custom policies if the built-in ones do not meet your needs.

Initially, all the policies are disabled. You must enable a policy if you want the Arctera Insight Classification to check for and tag the items that match the policy.

About built-in classification policies

The built-in classification policies are arranged in the following groups.

Table: Corporate Compliance

POLICY	DETECTS
Attorney Client Privileged Policy	Documents which are protected by attorney-client confidentiality.
Authentication Policy	Authentication information, such as user name and password credentials.
Auto-Generated Email Policy	Detects an email header commonly tied to auto-generated email from
	Microsoft Exchange and Exchange Online such as out of office replies,
	and read receipts.
Bribery Policy	Language in communication which suggests bribery or quid pro quo.
Company Confidential and Intellectual Property Policy	Documents that are confidential, secret, or internal-only, or that contain intellectual property source code.
Disclaimers Policy	Detects common e-mail disclaimers for confidential, contract, copyright, environment,
	external e-mail, financial, regulatory, liability, and promotional content.
Gifts and Entertainment Policy	Detects the offering or accepting of inappropriate gifts and entertainment. This policy aids in meeting compliance obligations for financial supervision.
Compensation Communication Policy	Communication about compensation.
Ethics and Code of Conduct Policy	Terms that may be unethical or against corporate code of conduct policy.
IP Address Policy	Internet Protocol version 4 (IPv4) and version 6 (IPv6) addresses.

POLICY	DETECTS
Offensive Language Policy	Language containing curses, insults, racism, sexism, or other offensive content.
Offensive Terms Policy	Detects offensive and vulgar content. While some terms overlap, this
	policy contains different keywords and logic that those included in the
	Offensive Language Policy and Ethics and Code of Conduct Policy.
	Note: This policy is composed of two patterns: High Confidence: contains
	terms that tend to be more offensive and aggressive and Low
	Confidence: contains terms that are either less offensive and aggressive
	and/or are more likely to generate matches.
	Note: This is not a comprehensive list of all offensive terms. Nor are all
	connotations of the terms included in this policy considered offensive.
	Please update this policy to best address your requirements.
	Note: As a transparent policy, this policy and its patterns can be fully
	edited. To modify, copy the policy and/or the pattern and then edit as
	needed. It is a best practice to use Transparent Policies in conjunction

POLICY	DETECTS
	with noise (junk) minimization techniques to reduce false positives. Please
	see online help for more assistance with transparent policies.
PCI-DSS Policy	Content that is subject to the Payment Card Industry Data Security Standard (PCI-DSS), including credit and debit card numbers.
Proposals / Bids Policy	Corporate proposal and bid documents.
Ransomware Policy	Ransom notes/demands/instructions from ransomware malware.
Resumes / CVs Policy	Detects resume and curriculum vitae (CV) documents.
Software Keys and Tokens Policy	Detect secret keys, tokens, and credentials used to authenticate usage of software APIs.
Workplace Harassment Policy	Content that may violate federal, state, local, corporate, and workplace or sexual harassment policies and regulations.
	This policy helps detect content subject to corporate/workplace policy and legislative regulations such as Title VII of the Civil Rights Act of 1964, California AB 1825 (Part of) Fair Employment and Housing Act, Connecticut Public Act No. 19-16 Time's Up Act, Delaware SB 360 (Part of) Delaware Discrimination in Employment Act (DDEA), Maine Revised Statutes Title 26, section 807 (Enhanced by) Maine Human Rights Act, New York Assembly Bill A8421 AKA Stop Sexual Harassment in NYC Act, and others.

Table: Financial Regulations

POLICY	DETECTS
Anti-Money Laundering (AML) Policy (V2)	Potential signs and symptoms of money laundering presented in conversation. This policy matches based on word patterns. This policy is provided for content classification purposes only.
Bank Account Number Policy	Country-specific or international bank account numbers.
Credit Card Policy	Credit and debit cards.
Gramm-Leach-Bliley Act (GLBA) Policy	Personal financial information for Financial Services Modernization Act of 1999, also known as Gramm-Leach-Bliley Act (GLBA) or Public Law 106-102.
Market Abuse Policy	Detects key phrases of market abuse, specifically market manipulation
	and insider trading. This policy has been tuned to departments composed
	of licensed employees. In other departments, running this policy may
	produce a higher volume of false positives. As a transparent policy, this
	policy and its patterns can be fully edited. To modify, copy the policy
	and/or the pattern and then edit as needed. It is a best practice to use
	Transparent Policies in conjunction with noise (junk) minimization
	techniques to reduce false positives. Please see online help for more

POLICY	DETECTS
	assistance with transparent policies.
Material Non-Public Information (MNPI) Policy (V2)	Potential signs and symptoms of material non-public information (MNPI) presented in conversation. This policy matches based on word patterns. This policy is provided for content classification purposes only.
Off-Channel Signaling Policy	Used to Detects Off-Channel Signaling in communication (language suggesting that a conversation is occurring on a non-monitored channel).
	The policy includes the following patterns:
	Signaling General Terms
	Signaling to Phone or Public Email Terms
	Signaling to SMS or Text Terms
	Signaling to Fileshares (for Citrix and Dropbox where both proximity logic and short phrases are used for detection)
	Policies and Patterns Arctera Classification Policies
	106
	Signaling to Low Frequency Channels (where the mere mention of the channel is used to trigger the policy)
	Signaling to Medium Frequency Channels (where both proximity logic and short phrases are used for detection)
	Signaling to High Frequency Channels (where channel name is common in communication (Facebook), or is a common word or initials

POLICY	DETECTS
	(Element or IG). Short phrases are used for detection)
	RECOMMENDATION: Before enabling, review/edit this policy to ensure proper targeting of channels and activities as they relate to un/approved channels and policies at your firm. Click on the hyperlink to each pattern below to see the details as to which channels are included and which terms are used.
	As a transparent policy, this policy and its patterns can be fully edited. To modify, copy the policy and/or the pattern and then edit as needed. It is a best practice to use Transparent Policies in conjunction with noise (junk) minimization techniques to reduce false positives. Please see online help for more assistance with transparent policies.
United States Sarbanes-Oxley (SOX) Policy	Forms and terms related to the U.S. Sarbanes - Oxley Act of 2002 (SOX, Public Law 107).
Swiss Financial Market Supervision Act (FINMASA) Policy	Data elements subject to the Federal Act on the Swiss Financial Market Supervisory Authority 956.1, Financial Market Supervision Act (FINMASA)
SWIFT Codes Policy	Society for Worldwide Interbank Financial Telecommunication (SWIFT) codes, also known as Bank Identifier Codes (BIC), Business Identifier Codes (BIC), or ISO 9362, and related content.
U.S. SEC Regulation Best Interest (RegBI) Policy	Potential signs and symptoms of communications to be reviewed for compliance with U.S. Securities and Exchange Commission (SEC) Regulation Best Interest (RegBI) rule. This policy matches based on word patterns. This policy is provided for content classification purposes only.

POLICY	DETECTS
United States Financial Forms / Documents Policy	U.S. financial forms and documents.

Table: Health Regulations

POLICY	DETECTS
Australia Individual Healthcare Identifier (IHI) Policy	Australia Individual Healthcare Identifiers (IHI) and related content.
Canada Healthcare Identifiers Policy	Canada Healthcare Identifiers and related content.
Coronavirus (COVID) Policy	Coronavirus disease names and related terms. This policy template can be extended with additional data elements to find content containing Coronavirus/COVID in specific contexts. For example, patient health, financial benefits, and so on.
ICD 10 Diagnosis Policy	Detect International Classification of Diseases (ICD) 10 diagnosis indexes (textual names) and codes (alpha-numeric).
Medical Diagnosis - Allergy Policy	Proposes possible Allergy medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Anemia Policy	Proposes possible Anemia medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is

POLICY	DETECTS
	provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Anxiety Policy	Proposes possible Anxiety medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Appendicitis Policy	Proposes possible Appendicitis medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Bladder Infection Policy	Proposes possible Bladder Infection medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Blood Clot Policy	Proposes possible Blood Clot medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes

POLICY	DETECTS
	only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Brain Tumor Policy	Proposes possible Brain Tumor medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Breast Cancer Policy	Proposes possible Breast Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Cervical Cancer Policy	Proposes possible Cervical Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Colon Cancer Policy	Proposes possible Colon Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes

POLICY	DETECTS
	only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Dengue Policy	Proposes possible Dengue medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Depression Policy	Proposes possible Depression medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Diabetes Policy	Proposes possible Diabetes medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Female Sexually Transmitted Disease (STD) Policy	Proposes possible Female Sexually Transmitted Disease (STD) medical
	condition or diagnosis based on medical signs and symptoms presented in
	a form or in conversation. This policy matches based on English patterns

POLICY	DETECTS
	and offers limit recommendation capabilities. This policy is provided for
	content classification purposes only. Patients should consult a qualified
	physician for all diagnosis and treatment.
Medical Diagnosis - Heart Attack Policy	Proposes possible Heart Attack medical condition or diagnosis based on medical
	signs and symptoms presented in a form or in conversation. This policy matches
	based on English patterns and offers limit recommendation capabilities. This policy
	is provided for content classification purposes only. Patients should consult a
	qualified physician for all diagnosis and treatment.
Medical Diagnosis - Hernia Policy	Proposes possible Hernia medical condition or diagnosis based on medical signs
	and symptoms presented in a form or in conversation. This policy matches based
	on English patterns and offers limit recommendation capabilities. This policy is provided
	for content classification purposes only. Patients should consult a qualified
	physician for all diagnosis and treatment.
Medical Diagnosis - Herpes Policy	Proposes possible Herpes medical condition or diagnosis based on medical signs

POLICY	DETECTS
	and symptoms presented in a form or in conversation. This policy matches based
	on English patterns and offers limit recommendation capabilities. This policy is provided
	for content classification purposes only. Patients should consult a qualified
	physician for all diagnosis and treatment.
Medical Diagnosis - HIV/AIDS Policy	Proposes possible HIV/AIDS medical condition or diagnosis based on
	medical signs and symptoms presented in a form or in conversation. This
	policy matches based on English patterns and offers limit recommendation
	capabilities. This policy is provided for content classification purposes
	only. Patients should consult a qualified physician for all diagnosis and
	treatment.
Medical Diagnosis - Hypertension / High Blood Pressure Policy	Proposes possible Hypertension or High Blood Pressure medical condition or diagnosis
	based on medical signs and symptoms presented in a form or in conversation.
	This policy matches based on English patterns and offers limit recommendation
	capabilities. This policy is provided for content classification purposes only.

POLICY	DETECTS
	Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Irritable Bowel Syndrome (IBS) Policy	Proposes possible Irritable Bowel Syndrome (IBS) medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Kidney Stone Policy	Proposes possible Kidney Stone medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Leukemia Policy	Proposes possible Leukemia medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Lung Cancer Policy	Proposes possible Lung Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes

POLICY	DETECTS
	only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Lupus Policy	Proposes possible Lupus medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Lyme Disease Policy	Proposes possible Lyme Disease medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Male Sexually Transmitted Disease (STD) Policy	Proposes possible Male Sexually Transmitted Disease (STD) medical
	condition or diagnosis based on medical signs and symptoms presented in
	a form or in conversation. This policy matches based on English patterns
	and offers limit recommendation capabilities. This policy is provided for
	content classification purposes only. Patients should consult a qualified
	physician for all diagnosis and treatment.
Medical Diagnosis - Menopause Policy	Proposes possible Menopause medical condition or diagnosis based on medical signs

POLICY	DETECTS
	and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Mononucleosis Policy	Proposes possible Mononucleosis medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Ovarian Cancer Policy	Proposes possible Ovarian Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Pancreatic Cancer Policy	Proposes possible Pancreatic Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Pneumonia Policy	Proposes possible Pneumonia medical condition or diagnosis based on medical signs and symptoms presented in a form or in

POLICY	DETECTS
	conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Pregnancy Policy	Proposes possible Pregnancy medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Prostate Cancer Policy	Proposes possible Prostate Cancer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Sinus Infection Policy	Proposes possible Sinus Infection medical condition or diagnosis based on medical
	signs and symptoms presented in a form or in conversation. This policy matches
	based on English patterns and offers limit recommendation capabilities. This policy
	is provided for content classification purposes only. Patients should consult a
	qualified physician for all diagnosis and treatment.

POLICY	DETECTS
Medical Diagnosis - Stomach Ulcer Policy	Proposes possible Stomach Ulcer medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Strep Throat Policy	Proposes possible Strep Throat medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
Medical Diagnosis - Urinary Tract Infection Policy	Proposes possible Urinary Tract Infection medical condition or diagnosis
	based on medical signs and symptoms presented in a form or in
	conversation. This policy matches based on English patterns and offers
	limit recommendation capabilities. This policy is provided for content
	classification purposes only. Patients should consult a qualified physician
	for all diagnosis and treatment.
Medical Diagnosis - Yeast Infection Policy	Proposes possible Yeast Infection medical condition or diagnosis based on medical signs and symptoms presented in a form or in conversation. This policy matches based on

POLICY	DETECTS
	English patterns and offers limit recommendation capabilities. This policy is provided for content classification purposes only. Patients should consult a qualified physician for all diagnosis and treatment.
United States Medical Record Number Policy	Medical record numbers (generically).
United States California Confidentiality of Medical Information Act (CMIA) Policy	Individually identifiable medical information for United States California Confidentiality of Medical Information Act (CMIA), specified in California Civil Code §§56.
U.S. Drug Enforcement Agency (DEA) Number Policy	U.S. Drug Enforcement Agency (DEA) numbers and related content.
US Stock Financial Advice Policy	Detect US stocks along with emojis which may suggest financial
	advice.
United States Health Insurance Portability and Accountability Act (HIPAA) Policy	Electronic patient health information (ePHI) information for United States Health Insurance Portability and Accountability Act (HIPAA), also known as Public Law 104-191.
US-COVID-19 Policy	Look for indications of Corona Virus related content.
US-COVID-19 ICD-10-CM Diagnosis Indexes (CDC Feb-Apr-2020) Policy	Detect COVID-19 specific ICD-10-CM diagnosis indexes per Feb 20, 2020 CDC Guidelines.
	https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf
	https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf

POLICY	DETECTS
	https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
	Policy version 04242020-1
US-COVID-19 ICD-10-CM Diagnosis Indexes possibly leading to COVID-19 (CDC-Feb-Apr-2020) Policy	Detect ICD-10-CM Diagnosis Indexes potentially leading to COVID-19 Per Feb & Apr 2020 CDC Guidelines.
	https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf
	https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf
	https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
	Policy version 04242020-1
US-COVID-19 ICD-10-CM Exposure Diagnosis Index (CDC Feb-Apr-2020) Policy	Detect COVID-19 ICD-10-CM diagnosis index for when where there is an actual exposure to someone who is confirmed or suspected (not ruled out) to have COVID-19. Per Feb 20, 2020 CDC Guidelines.
	https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf
	https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf
	https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
	Policy version 04242020-1

POLICY	DETECTS
US-COVID-19 ICD-10-CM Ruled Out Diagnosis Index (CDC Feb-Apr-2020) Policy	Detect COVID-19 ICD-10-CM diagnosis index for cases where there is a concern about a possible exposure to COVID-19 but it is ruled out.. Per Feb 20, 2020 CDC Guidelines.
	https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf
	https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf
	https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
	Policy version 04242020-1
US-COVID-19 ICD-10-CM Signs and Symptoms Diagnosis Indexes (CDC Feb-Apr-2020) Policy	Detect COVID-19 specific ICD-10-CM signs and symptoms diagnosis indexes where a definitive diagnosis has not been established. Per Feb 20, 2020 CDC Guidelines.
	https://www.cdc.gov/nchs/data/icd/ICD-10-CM-Official-Coding-Gudance-Interim-Advice-coronavirus-feb-20-2020.pdf
	https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf
	https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf
	Policy version 04242020-1

Table: International Regulations

POLICY	DETECTS
Australia Drivers License Number Policy	Australian driver's license numbers.
Australia Passport Policy	Australian passport numbers.
Australia Tax Policy	Australian tax file number and related content.
Canada Drivers License Number Policy	Canadian driver's license numbers.
Canada Passport Policy	Canadian passport numbers.
Canada Social Insurance Number Policy	Canadian social insurance numbers.
France National ID Policy	French National Identifiers and related content.
Italy Codice Fiscale Policy	Italian Codice Fiscale numbers, also known as Italian fiscal code card numbers.
Switzerland National ID Policy	Swiss National Identifiers and related content.
U.K. Drivers License Number Policy	U.K. driver's license numbers and related content.
U.K. National ID Policy	U.K. National Identifiers and related content.
U.K. National Insurance Number (NINO) Policy	U.K. National Insurance number and related content.
U.K. Passport Number Policy	U.K. passport number and related content.
U.K. Unique Tax Reference (UTR) Policy	U.K. Unique Tax Reference (UTR) number and related content.
United States Drivers License Number Policy	U.S. driver's license numbers.
United States Passport Policy	U.S. passport and passport card numbers.
United States Social Security Number (SSN) and Taxpayer ID Policy	U.S. taxpayer identification numbers. This is typically the U.S. Social Security Number (SSN).

Language detection

By default, Arctera Insight Classification determines the language in a message if there are at least 80 characters. Starting with release 2.4.0, the administrators can configure the minimum number of characters and a higher or lower confidence level for language detection. When multiple languages are present in small files, the administrator can specify a smaller size of each chunk that language detection is performed on. With this enhancement, Arctera Insight Classification attempts to detect languages with the configurable parameters.

Starting with release 3.0.0, you can create custom primary language detection policies for languages other than the languages part of the built-in Language Detection policies.

Table: Language Detection

POLICY	DETECTS
Multiple Languages Policy	Multiple languages are detected.
Primary Language - Arabic Policy	Primary language detected is Arabic.
Primary Language - Bengali Policy	Primary language detected is Bengali.
Primary Language - Chinese Policy	Detects primary language - Chinese.
Primary Language - Czech Policy	Primary language detected is Czech.
Primary Language - Danish Policy	Primary language detected is Danish.
Primary Language - Dutch Policy	Primary language detected is Dutch.
Primary Language - English Policy	Primary language detected is English.
Primary Language - Finnish Policy	Primary language detected is Finnish.
Primary Language - French Policy	Primary language detected is French.
Primary Language - German Policy	Primary language detected is German.
Primary Language - Greek Policy	Primary language detected is Greek.
Primary Language - Hebrew Policy	Primary language detected is Hebrew.
Primary Language - Hindi Policy	Primary language detected is Hindi.

POLICY	DETECTS
Primary Language - Hungarian Policy	Primary language detected is Hungarian.
Primary Language - Indonesian (Bahasa) Policy	Primary language detected is Indonesian (Bahasa).
Primary Language - Italian Policy	Primary language detected is Italian.
Primary Language - Japanese Policy	Primary language detected is Japanese.
Primary Language - Korean Policy	Primary language detected is Korean.
Primary Language - Norwegian Policy	Primary language detected is Norwegian.
Primary Language - Persian (Farsi) Policy	Primary language detected is Persian (Farsi).
Primary Language - Polish Policy	Primary language detected is Polish.
Primary Language - Portuguese Policy	Primary language detected is Portuguese.
Primary Language - Romanian Policy	Primary language detected is Romanian.
Primary Language - Russian Policy	Primary language detected is Russian.
Primary Language - Slovak Policy	Primary language detected is Slovak.
Primary Language - Spanish Policy	Primary language detected is Spanish.
Primary Language - Swedish Policy	Primary language detected is Swedish.
Primary Language - Tagalog (Filipino) Policy	Primary language detected is Tagalog (Filipino).
Primary Language - Thai Policy	Primary language detected is Thai.
Primary Language - Turkish Policy	Primary language detected is Turkish.
Primary Language - Unknown Policy	Primary language detected is Unknown.
Primary Language - Urdu Policy	Primary language detected is Urdu.
Primary Language - Vietnamese Policy	Primary language detected is Vietnamese.

Table: Personally Identifiable Information

POLICY	DETECTS
Argentina Personal Data Policy	Personal data as defined by Argentina's Personal Data Protection Law No. 25,326 (PDPL), Argentina Bill No. MEN-2018-147-APN-PTE, and a general similarity to the European General Data Protection Regulation (GDPR).
Australia Personal Data Policy	Personal data or personal information as defined by Australia's Federal Privacy Act 1988 (Privacy Act), Privacy (Tax File Number) Rule 2015, Taxation Administration Act 1953, , and cross-compatibility with European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Austria Personal Data Policy	Personal data as defined by the Data Protection Act 2000 (DSG 2000), Regulation (EU) 611/2013 on the measures applicable to the notification of personal data breaches under Directive 2002/58/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Belgium Personal Data Policy	Personal data applicable to Belgium's Data Protection Directive (DPL), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Bolivia Personal Data Policy	Detects personal data as defined by Bolivia Political Constitution of the Plurinational State of Bolivia, in Article Nº130, Bolivia Bill of Personal Data Protection, Bolivia Law 145 - Personal Identification Registry, and a general similarity to the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Brazil Personal Data Policy	Personal data applicable to Brazil's General Data Privacy Law, Law No. 13,709, Lei Geral de Proteção de Dados Pessoais (LGPD), and a general similarity to the European General Data Protection Regulation (GDPR).
Bulgaria Personal Data Policy	Personal data as defined by Bulgaria's Personal Data Protection Act, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Canada Personal Data (PIPEDA) Policy	Personal data as defined by Canada's Personal Information Protection and Electronic Documents Act (PIPEDA), the Data Privacy Act, the Model Care for the Protection of Personal Information, and a general similarity to the European General Data Protection Regulation (GDPR).
Chile Personal Data Policy	Personal data as defined by Chile's Privacy Protection Law, N° 19.628
	(1999) ("Personal Data Protection Law"), Law N°21.096 on the right to the protection
	of personal data (June/2018), Article 19 N°4 of the Chilean Constitution,
	Law N°21.081 Article 1 N°39 (2018) on consumer rights protection, and a general
	similarity to the European General Data Protection Regulation (GDPR).
China Personal Data Policy	Personal data as defined by China Personal Information Protection Law (PIPL), China National People's Congress (NPC) Cybersecurity Law, China National Information Security Standardization Technical Committee (TC260) Personal Information Security

POLICY	DETECTS
	Specification, and a general similarity to the European General Data Protection Regulation (GDPR).
Colombia Personal Data Policy	Personal data as defined by Articles 15 and 20 of Colombia's Constitution, Colombia Statutory Law 1266 of 2008, Colombia Statutory Law 1581 of 2012, Colombia Law 2157 of 2021, Colombia Decree 1377 of 2013, Colombia Decree 886 of 2014, and a general similarity to the European General Data Protection Regulation (GDPR).
Croatia Personal Data Policy	Personal data as defined by Croatia's Personal Data Protection Law ("DP Law") nos. 103/2003, 118/2006, 41/2008 and 130/2011, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Cyprus Personal Data Policy	Personal data as defined by Cyprus' Processing of Personal Data (Protection of the Individual) Law 138(I)/2001, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Czech Republic Personal Data Policy	Personal data as defined by the Czech Republic's Protection of Personal Data Act No. 101/2000 Coll. (PPD Act), Act on Certain Aspects of Information Society Services Information No. 480/2004 Coll., Electronic Communications Act No. 127/2005 Coll., European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Denmark Personal Data Policy	Personal data applicable to Denmark's Act on Processing of Personal Data (APPD), European Privacy Directive 95/46/EC on data

POLICY	DETECTS
	protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Egypt Personal Data Policy	Personal data as defined by Egypt Data Protection Law (Resolution/Law No. 151 of 2020), and a general similarity to the European General Data Protection Regulation (GDPR).
El Salvador Personal Data Policy	Detects personal data as defined by El Salvador's Personal Data
	Protection Act, Personal Data Law, Law for the Regulation of Information
	Services on People's Credit History, and a general similarity to the
	European General Data Protection Regulation (GDPR).
Ecuador Personal Data Policy	Personal data as defined by Ecuador's Organic Law on Personal Data
	Protection, Constitution of the Republic of Ecuador, and a general similarity to the
	European General Data Protection Regulation (GDPR).
Estonia Personal Data Policy	Personal data as defined by Estonia's Personal Data Protection Act, Electronic Communications Act, the Information Society Services Act, Directives 2002/58 and 2009/136/EC on Privacy and Electronic Communications, Electronic Communications Act, Directive 2006/24/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Finland Personal Data Policy	Personal data as defined by Finland's Constitution (731/1999), the Personal Data Act (523/1999) (henkilötietolaki) (PDA), the Act on the Data Protection Board and the Data Protection Ombudsman (389/1994), Information Society Code (917/2014) (ISC), the Act on the Protection of Privacy in Working Life (759/2004) (WPA), Credit Data Act (527/2007) (CDA), Regulation (EU) 611/2013 on the measures applicable to the notification of personal data breaches under Directive 2002/58/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
France Personal Data Policy	Personal data applicable to France's Federal and State laws, including France's Act No 78-17 on Information Technology, Data Files and Civil Liberties (DPA), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Germany Personal Data Policy	Detects personal data applicable to Germany's Federal and State laws,
	including Germany's Federal Data Protection Act (BDSG), European
	Privacy Directive 95/46/EC on data protection (Data Protection Directive),
	and the European General Data Protection Regulation (GDPR).
Greece Personal Data Policy	Personal data as defined by Greece's Data Protection Act no 2472/1999 (DPA), Directive 2002/58/EC (Privacy and Electronic Communications Directive), implemented by Law no 3471/2006 on the protection of

POLICY	DETECTS
	personal data and privacy in the telecommunications sector (TDPA), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Guatemala Personal Data Policy	Detects personal data as defined by Guatemala's Law of Access to Public Information, Political Constitution of the Republic of Guatemala 1985, and a general similarity to the European General Data Protection Regulation (GDPR).
Hong Kong Personal Data Policy	Personal data as defined by Hong Kong's Personal Data (Privacy) Ordinance (PDPO), Laws of Hong Kong (Cap 486) (PDPO), and a general similarity to the European General Data Protection Regulation (GDPR).
Honduras Personal Data Policy	Detects personal data as defined by Honduras's Constitution of the Republic of
	Honduras. Article 76, 100 and 182, Law of the National Civil Registry. Decree
	No. 62-2004, Draft of Law on Protection of Personal Data (not yet in force), and a
	general similarity to the European General Data Protection Regulation (GDPR).
Hungary Personal Data Policy	Personal data as defined by Hungary's Act No. CXII on the Right to Informational Self Determination and Freedom of Information, Act XLVII on Processing and Protection of Medical and Other Related Personal Data (Medical Data Act), Act LXVI on Personal Data and Address Records of Citizens, Act C on Electronic Communications (Electronic Communications Act), Act CXIX on Processing of Name and Address Data for Research and Direct Marketing Purposes, Act CXII on Credit

POLICY	DETECTS
	Institutions and Financial Undertakings (Credit Institutions Act), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Iceland Personal Data Policy	Personal data applicable to Iceland's Data Protection Directive; detects personal data applicable to European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
India Personal Data Policy	Personal data as defined by India's Personal Data Protection Bill (DPB) 2019, India's Information Technology Act 2000, and a general similarity to the European General Data Protection Regulation (GDPR).
Indonesia Personal Data Policy	Personal data as defined by Indonesia's Law No. 11 of 2008 on Electronic Information and Transaction, as amended by Law No. 19 of 2016 regarding the Amendment of EIT Law ("EIT Law Amendment"), Government Regulation No. 71 of 2019 regarding Provisions of Electronic Systems and Transactions, Minister of Communication & Informatics Regulation No. 20 of 2016 regarding the Protection of Personal Data in an Electronic System, and a general similarity to the European General Data Protection Regulation (GDPR).
Ireland Personal Data Policy	Personal data as defined by Ireland's Data Protection Act 1988, the Data Protection (Amendment) Act 2003 (DPA), European Communities (Electronic Communications Networks and Services) Regulations 2011 (Privacy and Electronic Communications) (e-Privacy Regulations), Regulation (EU) 611/2013 on the measures applicable to the

POLICY	DETECTS
	notification of personal data breaches under Directive 2002/58/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Israel Personal Data Policy	Personal data applicable to Israel's Protection of Privacy (Information Security) Regulations 5767-2017, Protection of Privacy Law 5741-1981, and a general similarity to the European General Data Protection Regulation (GDPR).
Italy Personal Data Policy	Personal data applicable to Italy's Legislative Decree No. 196/2003, which contains the Italian Personal Data Protection Code (Code), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Japan Personal Data Policy	Personal data as defined by Japan's Act on the Protection of Personal Information ("APPI"), and cross-compatibility with European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Kenya Personal Data Policy	Detects personal data as defined by Kenya Data Protection (Compliance & Enforcement) Regulation 2021, Data Protection (Registration of Data Controllers & Data Processors) Regulation 2021, Data Protection (General) Regulations 2021, and a general similarity to the European General Data Protection Regulation (GDPR).
Kingdom of Saudi Arabia Personal Data Policy	Personal data as defined by Kingdom of Saudi Arabia's Personal Data Protection Law (PDPL) (implemented by Royal Decree M/19 of 9/2/1443H (16 September 2021)) and a

POLICY	DETECTS
	general similarity to the European General Data Protection Regulation (GDPR).
Kuwait Personal Data Policy	Personal data as defined by a general similarity to the European General Data Protection Regulation (GDPR).
Latvia Personal Data Policy	Personal data as defined by Latvia's Personal Data Protection Law, Personal Data Act, Law on Electronic Communications, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Lithuania Personal Data Policy	Personal data as defined by Lithuania's Law on Legal Protection of Personal Data ("Data Protection Law"), Law No. I-1374 and X-1444, Data Retention Directive 2006/24/EC, Law on Electronic Communications ("Electronic Communications Law"), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Luxembourg Personal Data Policy	Personal data as defined by Luxembourg's protection of persons with regard to the processing of personal data (Data Protection Law), Regulation (EU) 679/2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Malaysia Personal Data Policy	Personal data as defined by Malaysia Personal Data Protection Act 2010 and a general similarity to the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Malta Personal Data Policy	Personal data as defined by Malta's Data Protection Act, Processing of Personal Data (Electronic Communications Sector) Regulations, Notification and Fees (Data Protection Act) Regulations, Processing of Personal Data (Protection of Minors) Regulations, Data Protection (Processing of Personal Data in the Police Sector) Regulations, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Mexico Personal Data Policy	Personal data as defined by Mexico's Federal Law on the Protection of Personal Data held by Private Parties, Privacy Notice Guidelines (April 2013), Recommendations on Personal Data Security (November 2013), Parameters for Self-Regulation regarding personal data (May 2014), General Law for the Protection of Personal Data in Possession of Obligated Subjects, and a general similarity to the European General Data Protection Regulation (GDPR).
Netherlands Personal Data Policy	Personal data applicable to Netherlands for Dutch Data Protection Act, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
New Zealand Personal Data Policy	Detects personal data as defined by New Zealand's Personal Data Protection Act, Personal Data Law, Law for the Regulation of Information Services on People's Credit History, and a general similarity to the European General Data Protection Regulation (GDPR)
New Zealand Sensitive Data Policy	Detects sensitive data for New Zealand as defined by the European Privacy Directive

POLICY	DETECTS
	95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR)
Nigeria Personal Data Policy	Detects personal data as defined by Nigeria's Data Protection Act of 2023, and cross-compatibility with European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Nigeria Sensitive Data Policy	Detects sensitive data for Nigeria as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR)
Norway Personal Data Policy	Personal data applicable to Norway's Personal Data Act (PDA), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Panama Personal Data Policy	Description Detects personal data as defined by Panama's Law 81 on the Protection of Personal
	Data, and a general similarity to the European General Data Protection Regulation
	(GDPR).
Peru Personal Data Policy	Personal data as defined by Peru Personal Data Protection Law - Law
	N° 29733, Political Constitution of Peru, and a general similarity to the European
	General Data Protection Regulation (GDPR).
Poland Personal Data Policy	Personal data as defined by Poland's Personal Data Protection Act of August 1997 (PDPA),

POLICY	DETECTS
	Telecommunications Act of July 2004, Telecommunications Act of July 2004, Insurance and Reinsurance Activity Act of September 2015, Act on Providing Services by Electronic Means of July 2002, Labour Code of June 1974, the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Portugal Personal Data Policy	Personal data as defined by the Data Protection Act (DPA) for Portuguese Law 67/98 (Data Protection Act), the Portuguese Constitution (Article 35 on the use of computerised data), Law 41/2004 implementing Directive 2002/58/EC on the protection of privacy in the electronic communications sector (E-Privacy Directive), Law 46/2012 implementing Directive 2002/22/EC on universal service and users' rights (Universal Service Directive), Regulation (EU) 611/2013 on the measures applicable to the notification of personal data breaches under Directive 2002/58/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Romania Personal Data Policy	Personal data as defined by Romania's Data Protection Law No. 677/2001 ("DPL"), Romanian E-Privacy Law No. 506/2004, E-Privacy Directive 2002/58/EC, Romanian E-commerce Law No. 365/2002, Electronic Commerce Directive 2000/31/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Russia Personal Data Policy	Personal data applicable to Russia Federal Law on Personal Data (No. 152-FZ), and

POLICY	DETECTS
	Russian Code on Administrative Infractions (No.195-FZ).
Singapore Personal Data Policy	Personal data as defined by Singapore's Personal Data Protection Act (PDPA), numerous laws relating to the processing of personal data in the public sector that apply to everyone, including secrecy and disclosure laws in the Official Secrets Act and the Electronic Transactions Act, and cross-compatibility with European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Slovakia Personal Data Policy	Personal data as defined by Slovakia's Act no. 122/2013 on Personal Data Protection, Act no. 84/2014 on Personal Data Protection, Law no. 136/2014 in Slovak Republic's Collection of Laws, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Slovenia Personal Data Policy	Personal data as defined by Slovenia's Personal Data Protection Act ("ZVOP-1" and "ZVOP-2"), Law no. 86/2004, Law no. 113/05, Law no. 51/07, Law no. 67/07, Law no. 94/07, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
South Africa Personal Data Policy	Personal data applicable to South Africa Protection of Personal Information Act (POPIA / POPI).
South Korea Personal Data Policy	Personal data as defined by South Korea's Personal Information Protection Act (Act No. 16930, Feb. 4, 2020), Enforcement Decree of the Personal Information Protection Act (Presidential Decree No. 30892, Aug. 04,

POLICY	DETECTS
	2020), Act on the Lapse of Criminal Sentences (Act No. 15258, Dec. 19, 2017, Partial Amendment), and a general similarity to the European General Data Protection Regulation (GDPR).
Spain Personal Data Policy	Personal data applicable to Spain's Data Protection Act (Law 15/1999 on the protection of personal data), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Sweden Personal Data Policy	Personal data applicable to Sweden for Swedish Personal Data Act (PDA), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Switzerland Personal Data Policy	Personal data as defined by Switzerland's Federal Act on Data Protection (FADP) and the Ordinance to the Federal Act on Data Protection (OFADP) at the Swiss federal and canton levels, Regulation (EU) 611/2013 on the measures applicable to the notification of personal data breaches under Directive 2002/58/EC, European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Taiwan Personal Data Policy	Personal data as defined by Taiwan Personal Data Protection Act, Enforcement Rules of the Personal Data Protection Act, and a general similarity to the European General Data Protection Regulation (GDPR).
Thailand Personal Data Policy	Personal data as defined by Thailand's Personal Data Protection Act ("PDPA"), and a general similarity to the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Turkey Personal Data Policy	Personal data applicable to Turkey's Law on Protection of Personal Data No 6698 (Data Protection Law, KVKK), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
U.K. Personal Data Policy	Personal data applicable to United Kingdom's Data Protection Act 1998 (DPA), European Privacy Directive 95/46/EC on data protection (Data Protection Directive), and the European General Data Protection Regulation (GDPR).
Ukraine Personal Data Policy	Personal data applicable to Ukraine "On Personal Data Protection" Law ("Data Protection Law") No. 2297 VI, No. 5491-VI, and No. 383-VII.
United States Personal Data Policy	Personal data applicable to United States's Federal and State laws, including The Federal Trade Commission Act (FTC Act), the Financial Services Modernization Act (Gramm-Leach-Bliley Act (GLB/GLBA)), the Health Insurance Portability and Accountability Act (HIPAA), the Fair Credit Reporting Act (FCRA), the Fair and Accurate Credit Transactions Act (FACTA), among many others.
United Arab Emirates (UAE) Personal Data Policy	Personal data as defined by United Arab Emirates (UAE) Article 379 of the UAE Penal Code, Cyber Crime Law (Federal Law 5 of 2012 regarding Information Technology Crime Control) (amended by Federal Law No. 12 of 2016 and Federal Decree-Law No. 2 of 2018), and general similarity to the European General Data Protection Regulation (GDPR).
Venezuela Personal Data Policy	Personal data as defined by Venezuela's Constitution of the Bolivarian Republic

POLICY	DETECTS
	of Venezuela (1998), Venezuela's International Covenant on Civil and Political
	Rights (1978), Venezuela's Act for the Protection of the Privacy of Communications
	(1991), Venezuela's Act for the Protection of Girls, Boys and Adolescents
	(2000), and a general similarity to the European General Data Protection Regulation
	(GDPR).
Vietnam Personal Data Policy	Personal data as defined by Vietnam's Draft Decree on Personal Data Protection in Vietnam dated February 9, 2021, the Constitution of the Socialist Republic of Vietnam 2013, Vietnam's Civil Code 2015, Vietnam's Criminal Code No. 100/2015/QH13 (November 27, 2015), Vietnam's Criminal Procedure Code Law No. 101/2015/QH13 (November 27, 2015), Vietnam's Law on Protection of Consumers' Rights No. 59/2010/QH12 (November 17, 2010), and a general similarity to the European General Data Protection Regulation (GDPR).

Table: Special Category Data

POLICY	DETECTS
Argentina Sensitive Data Policy	Argentina-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Australia Sensitive Data Policy	Australia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on

POLICY	DETECTS
	data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Austria Sensitive Data Policy	Austria-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Belgium Sensitive Data Policy	Belgium-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Bolivia Sensitive Data Policy	Detects sensitive data for Bolivia as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Brazil Sensitive Data Policy	Brazil-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Bulgaria Sensitive Data Policy	Bulgaria-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Canada Sensitive Data Policy	Canada-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Chile Sensitive Data Policy	Chile-specific sensitive data as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
China Sensitive Data Policy	China-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Colombia Sensitive Data Policy	Colombia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Croatia Sensitive Data Policy	Croatia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Cyprus Sensitive Data Policy	Cyprus-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Czech Republic Sensitive Data Policy	Czech Republic-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Denmark Sensitive Data Policy	Denmark-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and

POLICY	DETECTS
	the European General Data Protection Regulation (GDPR).
Egypt Sensitive Data Policy	Egypt-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
El Salvador Sensitive Data Policy	Detects sensitive data for El Salvador as defined by the European Privacy
	Directive 95/46/EC on data protection (Data Protection Directive) and the
	European General Data Protection Regulation (GDPR)
Ecuador Sensitive Data Policy	Ecuador-specific sensitive data as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Estonia Sensitive Data Policy	Estonia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Finland Sensitive Data Policy	Finland-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
France Sensitive Data Policy	France-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and

POLICY	DETECTS
	the European General Data Protection Regulation (GDPR).
Germany Sensitive Data Policy	Detects sensitive data for Germany as defined by the European Privacy
	Directive 95/46/EC on data protection (Data Protection Directive) and the
	European General Data Protection Regulation (GDPR).
Greece Sensitive Data Policy	Greece-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Guatemala Sensitive Data Policy	Detects sensitive data for Guatemala as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Honduras Sensitive Data Policy	Detects sensitive data for Honduras as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Hong Kong Sensitive Data Policy	Hong Kong-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Hungary Sensitive Data Policy	Hungary-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and

POLICY	DETECTS
	the European General Data Protection Regulation (GDPR).
Iceland Sensitive Data Policy	Iceland-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
India Sensitive Data Policy	India-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Indonesia Sensitive Data Policy	Indonesia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Ireland Sensitive Data Policy	Ireland-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Israel Sensitive Data Policy	Israel-sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Italy Sensitive Data Policy	Italy-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Japan Sensitive Data Policy	Japan-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the

POLICY	DETECTS
	European General Data Protection Regulation (GDPR).
Kenya Sensitive Data Policy	Detects sensitive data for Kenya as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Kingdom of Saudi Arabia Sensitive Data Policy	Kingdom of Saudi Arabia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Kuwait Sensitive Data Policy	Kuwait-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Latvia Sensitive Data Policy	Latvia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Lithuania Sensitive Data Policy	Lithuania-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Luxembourg Sensitive Data Policy	Luxembourg-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Malaysia Sensitive Data Policy	Malaysia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Malta Sensitive Data Policy	Malta-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Mexico Sensitive Data Policy	Mexico-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Māori Te Mana Raraunga Policy	Detects text that may be subject to protection under the Māori's (New Zealand indigenous people) Te Mana Raraunga. This policy matches based on word patterns. This policy is provided for content classification purposes only.
Netherlands Sensitive Data Policy	Netherlands-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Norway Sensitive Data Policy	Norway-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Panama Sensitive Data Policy	Detects sensitive data for Panama as defined by the European Privacy Directive

POLICY	DETECTS
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Peru Sensitive Data Policy	Peru-specific sensitive data as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Poland Sensitive Data Policy	Poland-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Portugal Sensitive Data Policy	Portugal-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Romania Sensitive Data Policy	Romania-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Russia Sensitive Data Policy	Russia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Singapore Sensitive Data Policy	Singapore-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and

POLICY	DETECTS
	the European General Data Protection Regulation (GDPR).
Slovakia Sensitive Data Policy	Slovakia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Slovenia Sensitive Data Policy	Slovenia-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
South Africa Sensitive Data Policy	South Africa-specific sensitive data as defined by the South Africa Protection of Personal Information Act (POPIA / POPI), the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
South Korea Sensitive Data Policy	South Korea-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Spain Sensitive Data Policy	Spain-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Sweden Sensitive Data Policy	Sweden-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).

POLICY	DETECTS
Switzerland Sensitive Data Policy	Switzerland-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Taiwan Sensitive Data Policy	Taiwan-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Thailand Sensitive Data Policy	Thailand-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Turkey Sensitive Data Policy	Turkey-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
U.K. Sensitive Data Policy	U.K.-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Ukraine Sensitive Data Policy	Ukraine-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
United States Sensitive Data Policy	U.S.-sensitive data as defined by the California Consumer Privacy Act of 2018 (CCPA).
United Arab Emirates (UAE) Sensitive Data Policy	United Arab Emirates (UAE) specific sensitive data as defined by the European Privacy

POLICY	DETECTS
	Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).
Venezuela Sensitive Data Policy	Venezuela-specific sensitive data as defined by the European Privacy Directive
	95/46/EC on data protection (Data Protection Directive) and the European General
	Data Protection Regulation (GDPR).
Vietnam Sensitive Data Policy	Vietnam-specific sensitive data as defined by the European Privacy Directive 95/46/EC on data protection (Data Protection Directive) and the European General Data Protection Regulation (GDPR).

Table: Transparent

POLICY	DETECTS
Client Concerns Policy	**This policy is rebranded and updated from Customer Complaints
	Policy**
	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision. This policy covers concerns associated to the
	following categories:
	- Communication
	- Employee Error
	- Fees and Commissions

POLICY	DETECTS
	- General Concerns
	- Legal
	- Negative Language
	- Operational
	- Performance and Losses
	- Promises and Guarantees
	- Trade Execution
	- Unauthorized Activity
	In addition, each one of these categories also has a discrete transparent
	policy made from the same keywords found in this policy.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Communication Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Communication
	As a transparent policy, key terms and phrases are shared for verification,

POLICY	DETECTS
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Employee Error Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Employee Error
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Fees and Commissions Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Fees and Commissions
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.

POLICY	DETECTS
Client Concerns - General Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- General Concerns
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Legal Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Legal Concerns
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Negative Language Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:

POLICY	DETECTS
	- Negative Language
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Operational Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Operational Concerns
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Performance and Losses Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Performance and Losses
	As a transparent policy, key terms and phrases are shared for verification,

POLICY	DETECTS
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Promises and Guarantees Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Promises and Guarantees
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Client Concerns - Trade Execution Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Trade Execution
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.

POLICY	DETECTS
Client Concerns - Unauthorized Activity Policy	Detect keywords and use sentiment analysis to identify potential alerts for
	financial supervision.
	This policy covers concerns associated to:
	- Unauthorized Activity
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Communication - 1on1 Policy	Classifies items with only one recipient.
	As a transparent policy, the policy logic is shared for verification, modification, inclusion, or exclusion of additional criteria. Please see online help for more assistance with transparent policies
Communication - Small Group Policy	Classifies items with only one recipient.
	As a transparent policy, the policy logic is shared for verification, modification, inclusion, or exclusion of additional criteria. Please see online help for more assistance with transparent policies.
Communication - Large Group Policy	Classifies items with 10 or more recipients.
	As a transparent policy, the policy logic is shared for verification, modification, inclusion, or exclusion of additional criteria. Please see online help for more assistance with transparent policies.

POLICY	DETECTS
Emoji - Angry or Frustrated Policy	Detects emoji that depict angry or frustrated emotions. As a transparent policy, key terms and phrases are shared for verification, modification, inclusion, or exclusion of additional criteria. Please see online help for more assistance with transparent policies.
Financial Distress Policy	Keywords and phrases used to identify employees that may be
	experiencing financial distress. This policy can be used to aid in meeting
	compliance obligations for financial supervision.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
High Risk Securities - Complete List Policy	Detects high risk securities (names or tickers) in proximity to transaction
	terms to aid in meeting compliance obligations for financial supervision.
	High risk securities are identified as securities that have had a recent US
	securities class action filling (current list covers Jan 1, 2021-Nov 17,
	2022).
	The following were removed due to abundance of false positives observed

POLICY	DETECTS
	in testing Tickers: BCS | CAN | CS | ET | GS | IBM | IQ | MDT | MP |
	OPEN | SAVE | SO | TSP | U | UBER | USB | WFC | XRAY | Z
	And Names: barclays | credit suisse | goldman sachs | international
	business machines | spirit airlines | u.s. bancorp | uber | wells fargo
	The following were identified as potentially causing numerous false
	positive but performed satisfactorily during testing and so remain in the
	policy - Tickers: ACAD | AI | AMZN | APPS | AZN | BLUE | BZ | CEI |
	COIN | DM | DNA | DOCU | EAR | EDU | EH | ERIC | EVA | FAT | GRAB |
	HOOD | JT | LIVE | META | MF | NFLX | OM | PCT | PLUG | PTE | RAD |
	RENT | RIDE | ROOT | SAM | SNAP | SOS | SQ | TALK | TASK | TLS |
	TMC | TSN | UI | UL | VIEW | WDH | WISH | WM | YQ
	And Names: amazon.com incorporated | amazon.com, incorporated |
	block inc | block incorporated | block, inc | block, incoporated | netflix
	incorporated | netflix, incorporated | okta inc | okta incorporated | okta, inc

POLICY	DETECTS
	| okta, incorporated | on24 inc | on24, inc | rivian automotive inc | rivian
	automotive incorporated | rivian automotive, inc | rivian automotive,
	incorporated | twitter inc | twitter inc. | twitter, inc | twitter, inc.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
High Risk Securities - Large CAP Policy	Detects large cap high risk securities (names or tickers) in proximity to transaction terms to aid in meeting compliance obligations for financial
	supervision. High risk securities are identified as securities that have had
	a recent US securities class action filling (current list covers Jan 1, 2021-
	Nov 17, 2022). Securities are separated into Very Large CAP ($100B+)
	and Large CAP ($10B-$100B) for easy editing and/or removal.
	The following were removed due to abundance of false positives observed
	in testing - Tickers: BCS | CS | ET | GS | IBM | MDT | SO | U | UBER |
	USB | WFC

POLICY	DETECTS
	And Names: barclays | credit suisse | goldman sachs | international
	business machines | u.s. bancorp | uber | wells fargo
	The following were identified as potentially causing numerous false
	positive but performed satisfactorily during testing and so remain in the
	policy - Tickers: AMZN | AZN | DOCU | ERIC | GRAB | META | NFLX |
	SNAP | SQ | TSN | UI | UL | WM
	And Names: amazon.com incorporated | amazon.com, incorporated |
	block inc | block incorporated | block, inc | block, incoporated | netflix
	incorporated | netflix, incorporated | okta inc | okta incorporated | okta, inc
	| okta, incorporated | rivian automotive inc | rivian automotive incorporated
	| rivian automotive, inc | rivian automotive, incorporated | twitter inc | twitter
	inc. | twitter, inc | twitter, inc.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.

POLICY	DETECTS
High Risk Securities - Mid CAP Policy	Detects mid cap high risk securities (names or tickers) in proximity to
	transaction terms to aid in meeting compliance obligations for financial
	supervision. Mid cap high risk securities are identified as securities that
	have had a recent US securities class action filling (current list covers Jan
	1, 2021-Nov 17, 2022) and a market cap of $2B-$10B.
	The following were removed due to abundance of false positives observed
	in testing - Tickers: IQ | MP | SAVE | XRAY | Z
	And Name: spirit airlines
	The following were identified as potentially causing numerous false
	positive but performed satisfactorily during testing and so remain in the
	policy - Tickers: ACAD | BZ | COIN | DNA | EDU | EVA | HOOD | PLUG |
	SAM
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.

POLICY	DETECTS
High Risk Securities - Other Policy	Detects high risk securities in proximity to transaction terms to aid in
	meeting compliance obligations for financial supervision. High Risk
	Securities -Other are identified as securities that have had a recent US
	securities class action filling (current list covers Jan 1, 2021-Nov 17, 2022)
	and are currently not listed on an SEC-regulated exchange as of Dec 21,
	2022.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
High Risk Securities - Small CAP Policy	Detects small cap high risk securities (names or tickers) in proximity to
	transaction terms to aid in meeting compliance obligations for financial
	supervision. Small cap high risk securities are identified as securities that
	have had a recent US securities class action filling (current list covers Jan
	1, 2021-Nov 17, 2022) and a market cap of less than $2B.

POLICY	DETECTS
	The following were removed due to abundance of false positives observed
	in testing - Tickers: CAN | OPEN | TSP
	The following were identified as potentially causing numerous false
	positive but performed satisfactorily during testing and so remain in the
	policy - Tickers: AI | APPS | BLUE | CEI | DM | EAR | EH | FAT | JT | LIVE
	| MF | OM | PCT | PTE | RAD | RENT | RIDE | ROOT | SOS | TALK |
	TASK | TLS | TMC | VIEW | WDH | WISH | YQ
	And Names: on24, Inc | on24 Inc
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Outside Business Activities Policy	Key phrases of potential outside activities that may require firm pre-approval or could be prohibited per firm policy. This policy is designed to aid in meeting compliance obligations for financial supervision.
Secrecy Policy	Detects key phrases associated to secrecy: the action of keeping
	something secret or the state of being kept secret. As a transparent policy,

POLICY	DETECTS
	this policy and its patterns can be fully edited. To modify, copy the policy
	and/or the pattern and then edit as needed. It is a best practice to use
	Transparent Policies in conjunction with noise (junk) minimization
	techniques to reduce false positives. Please see online help for more
	assistance with transparent policies.
Selling Away Policy	Detects language that may indicate selling away, the offering or obtaining
	of financial products for a client that are not approved by the firm. This
	policy is designed to aid in meeting compliance obligations for financial
	supervision.
	This policy contains the following editable conditions:
	- Identifying a named outside firm in proximity to terms referencing the selling of financial products
	- Identifying the 'hand delivery' of a receipt, check, certificate, or confirm or the delivery of a 'certificate of authenticity' or 'private receipt'
	- Identifying ACH in proximity to terms such as 'my private bank account
	- Identifying outbound email referencing the sending of a payment to a personal address, or financial institution.

POLICY	DETECTS
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.
Subscriptions Policy	Detects words or phrases that are often associated with subscription
	based electronic communications. This policy may be used to isolate and
	exclude such communications from the institution's review sets.
	As a transparent policy, key terms and phrases are shared for verification,
	modification, inclusion, or exclusion of additional criteria. Please see
	online help for more assistance with transparent policies.

Table: U.S. Federal Regulations

POLICY	DETECTS
Criminal History Policy	Identity history summary (Criminal History Record, or Rap sheet) for Criminal Justice Systems.
United States Delaware Personal Data Privacy Act (DPDPA) Policy	Detects U.S.- and Delaware-centric personally identifiable information in support of Delaware Personal Data Privacy Act (DPDPA).

POLICY	DETECTS
United States Federal Financial Institutions Examination Council (FFIEC) Policy	Personally identifiable and financial information for the Federal Financial Institutions Examination Council (FFIEC).
United States Export Controlled Information Policy	Content subject to export control under the U.S. International Traffic in Arms Regulations (ITAR), the Arms Export Control Act, the Export Administration Act, export-controlled materials under U.S. Department of Commerce (DoC) Export Administration Regulations (EAR) 15 CFR §§730-774, the Export Administration Regulations (EAR), U.S. Executive Order 13556 Controlled Unclassified Information (CUI) 32 CFR Part 2002, or articles, services, and related technical data designated as defense articles or defense services pursuant to sections 38 and 47(7) of the Arms Export Control Act and constitute the U.S. Munitions List (USML).
United States Federal Information Security Management Act (FISMA) Policy	Personal and security information and for Federal Information Security Management Act (FISMA) of 2002, also known as E-Government Act of 2002, also known as Public Law 107-347. The type of information that should be protected depends on the relevant agency/sector.
United States Florida Senate Bill 264 (SB 264) Policy	Detects electronic patient information for Florida's medical privacy law and medical identity theft to support Florida Senate Bill 264 (Forida/FL SB 264).
United States Indiana Consumer Data Protection Act (ICDPA) Policy	Detects U.S.- and Indiana-centric personally identifiable information in support of Indiana Consumer Data Protection Act (ICDPA/ INCDPA)
United States Iowa Consumer Data Protection Act (ICDPA) Policy	Detects U.S.- and Iowa-centric personally identifiable information in support of Iowa

POLICY	DETECTS
	Consumer Data Protection Act (ICDPA/ IACDPA)
United States Kentucky Consumer Data Protection Act (CDPA) (HB 15) Policy	Detects U.S.- and Kentucky-centric personally identifiable information in support of Kentucky Consumer Data Protection Act (CDPA) (HB 15).
United States Montana Consumer Data Protection Act (MCDPA) Policy	Detects U.S.- and Montana-centric personally identifiable information in support of Montana Consumer Data Privacy Act (MCDPA/ MTCDPA).
United States National Institute of Standards and Technology (NIST) 800-171 Policy	Detect forms and terms relating to the U.S. National Institute of Standards and Technology (NIST) 800-171 publication.
United States New Hampshire Consumer Data Privacy Act (CDPA) Policy	Detects U.S.- and New Hampshire-centric personally identifiable information in support of New Hampshire Consumer Data Privacy Act (CDPA) (SB 255)
United States New Jersey Data Privacy Law (DPL) Policy	Detects U.S.- and New Jersey-centric personally identifiable information in support of New Jersey Data Privacy Law (NJDPL) (SB 332)
United States Oregon Consumer Privacy Act (CPA) Policy	Detects U.S.- and Oregon-centric personally identifiable information in support of Oregon Consumer Privacy Act (CPA).
United States Internal Revenue Service (IRS) 1075 Policy	IRS tax forms and financial information for the U.S. Internal Revenue Service (IRS) 1075 and Internal Revenue Code (IRC) 6103.
United States Securities and Exchange Commission (SEC) Forms Policy	U.S. Securities and Exchange Commission (SEC) Forms.
United States International Traffic in Arms Regulations (ITAR) Policy	Content subject to export control under the International Traffic in Arms Regulations (ITAR), the Arms Export Control Act, the Export Administration Act, the Export Administration

POLICY	DETECTS
	Regulations (EAR), or articles, services, and related technical data designated as defense articles or defense services pursuant to sections 38 and 47(7) of the Arms Export Control Act and constitute the U.S. Munitions List (USML).
United States Tennessee Information Protection Act (TIPA) Policy	Detects U.S.- and Tennessee-centric personally identifiable information in support of Tennessee Information Protection Act (TIPA)
United States Texas Data Privacy and Security Act (TDPSA) Policy	Detects U.S.- and Texas-centric personally identifiable information in support of Texas Data Privacy and Security Act (TDPSA).

Table: U.S. State Regulations

POLICY	DETECTS
California Assembly Bill 1298 (HIPAA) Policy	Electronic patient personally identifiable information (PII) for California Assembly Bill 1298 (California/CA AB 1298).
United States California Consumer Privacy Act (CCPA) and California Privacy Rights Act (CPRA) Policy	U.S.- and California- centric personally identifiable information in support of California AB-375, California Consumer Privacy Act (CCPA), and California Privacy Rights Act (CPRA).
United States California Privacy Rights Act (CPRA) Sensitive Personal Information (SPI) Policy	U.S.- and California- centric sensitive personal information (SPI) in support of U.S. California Privacy Rights Act (CPRA) of 2020.
California Financial Information Privacy Act (SB1) Policy	Personal financial information for California Financial Information Privacy Act, also known as CA SB1.
United States Maryland Online Data Privacy Act (MODPA) and Maryland	Detects U.S.- and Maryland-centric personally identifiable information in

POLICY	DETECTS
Personal Information Protection Act (PIPA) Policy	support of Maryland Online Data Privacy Act (MODPA) (HB 567) and
	Maryland Personal Information Protection Act (PIPA) (Law 14-3504)
United States Massachusetts Regulation 201 CMR 17.00 (MA 201 CMR 17) Policy	U.S.- and Massachusetts- centric personally identifiable information in accordance with Massachusetts Regulation 201 CMR 17.00.
United States Minnesota Consumer Data Privacy Act (CDPA / "Minnesota	Detects U.S.- and Minnesota-centric personally identifiable information in
Act") Policy	support of Minnesota Consumer Data Privacy Act (CDPA / "Minnesota Act").
United States Nebraska Data Privacy Act (NDPA) (LB 1074) Policy	Detects U.S.- and Nebraska-centric personally identifiable information in
	support of Nebraska Data Privacy Act (NDPA) (LB 1074)
United States New York State Department of Financial Services (NYDFS) Cybersecurity Regulation 23 NYCRR 500 Policy	Personal data applicable to the U.S. Cybersecurity Regulation; detects personal data applicable to the New York State Department of Financial Services (NYDFS) Cybersecurity Regulation , 23 NYCRR 500.
United States Utah Consumer Privacy Act (UCPA) Policy	Detects U.S.- and Utah-centric personally identifiable information in support of Utah
	Consumer Privacy Act (UCPA).
United States Connecticut Data Privacy Act (CDPA) Policy	Detects U.S.- and Connecticut-centric personally identifiable information in support
	of Connecticut Data Privacy Act (CDPA).
United States Colorado Privacy Act (CPA) Policy	Detects U.S.- and Colorado-centric personally identifiable information in support of
	Colorado Privacy Act (CPA).

POLICY	DETECTS
United States Virginia Consumer Data Protection Act (VCDPA) Policy	Detects U.S.- and Virginia-centric personally identifiable information in support of
	Virginia Consumer Data Protection Act (VCDPA).

Creating policies

Creating policies

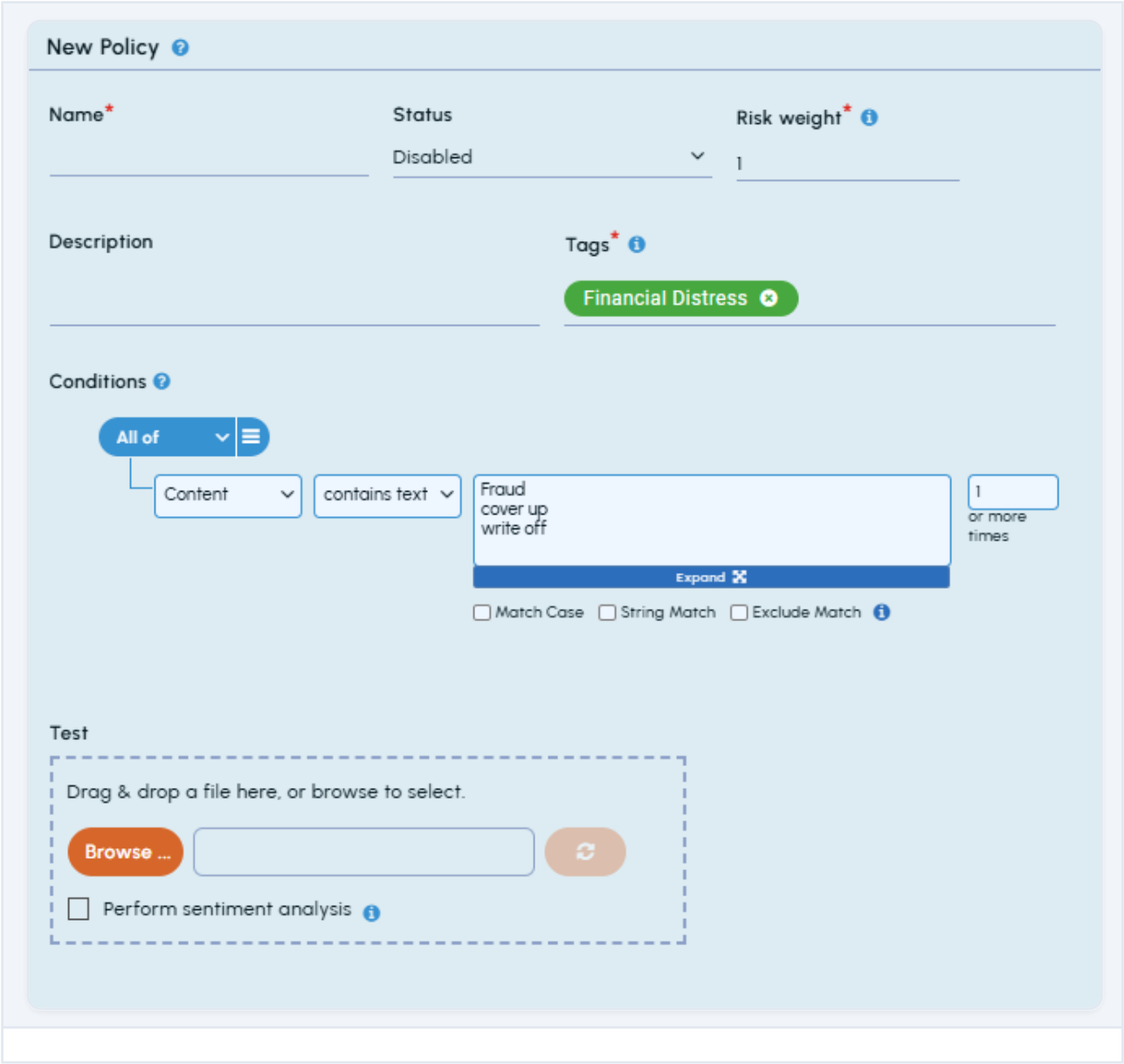
The Arctera Insight Classification comes with a large number of built-in policies, but you can create custom policies if the built-in ones do not meet your needs.

You can also edit existing policies. However, in the case of the built-in policies, the changes that you can make are quite limited.

To create a new policy

1. In the left navigation pane, click **Policies**.
2. Click **New**.

The **New Policy** dialog box appears.

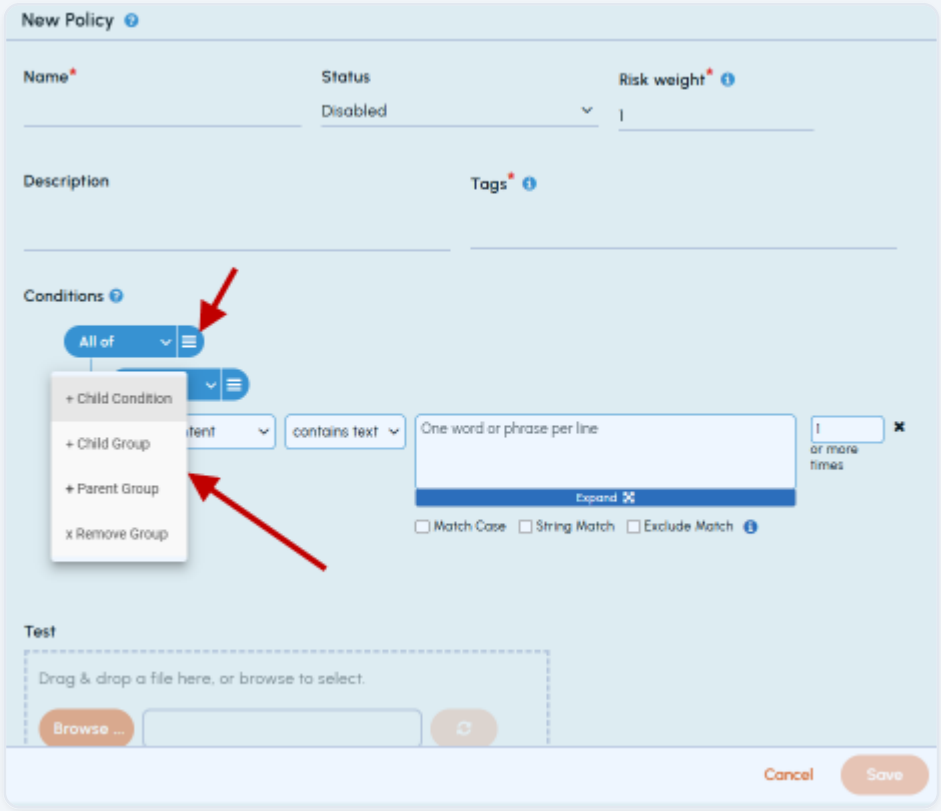


1. Specify the following details:

NAME	SPECIFIES THE POLICY NAME. THE NAME MUST BE UNIQUE, AND IT CAN
	CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
Status	Enables or disables the policy. You must enable the policy if you want the
	Arctera Insight Classification to check for and tag the items that match the policy.

NAME	SPECIFIES THE POLICY NAME. THE NAME MUST BE UNIQUE, AND IT CAN
	CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
Description	(Optional.) Provides a short description of the policy for display
	in the Arctera Insight Classification.
Risk weight	Specify the risk weight for the policy. This is a mandatory field.
	By default, the risk weight value of all the custom policies and most of the built-in policies is configured as 1. Users can modify the risk weight value in the range of 0 to 10.
Tags	Nominates one or more tags that you want to apply to the items that match the policy conditions. Click the Tags field to choose from a list of the available tags.
Conditions	Specifies one or more conditions that an item must meet for the Arctera Insight Classification to consider it a match.
	Click action menu and click: + Child Condition to add a new condition for this policy.
	+ Child Group to add a new group of conditions for this policy. You can add maximum 10 child groups in one hierarchy.
	+ Parent Group to add a new parent group to the existing group of conditions. You can add maximum 10 parent groups in one hierarchy.
	X Remove Group to remove a specific group.
	Note:

NAME	SPECIFIES THE POLICY NAME. THE NAME MUST BE UNIQUE, AND IT CAN
	CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
	- If the tree structure reaches a predefined maximum depth of 10 , adding a parent node will be disabled.
	- A parent can only be added up to 10 nested levels in condition tree to any group in hierarchy.
	- You can only add parent to a group and not to a condition
	- A group can only be removed if it has only ONE child .
	See About policy conditions .



2. To test the policy that you are creating, under **Test**, click**Browse** and then select an item that ought to match it.

    - To test the same document again, click the **Refresh** icon.
    - To perform sentiment analysis on the selected item (to determine whether the sentiment associated with the item is positive, negative, or neutral), select the **Perform sentiment analysis** check box. If this check box is selected, a sentiment score tag is displayed to show the sentiment analysis score. The sentiment analysis score details are displayed if the policy has a condition associated with the sentiment, and the item fulfills this criteria.
    - To extract information from images and perform classification using Optical Character Recognition (OCR), select the **Include text in images** check box. It extracts the English language text.

      > “ ”

      Note: TheInclude text in imagescheck box is displayed only when the Tesseract software is installed on the system where Arctera Insight Classification is running.

      > “ ”

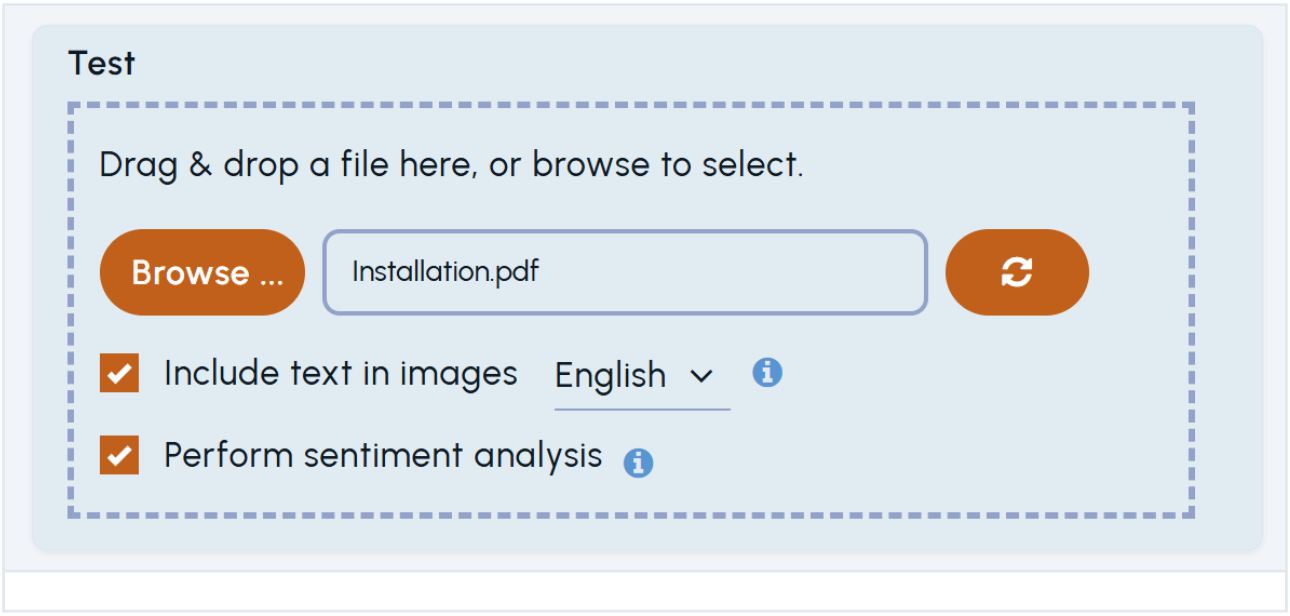
Limitations

Optical Character Recognition processing rate for PDF files is considerably slow.

In addition to this, OCR will not extract information from:

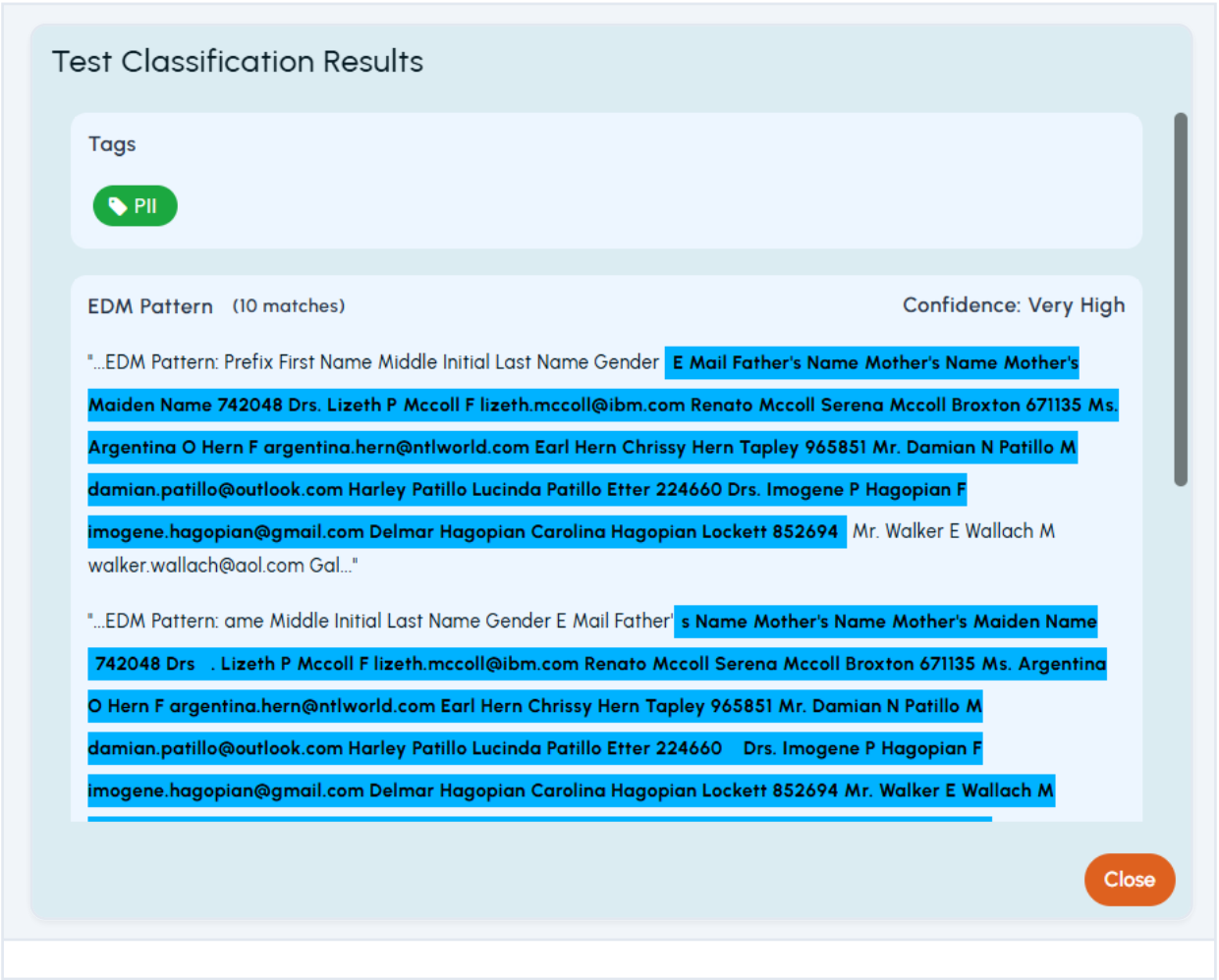
- images that are rotated for 45 degree.
- a single *tif*or*gif* file having multiple pages.
- handwritten image.
- passport image.

Refer to the test functionality results of the sample file for better understanding.

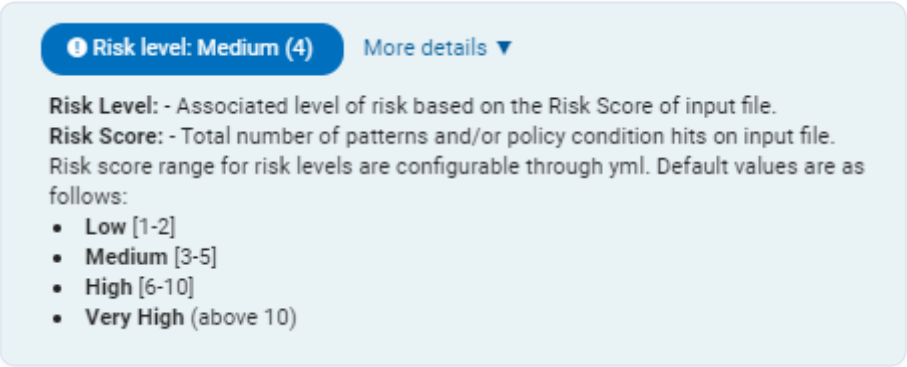


After finding a match, do the following:

- Click **Show details** to see the matching text and confidence levels as shown in the following sample image.



- Click **More details** to see the risk level and risk score of the document as shown in the following sample image.



1. After testing the policy, click **Save**.

About policy conditions

A condition specifies the criteria that an item must meet for the Arctera Insight Classification to consider it a match. Your policies can contain any number of conditions.

This topic provides information on the following:

- Basic components of a condition
- Custom fields
    - Create new property by using custom property fields
    - Create a new property by using YAML configuration file
- Text matches
- Emoji Support
- Variable support
- Regular expression matches
- Author and Recipient department based policy condition
- Pattern matches
- Exact Data Matches
- Language matches
- Entity matches
- Sentiment score
- Risk score and risk level information on classification

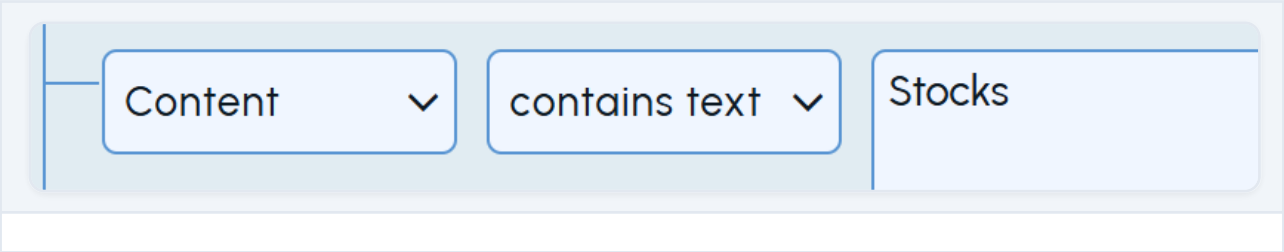
- Condition groups
- []

Basic components of a condition

All conditions have this basic form:

property operator value

For example, in the following condition, "Content" is the property, "contains text" is the operator, and "Stocks" is the value:



The property specifies the part or characteristic of an item that you want to evaluate: its content, title, modified date, file size, and so on. When you choose a property from the list, the options in the two other fields change to suit it. For example, if you choose the "Modified date" property, the other fields provide options with which you can set one or more dates. For properties such as "Content" the available operators are as follows:

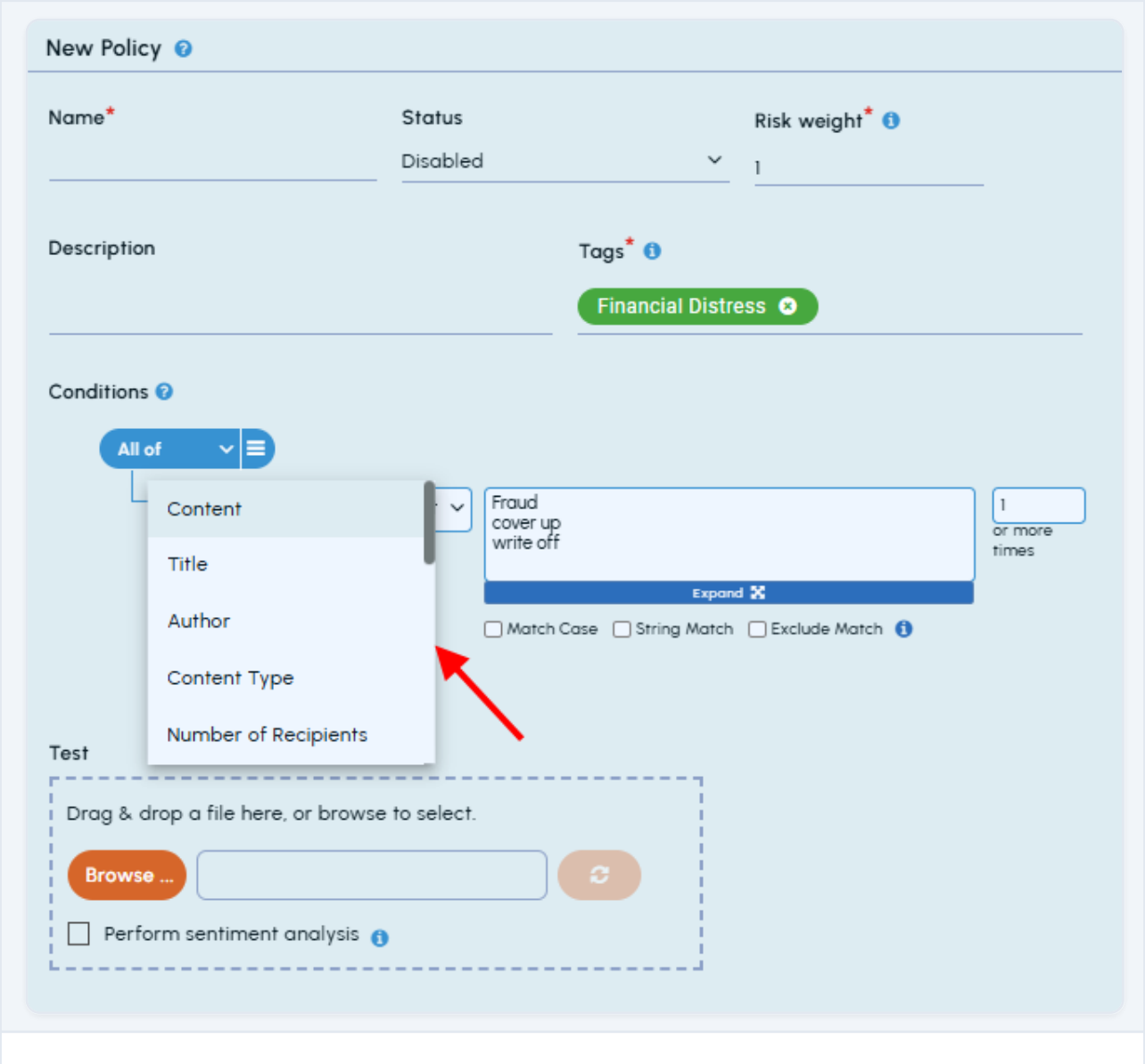
- contains text
- matches regex
- department based policy conditions
- matches pattern
- is similar to
- contains exact data match in
- language is
- contains entity
- sentiment score

At the right of each condition, you can specify the minimum number of times that an item must meet the criteria for the Arctera Insight Classification to consider it a match.

Custom fields

Various applications that you use in your organization may add custom property information to the items that you want to classify. For example, when Enterprise Vault processes an item, it populates a number of the item's metadata properties with information and stores this information with the archived item: the date on which Enterprise Vault archived the item, the number of attachments that it has, and so on.

If you know the name of a property that particularly interests you, you can enter it as a custom field in your policy conditions.

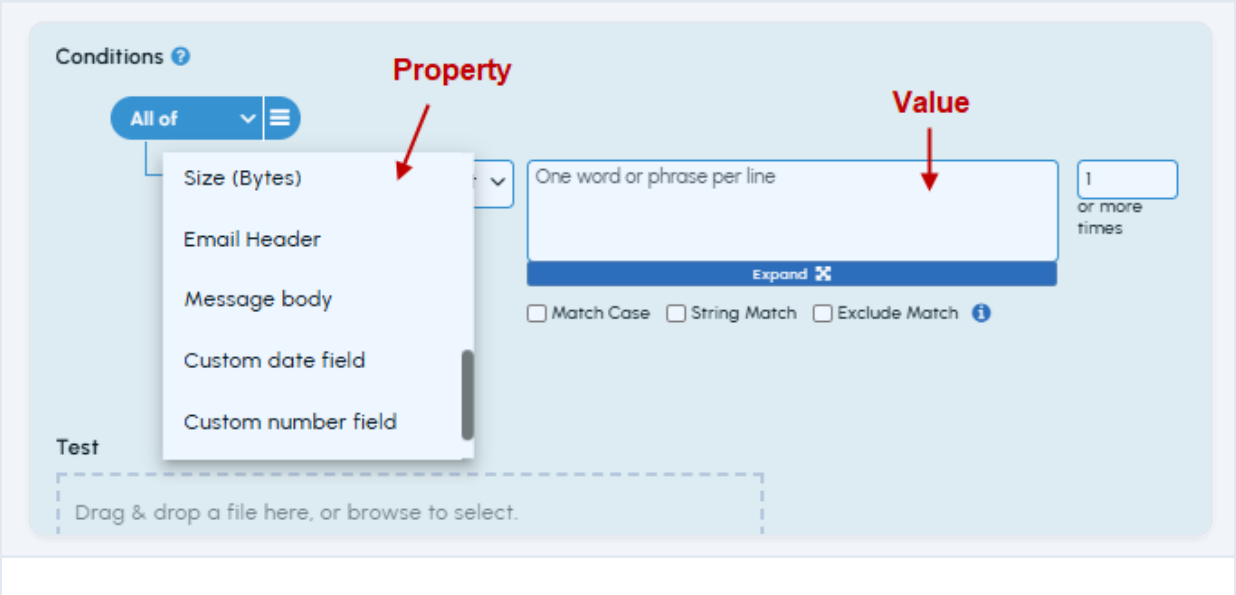


Create new property by using custom property fields

While creating a policy if a required property is not available in the property list, you can create a new property by using custom property fields.

To create a new property, use custom property fields while creating or editing a policy as follows:

1. Set the other fields as per steps given in the topic See Creating policies.
2. Under Conditions section, from the Property drop-down list, select a required custom property field: Custom date field, Custom number field, or Custom string field.

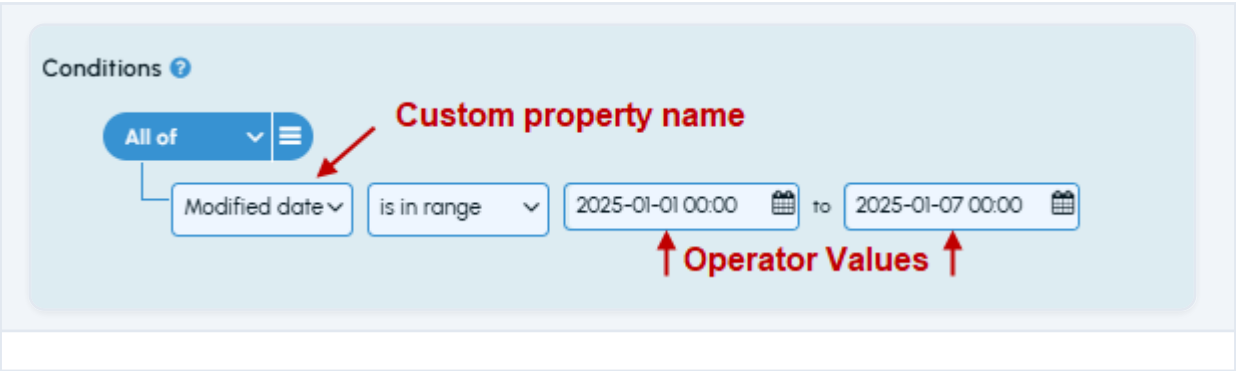


3. Specify the name for the new custom property.

> “ ”

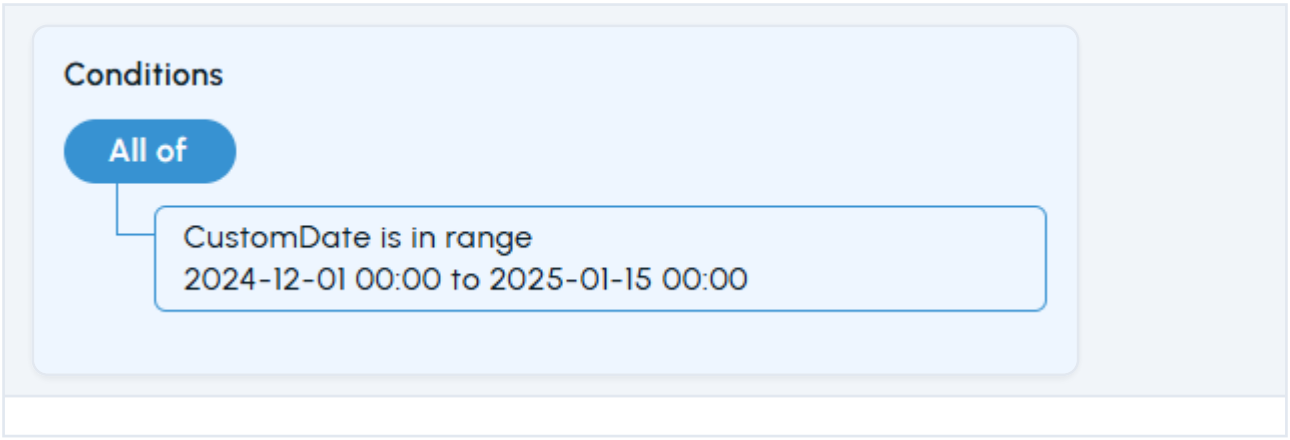
Note: Custom property name must be same as the metadata property name as identified by text extraction engine, for example Apache TIKA. In case of Arctera Enterprise Vault, custom property name must match with one of the indexing properties.

> “ ”



4. Complete the rest of the steps to create a policy.

The new policy is created with a new custom property.



> “ ”

Note: You can add up to 10 group condition levels while creating a new policy.

> “ ”

Create a new property by using YAML configuration file

Use the Arctera Insight Classification's YAML file to add a custom property under the property list on the UI.

The *metadataDefinitions* section of YAML file lists all the existing properties in the property list as follows:

```
metadataDefinitions:
  - name: vic-content    # This is a special name that means 'document body' to VIC, so do not change.  Ok to change displayName
    displayName: Content
    type: String
  - name: title
    displayName: Title
    type: String
    aliases: [subj, 'dc:title']
  - name: auth
    displayName: Author
    type: String
    aliases: ['dc:creator']
  - name: type
    displayName: Content Type
    type: String
    aliases: [dtyp, Content-Type]
  - name: recp
    displayName: Recipients
    type: String
  - name: mdat
    displayName: Modified date
    type: Datetime
    aliases: [last_modified, 'dcterms:modified']
  - name: cdat
    displayName: Creation date
    type: Datetime
    aliases: ['dcterms:created']
  - name: sens
    displayName: Sensitivity
    type: Number
```

The following table shows the data structure for an existing property:

PROPERTY ITEM	DESCRIPTION
name	Specifies the metadata property recognized by the text extractor engine like Apache TIKA.
	In case of Arctera Enterprise Vault, specify the indexing properties captured.
displayName	Name of the property as displayed in the property list on the UI, for example "Title".
type	Associated property type, for example String, Datetime, or Number.
aliases	Specifies the additional metadata properties to be mapped to displayName.

To make this property available in UI under policy condition page

1. Add the new property details as shown in the previous table to the *metadataDefinitions* section in YAML.
2. Restart the Arctera Insight Classification service of respective application.

> “ ”

Note: For updating tenant level settings, reach out toArctera ops team

> “ ”

Text matches

Observe the following guidelines when you set up a condition to look for specific words or phrases in the items that you submit for classification:

- The condition can look for multiple words or phrases, if you place each one on a line of its own. An item needs to contain just one word or phrase in the list to meet the condition.
- Select Match Case to find only exact matches for the uppercase and lowercase characters in the specified words or phrases.

- Select String Match to find instances where the specified words or phrases are contained within other ones. For example, if you select this option, the word `enter` **matches *enters*, *entertainment* and *carpenter*. If you clear the option,** `enter` matches only *enter*.

Similarly, if you select String Match, the phrase `call me` matches *call media* and *recall meeting*, but not *surgically mend*.

Leading and trailing spaces are now considered as part of string match conditions in both policies and patterns. It means if there is a space before first character or after the last character of the policy condition, it will still be considered as part of the evaluation criteria.

For example, if you want to generate a match for the term so sorry, good hits will be *we are so sorry about this* or *I am so sorry this happened to you* but without leading and trailing space there is a possibility to generate false positive like *alfonso sorry about that - no biggie!*. Inclusion of leading and trailing spaces will help in avoiding such false positives and you can make use of a space character before and after to classify only so sorry.

- You can place the proximity operators NEAR and BEFORE between two words in the same line. For example, `tax NEAR/10 reform` **matches instances where there are no more than ten words between *tax* and *reform*.** `sales BEFORE/5 report` matches instances where *sales* precedes *report* and there are no more than five words between them. The number is mandatory in both cases.

  > “ ”

  Note: These proximity operators may not work as expected when evaluating formatted data, such as tables and spreadsheets. The conversion process that this data undergoes before it is classified can swap the order of the table cells. For example, suppose that a spreadsheet contains the word *sales* in one cell and *report* in the cell immediately to the right. This should match the operator *sales BEFORE/5 report* but may not do so after the spreadsheet has been converted, because the conversion process has transposed the two words.

  > “ ”

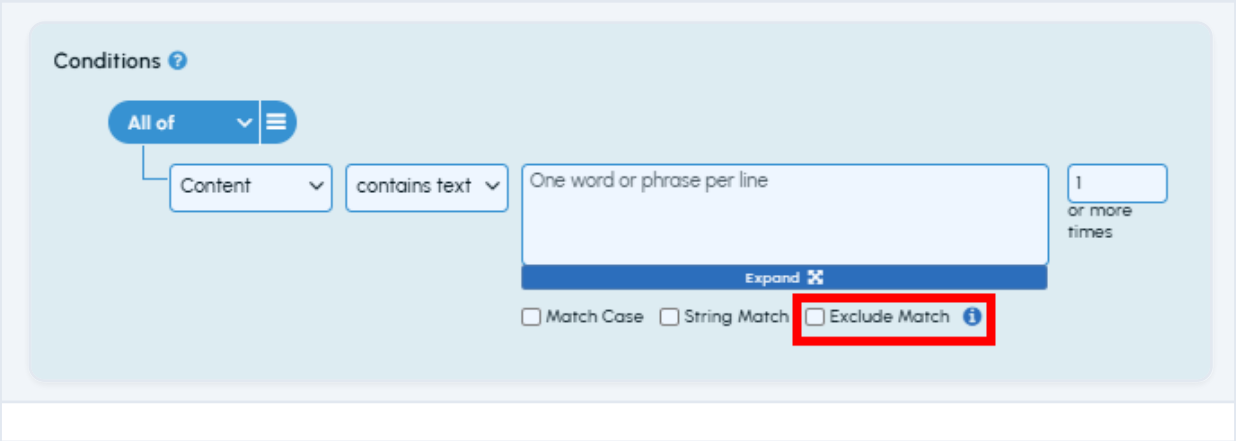
- Word and phrases can include the asterisk (*) and question mark (?) wildcard characters. As part of a word, an asterisk matches zero or more characters. On its own, the asterisk matches exactly one word. A question mark matches exactly one character. For example:

  - `stock*` matches *stock*, *stocks*, and *stockings*.

- `*ock` matches *stock*and*clock*.
- `*ock*` matches *stock*and*clocks*.
- `??ock` matches *stock*and*clock*, but not*dock*.
- `sell * stock` matches *sell the stock*and*sell some stock*, but not*sell stock*.

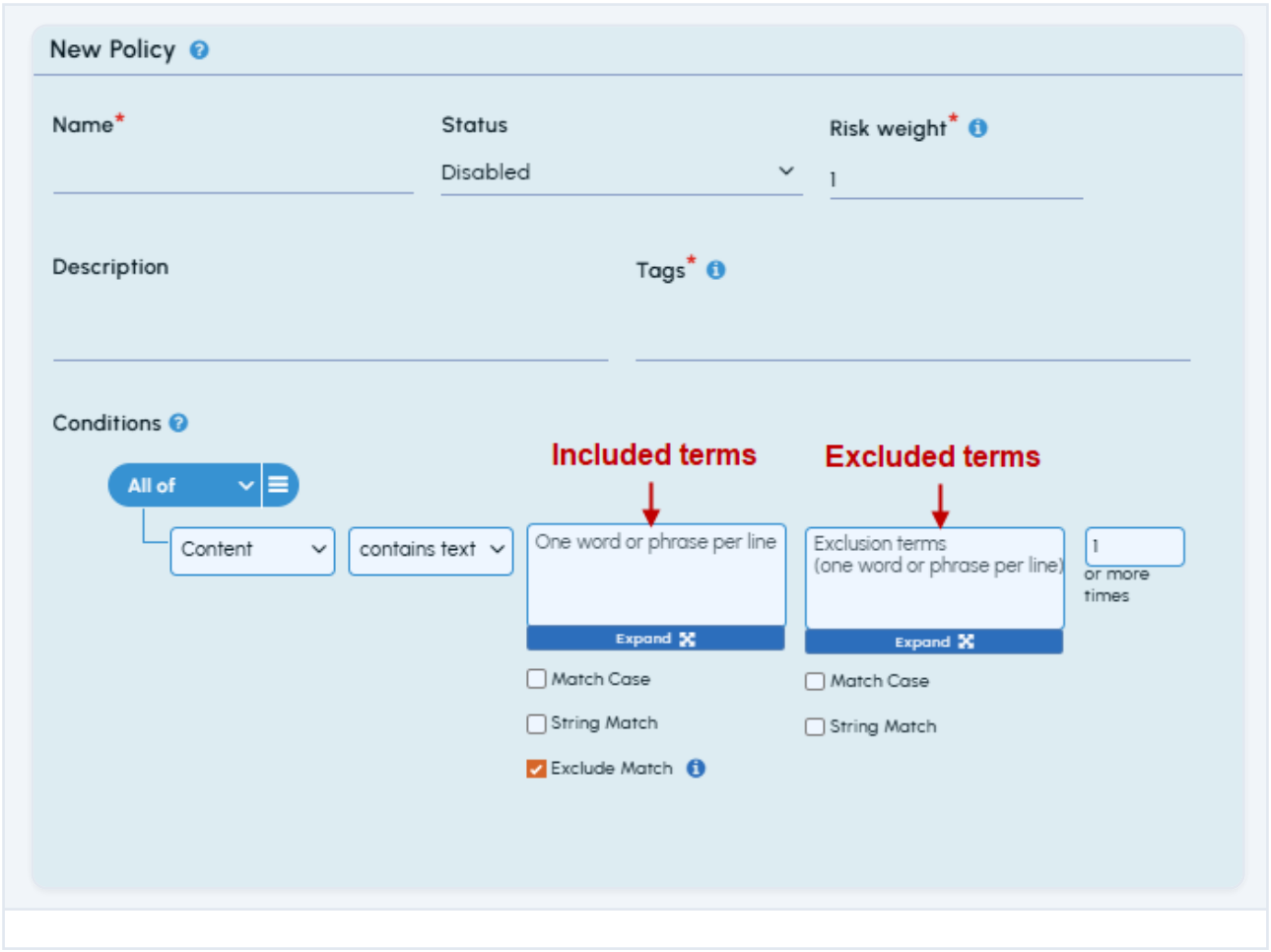
You can use wildcards in combination with the NEAR and BEFORE operators. For example:

- `s?l? BEFORE/1 stock*` matches *sold the stock*,*sell stocks*, and*sale of stockings*.
- Select Exclude Match if you want to exclude specific words or phrases while evaluating the policy condition criteria.



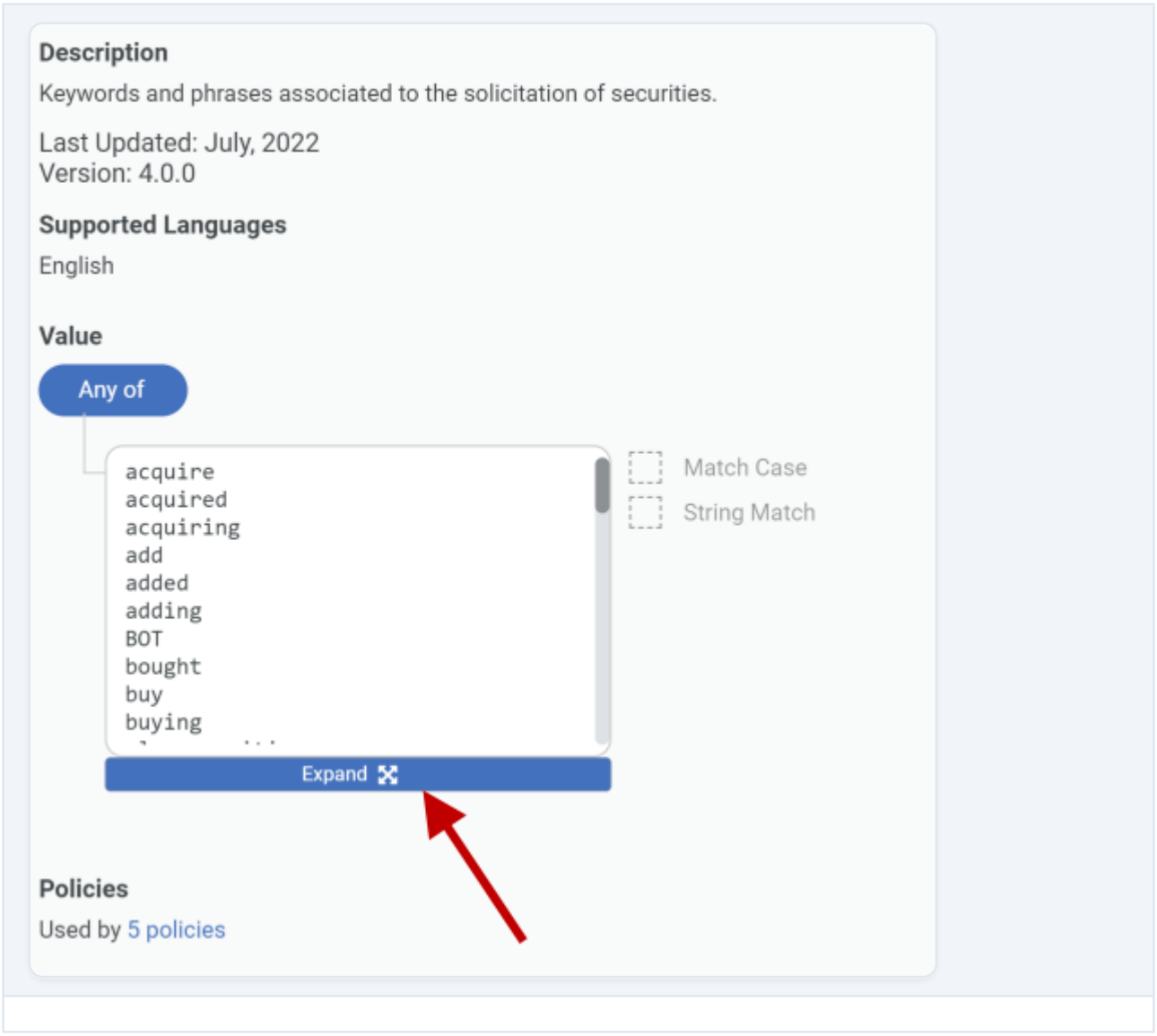
When you select this option, along with the inclusion terms, you can also define the terms that you want to exclude from the matching criteria.

For example, assume that a document contains a sample text "Admin: There is a spoofing activity detected. Bob: Can you help me locate a spoofed account for spoofed email account?". You want to hit on the terms "*spoof, spoofed, spoofing*" only, and want to avoid or exclude the terms*spoofed email account, an email spoof*. In such a scenario, you can provide the keywords "*spoof, spoofed, spoofing*" in the inclusion terms field, and provide the terms "*spoofed email account, an email spoof* " in the exclusion terms field as shown in the sample image below.

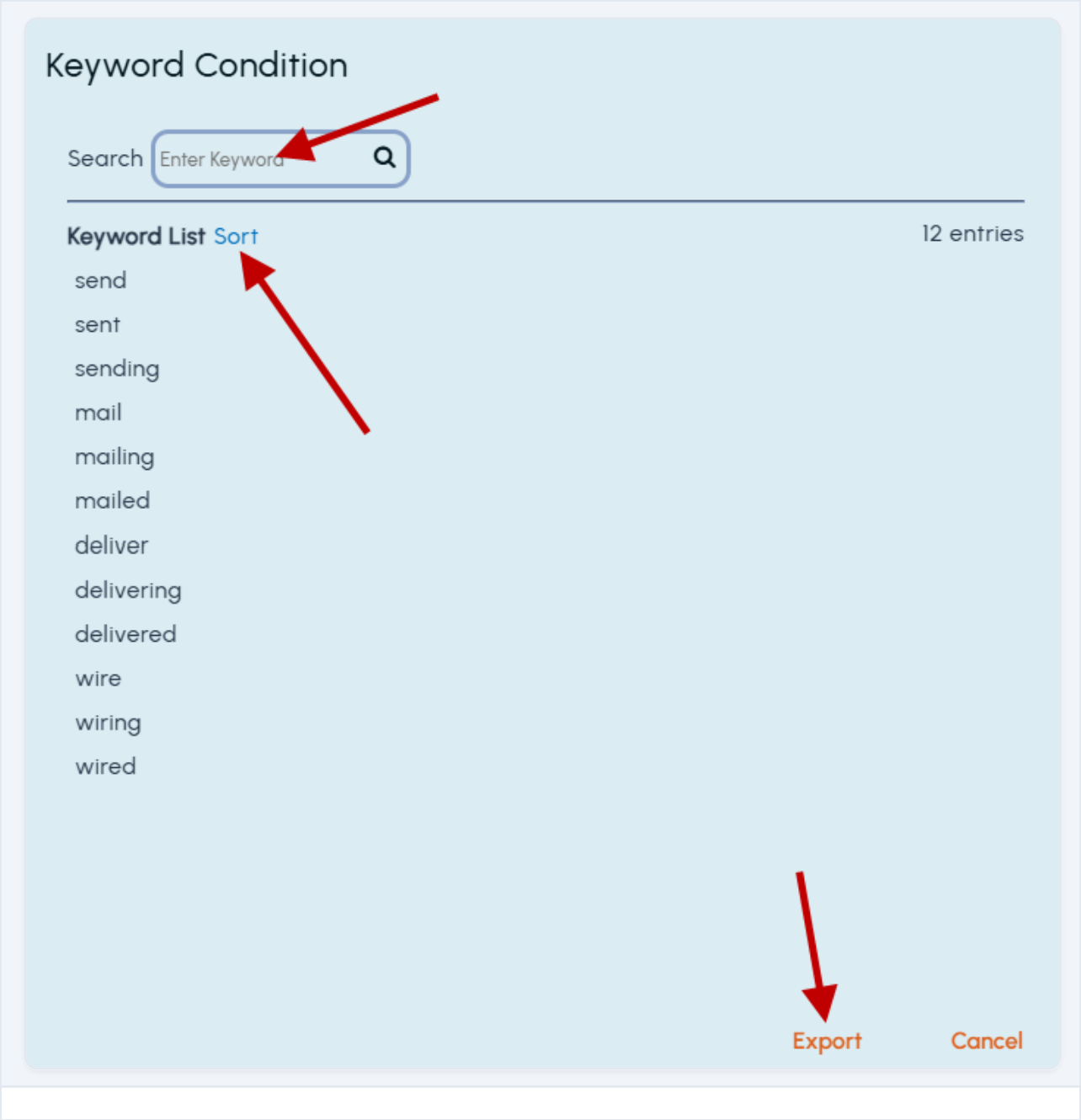


Keyword explorer

In case the list of keyword is long, you can manage it by clicking Expand at the bottom of the list.

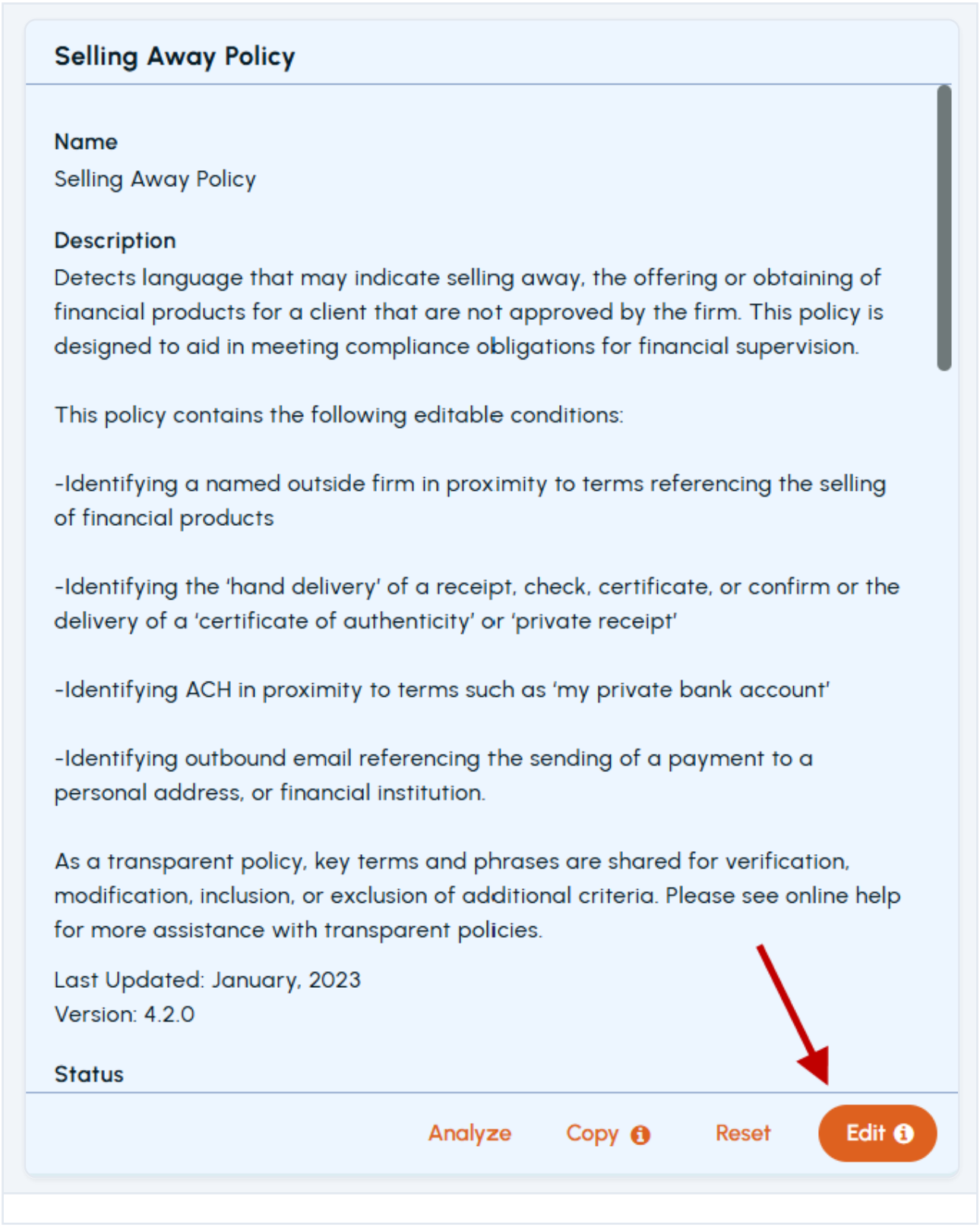


A new pop-up with list of all keywords appears on the screen.



You can search and sort this list by using the search box or clicking Sort. You can also export the list in CSV format by clicking Export.

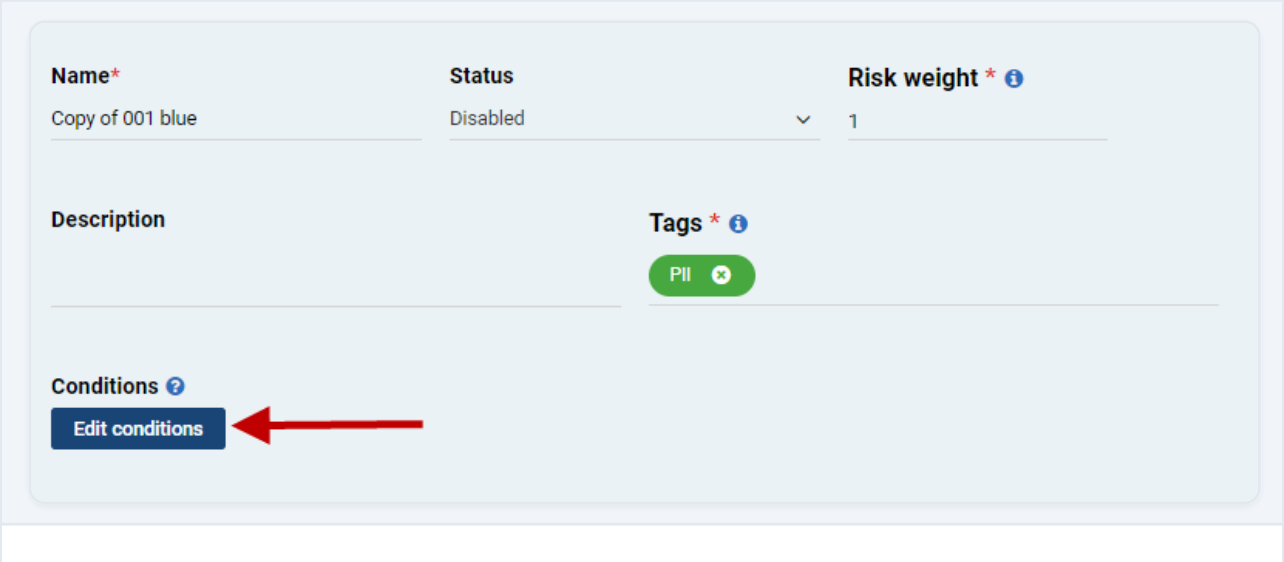
If you want to add or import list of keywords, click Edit.



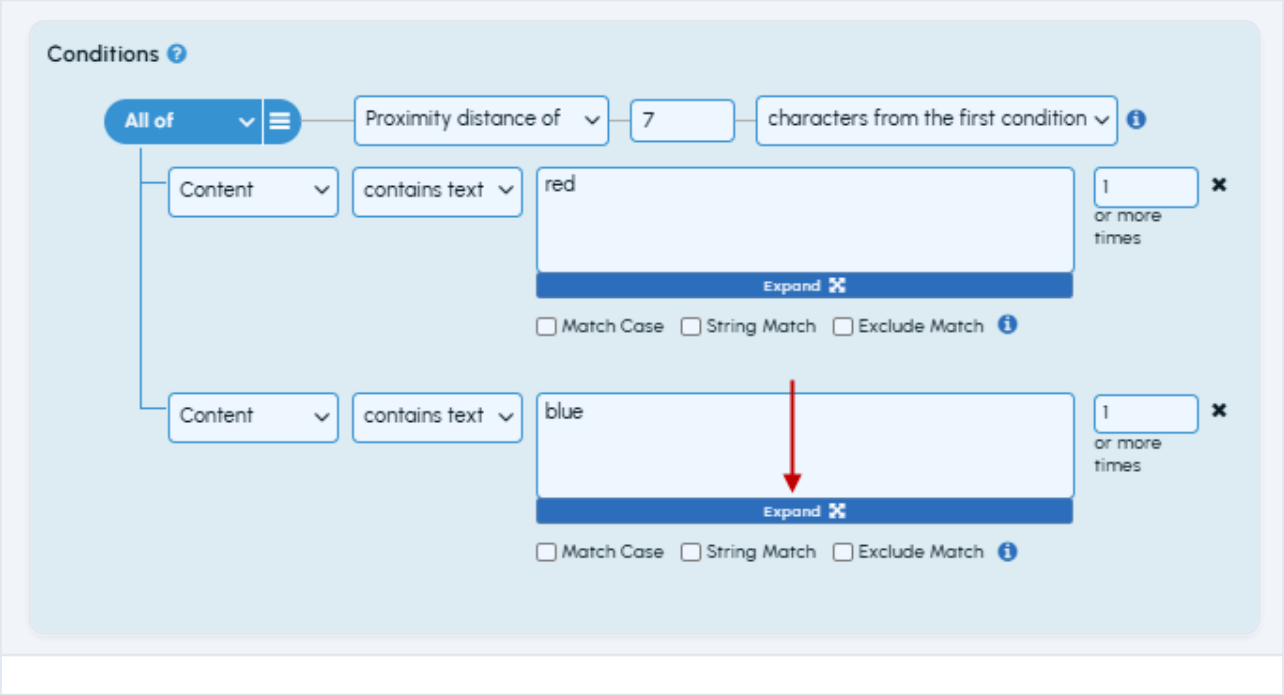
- For *built-in policy* , click Edit to
  - enable/disable status
  - change tags, risk weight, confidence levels
  - adjust minimum match counts

- For *custom policy* , click Edit to
  - enable/disable status
  - change policy name and description, tag, risk, and weight
  - update policy conditions
- Click Copy to create a fully editable version of the existing built-in or custom policy.

In the right pane, click Edit Conditions.



Click Expand in the Conditions section.



On the Keyword Condition pop-up, you can perform following actions.

- Add new keyword
- Search a keyword in the existing list.
- Import keywords in CSV format.
- Export keyword list in CSV format.

> “ ”

Note: Keyword Condition can containmaximum 11,000keywords. For example, if you already have 2000 keywords in the condition, you can add 9,000 more keywords only.

> “ ”

Using Special Standalone Characters in Keyword-Based Policies

When creating keyword-based conditions in policies, it is important to understand how the application handles special standalone characters like @, #, or $.

- Special standalone characters are not allowed by default in keyword conditions.
- If such characters are entered, the system will display a warning message.
- The policy can still be saved, but any standalone special characters will be automatically removed in the back-end, unless the String Match option is explicitly selected.

To Retain Special Characters

If you want to retain standalone special characters in your keyword conditions, you must select the String Match option. This setting ensures that the characters are preserved exactly as entered and bypasses the automatic cleanup process.

Scenarios and Behavior

SCENARIOS	BEHAVIOR
Policy with keyword condition containing special characters	Warning message is shown. Special characters will be removed unless String Match is selected.
Policy with keyword condition and String Match selected	Warning message is not shown. Special characters are retained.

Limitations of the exclusion policy condition

- This field allows only keywords based exclusion. It means the input field for exclusion terms will only accept keywords, and not regular expression, pattern, and so on.
- Keyword based exclusion works only for the scenario where every inclusion term is completely contained within an exclusion term.
- The group-level condition proximity option will not be available for a group if any of the underlying conditions has exclusions.

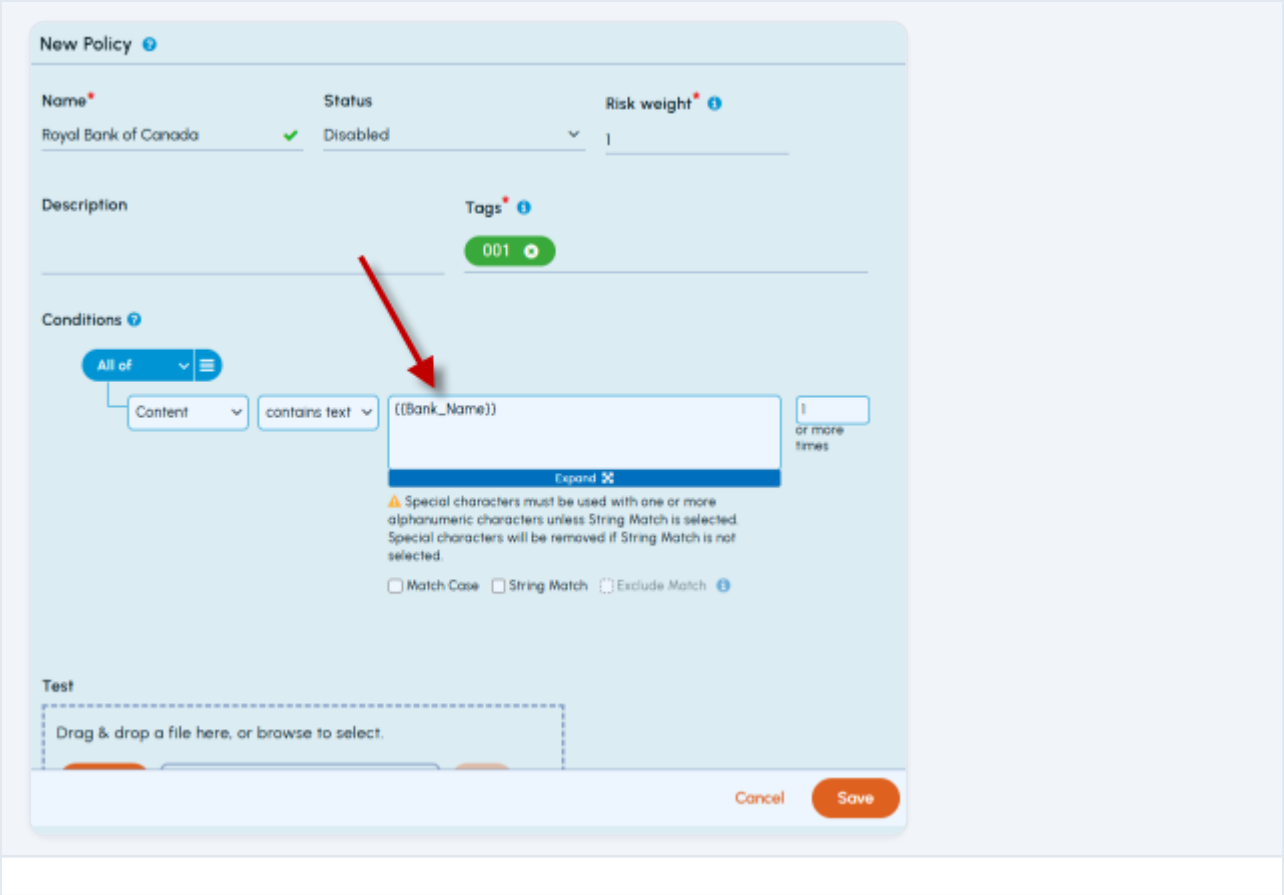
To use the exclusion policy condition, See Using a keywords-based exclusion policy condition.

Emoji Support

Classification of emoji based policy conditions is supported. A policy can be created which has a condition containing emojis and the content can be classified.

Variable support

Variable Support allows you to insert dynamic values into text conditions across policies and patterns using simple placeholders such as in the text box.



See Creating or editing patterns.

Limitations

- A variable can have up to 100 newline-separated values.
- In a text-type condition, a single phrase can include up to two variables, including duplicates.
- In a policy's text-based condition, if variables are used, the exclude match feature cannot be applied.
- Variables are not permitted in the exclude match text box.
- If a variable is misspelled, not defined, or not used with the correct syntax (for example,), it is treated as plain text during classification and matched accordingly.

Regular expression matches

A regular expression, or regex for short, is a pattern of text that consists of ordinary characters (for example, letters *a*through*z*) and special characters, called*metacharacters*. The pattern describes one or more strings to match when searching text. For example, the following regular expression matches the sequence of digits in all Visa card numbers:

```
\b4[0-9]{12}(?:[0-9]{3})?\b
```

Your regular expressions must conform to the Perl regular expression syntax.

You may find it helpful to build and test your regular expressions using the free online tool at https://regex101.com. This tool displays an explanation of your regular expression as you type it, and also lists all matches between the regular expression and a test string of your choice. The default regular expression flavor, pcre (php), is compatible with the Arctera Insight Classification.

> “ ”

Note: Looking for regular expression matches is considerably slower than looking for matches for specific words or phrases. You can greatly improve performance and accuracy by looking for instances where both types of matches occur in proximity to each other. To do this, set up anAll ofcondition group that contains both a regular expression condition and acontains textcondition for finding specific words and phrases, and specify the required distance within which matches must occur. The Arctera Insight Classification first evaluates thecontains textcondition and only then looks for a regular expression match.

> “ ”

Author and Recipient department based policy condition

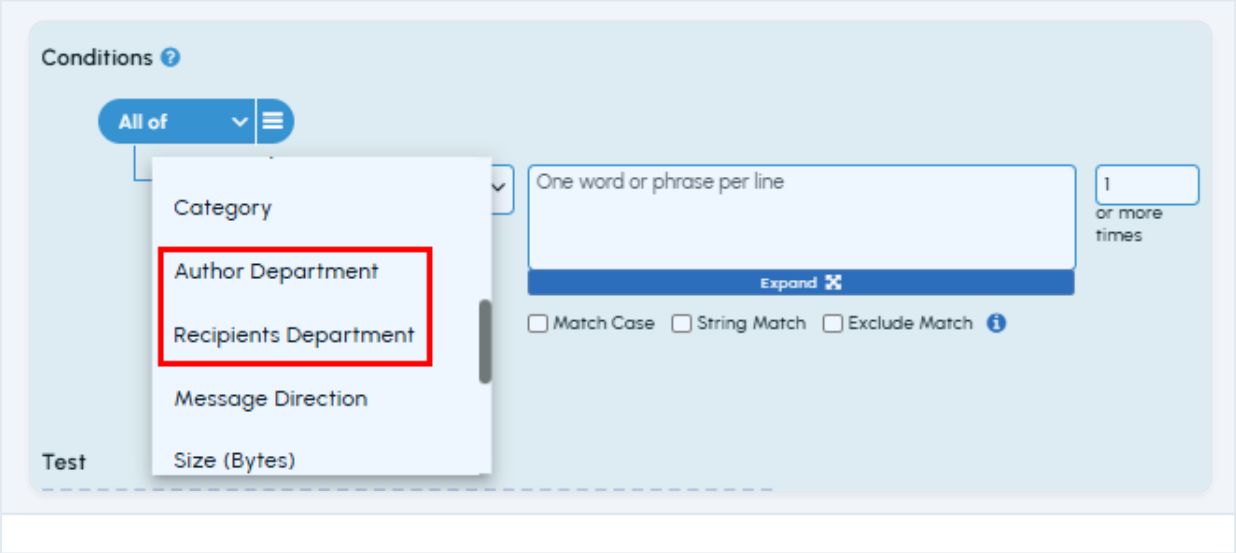
With this feature, you will be able to create policy conditions based on departments configured in Arctera Advanced Supervision. This will enable organizations to create policies that apply only to content sent or received by monitored employees from specific departments. The departments get updated over time, as people leave or join a department, the policy will automatically adjust itself. This new capability avoids the need to create multiple policies and applying them broadly. So, with this level of granularity one can configure multiple similar policies that might be slightly different for each department such as a Market Abuse Policy for departments in AMS and a Market Abuse Policy for departments in EMEA.

You can enable Author and Recipient department based policy condition by using `departmentApiEnabled` parameter in *yaml*.

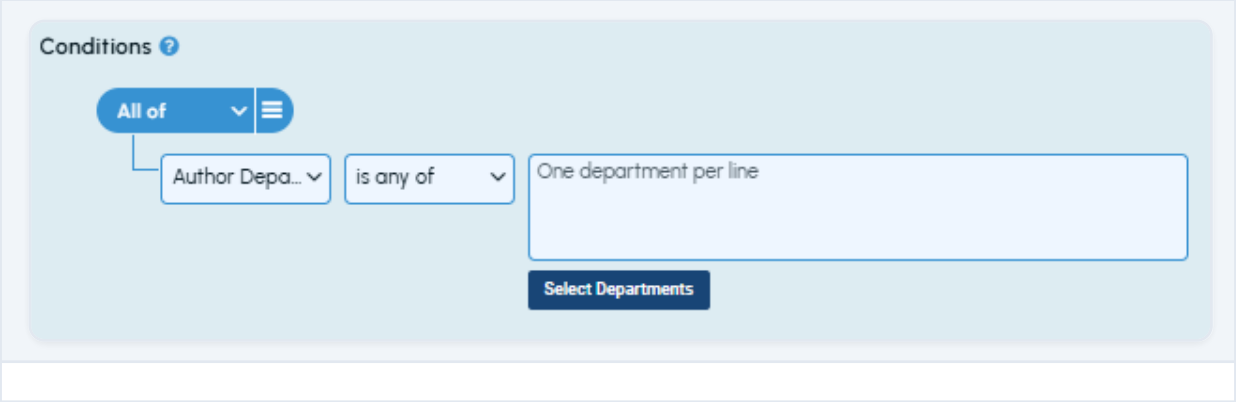
To use Author and Recipient department based policy condition,

1. Navigate to Policies and click New at the bottom of the page.

2. In the condition field, select Author Department or Recipient Department.



3. is any of option is selected by default.
4. The condition value field next to it will display all the selected departments. Note that this is a read-only field.



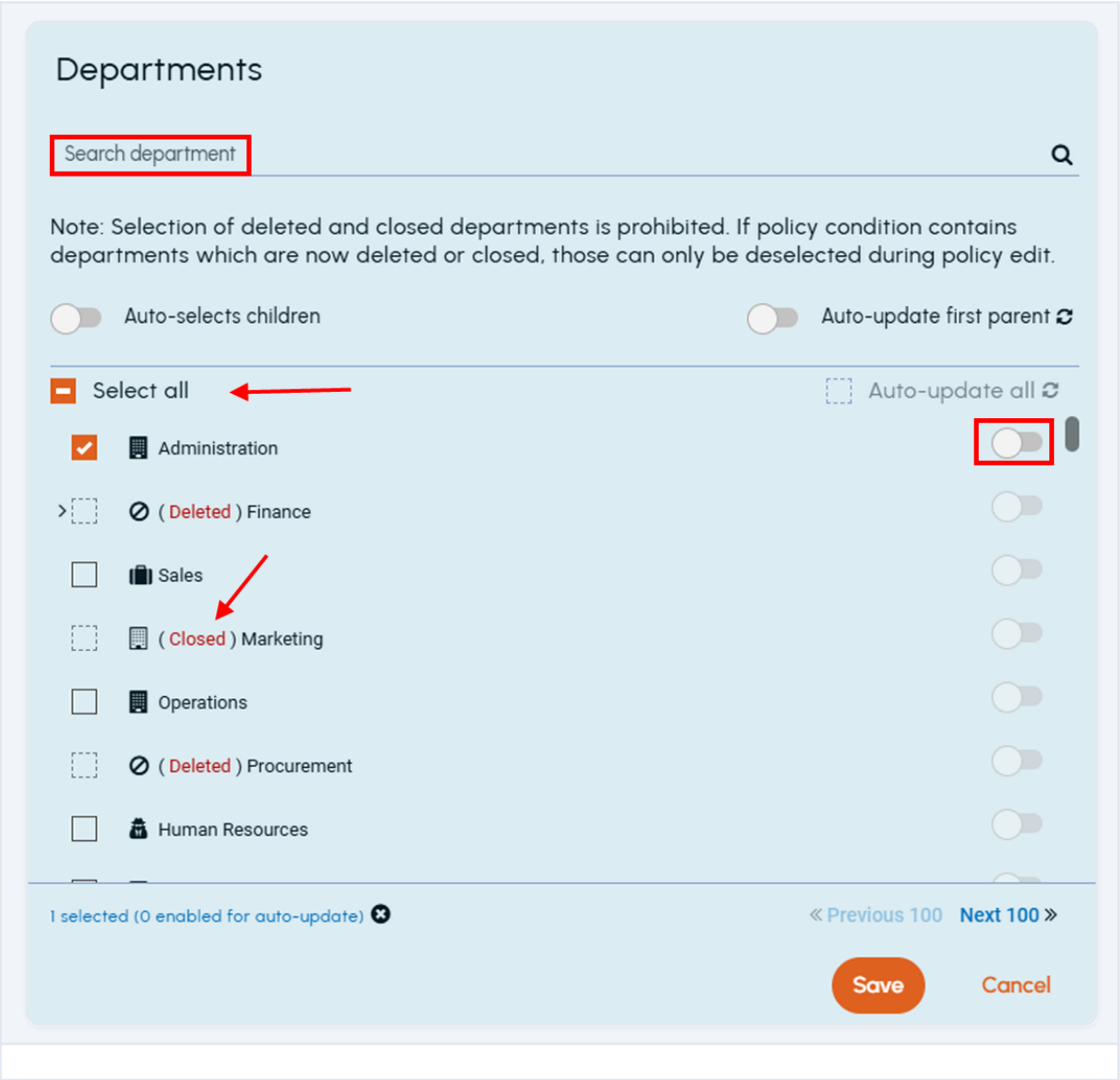
5. Click Select Departments and a pop-up with the list of department will appear on screen. Check or uncheck the box next to the department to add or remove any department.
6. You can also search for the departments by using the *search box* at the top.

> “ ”

Note: The searched departments will be listed along with their parents, even if the parent is not a match for the search. Selection of deleted and closed departments is not permitted. If policy condition contains departments which are now deleted or closed, those can only be deselected during policy edit.

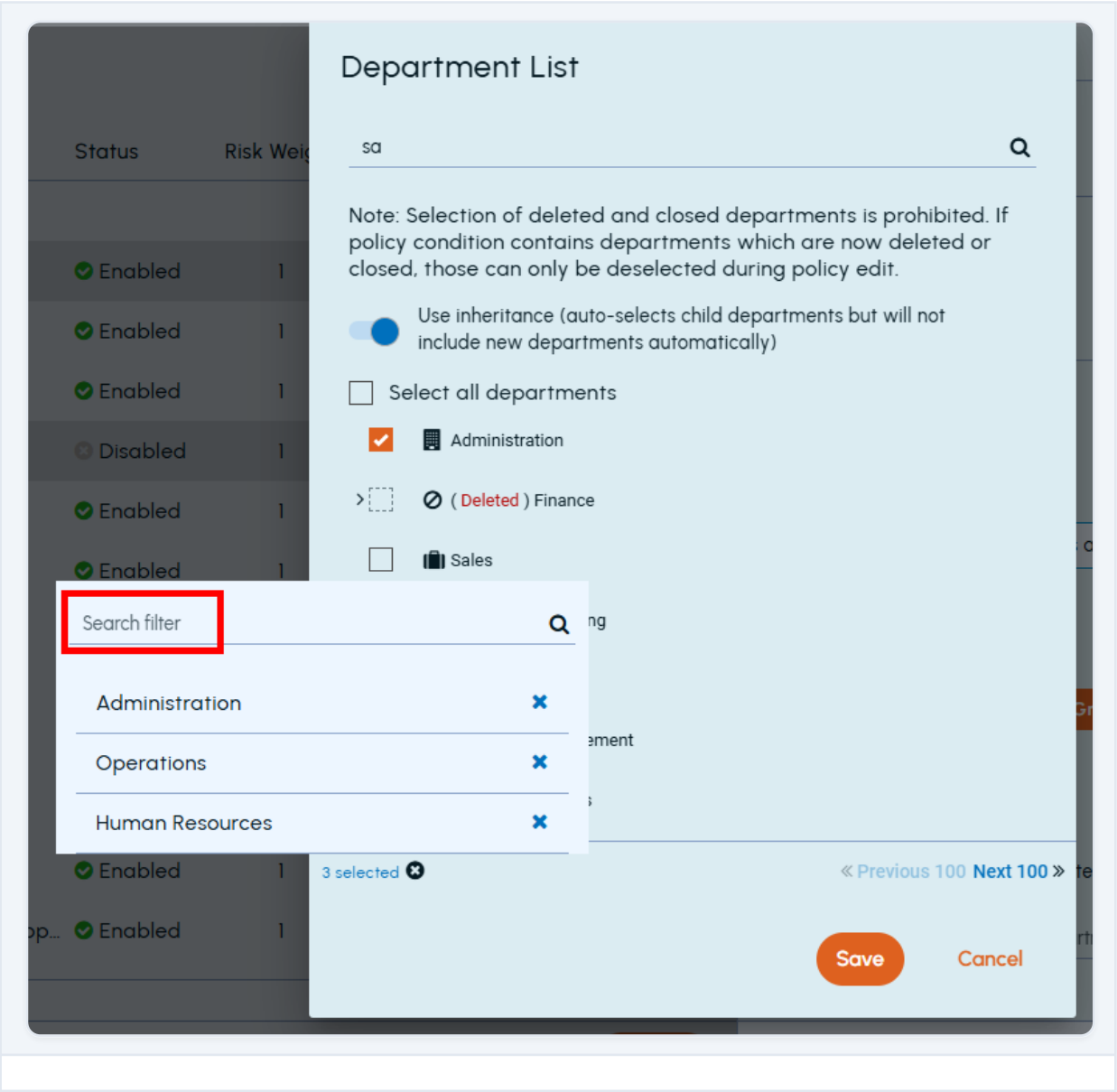
> “ ”

7. If the Use inheritance toggle on the same screen is turned on, all the child departments under the department will be selected automatically. By default, this toggle will be Off.
8. Select all departments checkbox can be used to select or deselect all the departments with a single click. This selection will not impact closed and deleted departments marked in red .
9. If you want to automatically include all newly added child departments under a specific department (for example, Department A) in a policy condition, you must first select that department and then enable the Auto-update toggle on the right for it. To enable auto-update for all departments at once, you can use the Auto-update all checkbox; however, this option is available only when all departments are selected. Additionally, if you want to automatically include any new departments that are added at the root level, you need to turn on the Auto-update root department toggle.



- The auto-update department syncing activity runs once every 24 hours at 1 AM.
- During this process, all newly added child departments of a parent department (for which auto-update has been enabled) are automatically included in the policy condition. However, existing child department that were not selected before enabling auto-update are not included.
- If a complete hierarchy is added as a child department, only the root-level department is included in the policy condition, and not the entire nested hierarchy.
- Auto-update for a department can only be enabled when that department is selected.
- The Auto-update root department option is not available if the departments are filtered using the search function.

10. All selected departments can be viewed or searched by clicking the number at the bottom. All selected departments can also be deleted by clicking X next to the number.



Currently, only 100 parent departments are displayed on one page, irrespective of the number of their child departments. If there are more than 100 parent departments, you can navigate across multiple pages one by one.

Limitations of the department based policy conditions

- After removing multiple items from pinned department panel rapidly and closing it, reopening that panel might take a while.

Pattern matches

A pattern match evaluates the selected item property against an existing Arctera Insight Classification pattern. Depending on the selected pattern, you may be able to set the confidence

levels that you are willing to accept. A high confidence level is likely to produce fewer but more relevant matches.

Note the following if you do not get the expected results when you test a policy that makes use of a built-in pattern:

- It is important to check that your test item meets the pattern confidence levels. For example, by default, the Credit Card Policy looks for content that matches the pattern "Credit/Debit Card Number" with medium to very high confidence. To meet the requirements of the medium confidence level, an item must contain either of the following:
  - A delimited credit card number (one that contains spaces or dashes between the numbers).
  - Both a non-delimited credit card number and one or more credit card keywords, such as "AMEX" or "Visa".

So, an item does not meet these requirements if it contains a non-delimited credit card number but it does not also contain credit card keywords.

- After you click Show details to view the results of a test, the Test classification results window may fail to highlight some or all of the matches. This is a known issue with certain patterns only. A future version of the Arctera Insight Classification will correct the issue.

Exact Data Matches

Unlike most classification techniques that rely on pattern matching to identify sensitive data, Exact Data Match (EDM) triggers a classification response when the actual data that needs to be protected is detected. By matching on the exact data, this reduces the rate of false positives and allows for much higher levels of accuracy in automatic classification. EDM uses a fingerprint method whereby an extract of a database or table is provided as source file in either CSV or TXT format. The table is ingested, and rules are created that indicate a match when one or more columns of a single row are detected in proximity. EDM is ideal when the identification of discrete customer data, employee data, and any other sensitive data repository maintained within a table is required.

To classify information using Exact Data Match

- Create an EDM pattern by setting the configuration options and providing the source document (typically containing the desired fields exported from a data store, such as a database). See “To create an Exact Data Match based pattern”.
- Use the resulting EDM pattern in any policy to be used for EDM based classification.

Exact Data Match can be enabled or disabled using YAML.

The Exact Data Match feature allows you to detect the specific data sets from a database. For example, employee records. You can match one or more fields and optional fields as per the configured proximity value. It supports large data sets (like database records) and text in all languages and provides data protection by hashing the stored data. The main benefit of using Exact Data Match is to reduce false positives by matching data exactly (unlike pattern-based matching).

For example, if you have the following content in the document to classify:

Name: Teresa M. Brown

Employee ID: 624828

and you are trying to match against the following EDM source document,

First name	Last name	Employee ID
Nick	Jones	317419
Teresa	Brown	624828
John	Miller	163332

Then this will trigger a match.

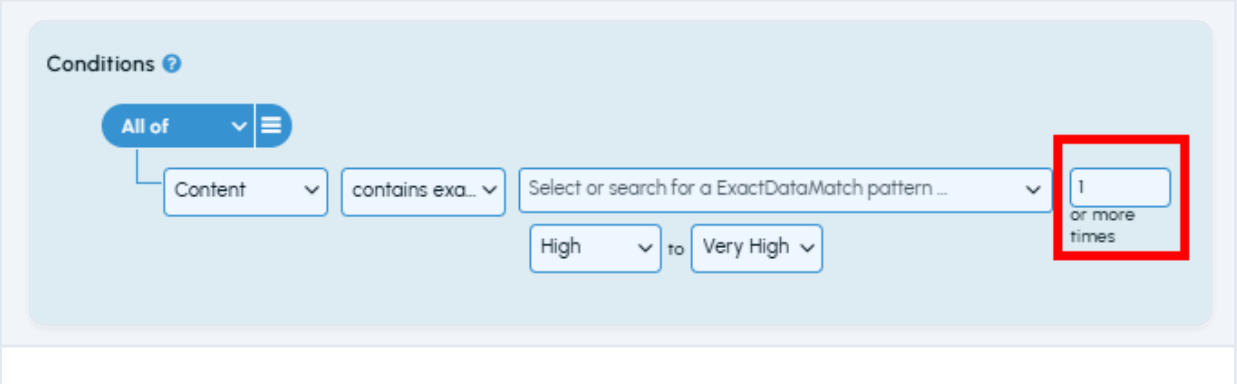
Exact Data Match provides following benefits:

- Provides the ability to detect specific data sets from a database. For example, employee records.
- Supports matching of combinations of data. For example, matching one or more fields and optional fields as per configured proximity value.
- Supports large data sets like database records.
- Provides data protection by hashing of stored data.
- Automatically synchronizes the exact data match rule pack (which is required for classification) on the remote classification servers. Manual intervention is not required.
- Supports file encryption for the exact data match rule pack files with the mechanism similar to tenant-specific patterns, policies, and tags.

- Supports the Min/Max disposition for exact data match type policy conditions while configuring policies.
- Supports text in all languages.

To create a policy using an Exact Data Match pattern

1. Follow the initial steps for creating or editing a policy as described earlier.
2. In the operator list box, select **contains exact data match in** and then select the required exact data match pattern from the value list box next to it.



Note: UnderConditions, the Min/Max disposition support is added for exact data match type policy conditions while configuring policies. You can specify the exact or more counts for keywords match. When you select the value more than 1, theExclude repeatscheck box appears. If you select this check box, matches that are different from each other. For example, a credit card condition with a minimum count of two requires two different credit cards in a single document.

3. Click **Save**.

When you test a document against a EDM based policy, Arctera Insight Classification shows the result. Also, the first column of the matching row is highlighted.

Example 1:

If source document content is as follows,

FirstName	LastName	Country
Stuart	Broad	England
Trent	Boult	New Zealand
James	Anderson	England

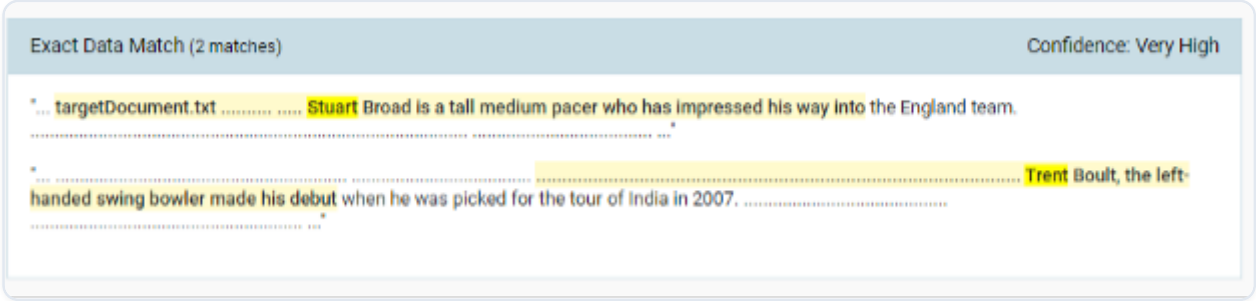
with Exact Data Matching Options as follows:

NAME	VALUE
First row contains column headers	Yes
Column delimiter	,
Perform hashing to secure data fields	No
Use case-sensitive matching	No
Proximity for matches	200
Minimum columns to match	2
All columns	Not selected

And if test document content is as follows:

```
..........
.....
Stuart Broad is a tall medium pacer who has impressed his way into the England team.
.....
.....
.....
Trent Boult, the left-handed swing bowler made his debut when he was picked for the tour of India in 2007.
.....
.....
James Born in Burnley, Lancashire.  He is perhaps the best swing bowler in the world at the present moment.
.....
.....
.....
```

The classification result will show a match for two records Stuart, and James.



Example 2:

For same source document and test document as stated in earlier example, if Minimum Columns value is set to 3 as follows:

NAME	VALUE
First row contains column headers	Yes
Column delimiter	,
Perform hashing to secure data fields	No
Use case-sensitive matching	No
Proximity for matches	200
Minimum columns to match	3
All columns	Not selected

The classification result will show a match for single record, that is Stuart. Because all 3 fields from first record is present in test document.

Example 3:

For same source document and test document as stated in first example, if proximity value is set to 50 as follows:

NAME	VALUE
First row contains column headers	Yes
Column delimiter	,
Perform hashing to secure data fields	No
Use case-sensitive matching	No
Proximity for matches	50

NAME	VALUE
Minimum columns to match	3
All columns	Not selected

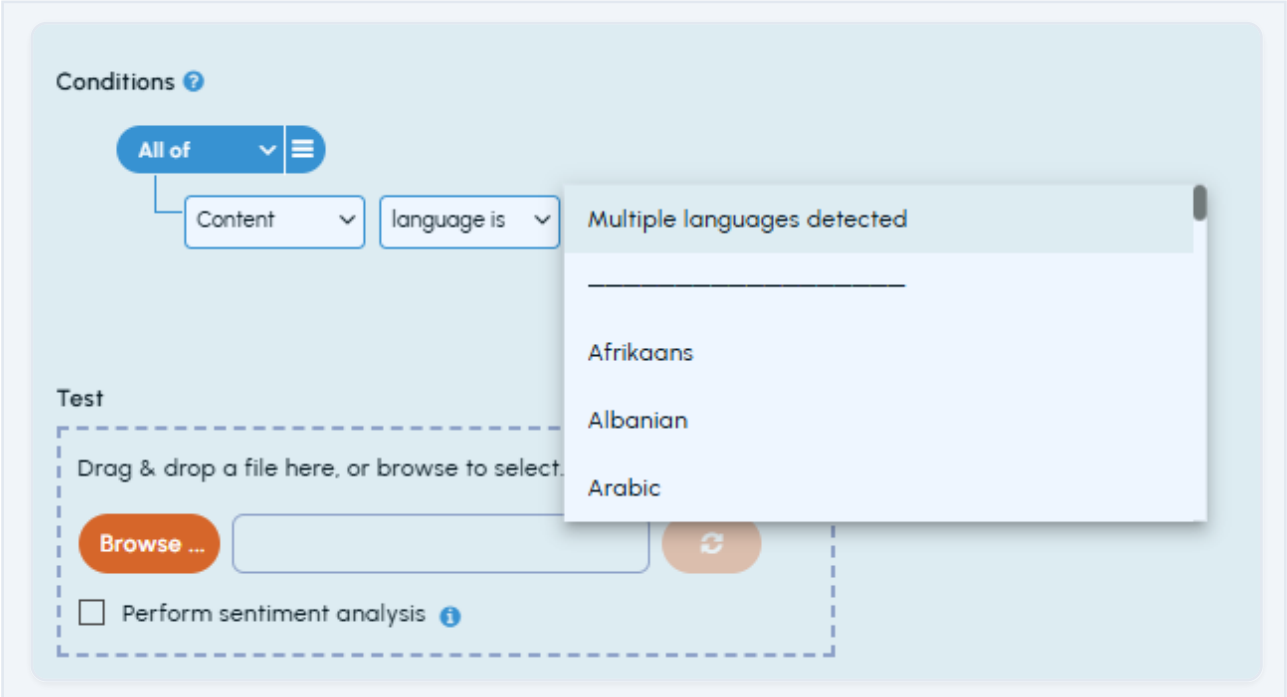
In this case, required words are not within proximity of 50 characters. Therefore the result will show no match.

Classification performance for Exact Data Match based policy depends on following factors.

- Number of records to be matched
- Number of fields and field size
- Data being classified
- Number of matches
- Proximity and column matches found
- Compute hardware and available resources

Language matches

You can set up a condition to restrict policy matching to items in a particular language. For example, set the condition like the one below to find items whose content is primarily in French:

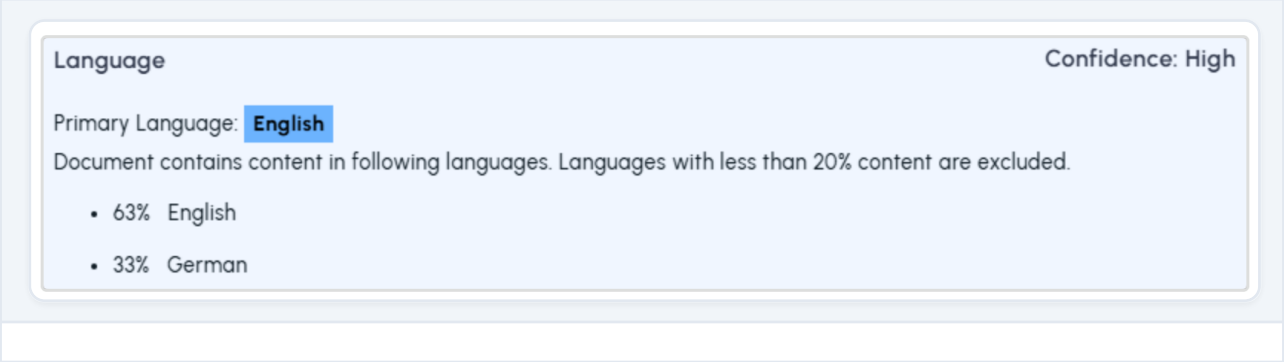


One of the options in the language list is Multiple languages detected. This option matches items that contain at least two languages.

To safeguard against the Arctera Insight Classification ignoring items because it cannot determine their primary language, select Or Primary Language Unknown. The most common reason why the Arctera Insight Classification may be unable to determine an item's primary language is that the item has a very small amount of content.

Language Proportion Detection

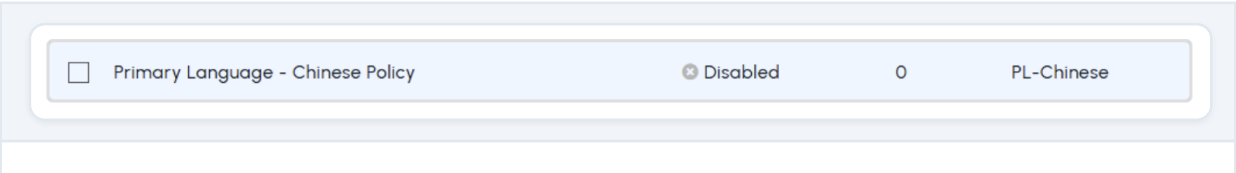
Arctera Classification supports detailed language analysis by identifying and reporting the proportion of each language found within a document. When a document contains content written in multiple languages, the system calculates and displays the percentage of text in each language. For example, a document may be identified as containing 63 percent English and 33 percent German. This capability is helpful for organizations working with multilingual content.



Choosing a Content Language Detection Engine

You have the flexibility to select between two content language detection engines, depending on your specific use case.

- Apache Tika Engine: The default option, which has comparatively lower speed and accuracy than FastText.
- FastText Engine: A newly introduced engine that offers faster and more accurate detection for both single-language and multilingual text.



To configure your preferred engine, navigate to the system settings and select the engine that best aligns with your performance or compatibility requirements.

Entity matches

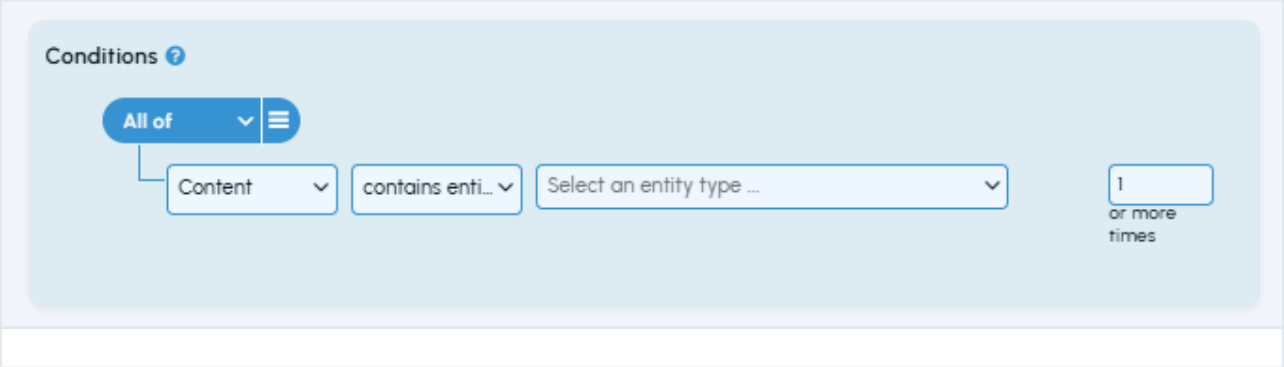
You can set up a condition to restrict policy matching to content that includes a person name or location.

> “ ”

Note: The "contains entity" condition will only be available ifnlp-service.jaris used while running the Arctera Insight Classification application. Also, Named Entity Recognition (NER) is available only for English.

> “ ”

For example, set the condition like the one below to find content including the person name.



> “ ”

Note: Named Entity Recognition (NER) consumes more time and resources compared to normal classification. NER is not suitable for large documents, especially documents bigger than 10 MB.

> “ ”

Sentiment score

You can set up a condition to perform sentiment analysis at item level and determine whether the sentiment associated with the item is positive, negative, or neutral based on the sentiment analysis score:

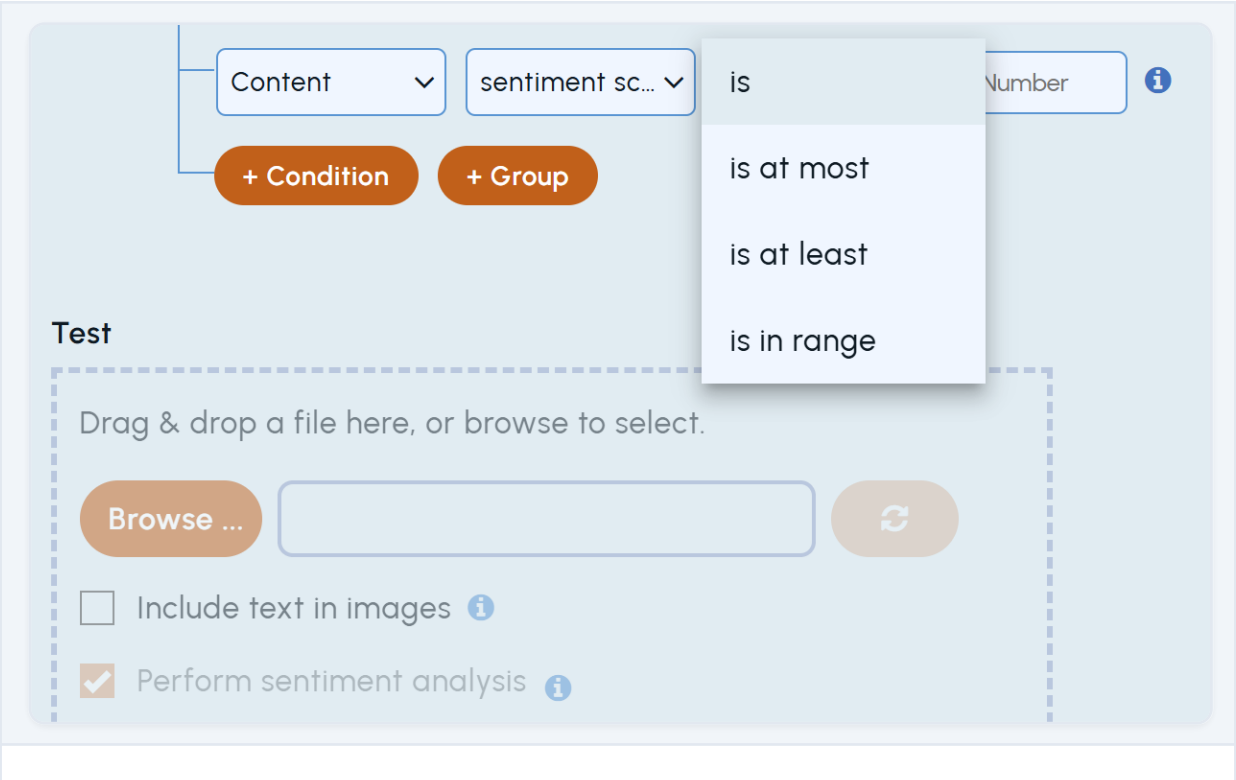
- score = 0 or score < 50 is negative

- score = 50 is neutral
- score = 100 or score > 50 is positive

Depending on how you want to interpret data, you can define and tailor policies with sentiment conditions to meet your sentiment analysis needs. Users can input an items of their choice and gauge the underlying sentiment based on the sentiment analysis score.

Sentiment analysis processing depends on the conditions set for sentiment score. If the set condition is met, then only the policy related tag is returned. The following conditions are available for sentiment score:

- is \- Sentiment analysis is performed, and the policy tag will be returned only if the sentiment score matches this number
- is at most \- Sentiment analysis is performed, and the policy tag will be returned only if the sentiment score is less than or equals this number
- is at least \- Sentiment analysis is performed, and the policy tag will be returned only if the sentiment score is equal to or higher than this number
- is in range \- Sentiment analysis is performed, and the policy tag will be returned only if the sentiment score falls in this range



For example, you set the condition as sentiment score is at most 70, then sentiment analysis will be performed, and the policy tag will be returned only if the score is 70 or less than 70.

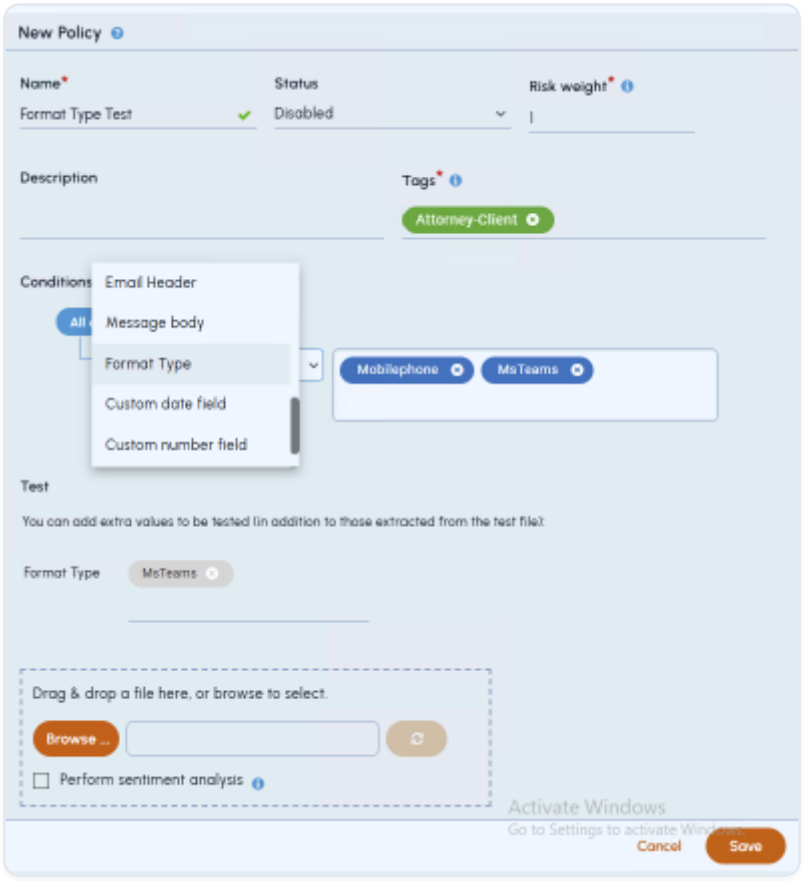
Note: If sentiment analysis fails, the classification process will continue without evaluating the sentiment analysis conditions within the policies. As a result, hits and matches based on the sentiment condition will be affected.

Format Type

You can set up a condition using format type.

To create a policy using Format Type condition,

1. Click Policy in the left pane
2. Click New and enter the mandatory details.
3. From the content drop-down, select Format Type
4. Select the options from the available list.
5. Click Save



Limitations of the Format Type policy condition

- Export is not allowed with format type policies.
- If multiple policies are selected for export and any policy is of format type, then all policies except the format type policy will be exported, depending on their types.
- After a format type policy is created, if the feature is disabled, classification for format type policies will not be allowed.

Risk score and risk level information on classification

Risk score and risk level for each classified item is sent to the consuming applications. Consuming applications can analyze this information and support features such as sort, filter, search, and report on items by risk score and/or risk level. By understanding the level of risk, you can optimize efforts on data management, review, and control. You can prioritize activities and resources on items of highest risk.

The risk score and risk level are based on the number of pattern or policy condition hits. Items with more hits are categorized as high risk. Items with fewer hits are categorized as low risk.

Configuring risk level settings through YAML file

In the YAML file, the previously used *lowerRiskRuleNameParts*parameter is deprecated, and the three new parameters -*lowRiskUpperLimit*,*mediumRiskUpperLimit*, and*highRiskUpperLimit* \- are added. These parameters provide control over different risk level definitions based on the risk score value. This configuration defines the upper limit of the risk score range for the low, medium, and high-risk levels.

- lowRiskUpperLimit \- It can be zero or greater than zero. By default, it is set to 2.
- mediumRiskUpperLimit -It can be any non-zero positive integer. But must be greater than *lowRiskUpperLimit* value. By default, it is set to 5.
- highRiskUpperLimit -It can be any non-zero positive integer. But must be greater than *mediumRiskUpperLimit* value. By default, it is set to 10.

This setting defines the upper limit of risk score range for low, medium, and high risk levels.

CONDITION	RISK LEVEL
Risk score > highRiskUpperLimit	Very high

CONDITION	RISK LEVEL
highRiskUpperLimit >= Risk score > mediumRiskUpperLimit	High
mediumRiskUpperLimit >= Risk score > lowRiskUpperLimit	Medium
lowRiskUpperLimit >= Risk score >= 1	Low
Risk score = 0	No risk

Risk score and risk level information in classify API response

The risk information is sent as part of classify response only if following conditions are met:

- The *matchDetailLevel* is configured in classify request as either LOW/MEDIUM/HIGH
- The item must have some risk based on the risk score and the risk level limits settings in the YAML file.

Risk score calculation

The risk computation of potentially sensitive content is based on the degree of hits against patterns or policy conditions and policy risk weight.

> “ ”

Note: By default, the Risk weight value of all the custom policies and most of the built-in policies is configured as 1. For Subscription policy and all the Language detection policies risk weight is set to 0 by default.

> “ ”

Consider the following example. A document has the following classification result.

Policies

Classification result

POLICY NAME	PATTERN NAME (MATCH COUNT)	RISK WEIGHT
Policy-1	Pattern-A (2), Pattern-B (3)	2
Policy-2	Pattern-B (2), Pattern-C (1)	0
Policy-3	Pattern-C (3), Pattern D (5)	1
Policy-4	Pattern-C (1), Pattern E (1)	5

Unique policy match table

The following table describes a sample scenario for risk score calculation.

PATTERN NAME	MATCH COUNT	FROM POLICY	RISK WEIGHT
Pattern-A	2	Policy-1	2
Pattern-B	3	Policy-1	2
Pattern-C	3	Policy-3	1
Pattern-D	5	Policy-3	1
Pattern-E	1	Policy-4	5

To calculate risk score, the policy with the highest risk weight is considered in case the pattern hitting on item is present in multiple policies as shown in the unique policy match table.

The risk score is a sum of the products of match count and the policy risk weight. The following susing an Exact Data Match patternteps explain the step-level actions for risk score calculation.

Step1: Multiply match count with risk level.

Step2: Repeat step1 for all the rows in the unique rule match table.

Step3: Add the results of step2.

Risk Score = 2*2* + *3*2 + 3*1* + *5*1 + 1*5

= 4 + 6 + 3 + 5 + 5

= 23

Risk levels

Risk is categorized in different risk levels as per the risk score and is described in the About policy conditions section.

In the above example, the risk score is 21. Therefore, the Risk is categorized as:Risk level : Very high

Facts and limitations about risk score:

- Sentiment score/Named Entity based policy condition hits does not contribute towards risk score.
- Contribution to the total item risk score will be zero due to Subscription policy and any language detection policy as all these policies have risk weight zero by default.
- Due to language detection policy hits, some discrepancies may be observed in Most Common Sensitive Data results from analyzer overview page. The result of the analyzer overview page is accurate.
- Following policy conditions contribute to risk score.
    - Content
    - Title
    - Author
    - Content Type
    - Recipient
    - Modified Date
    - Creation Date
    - Sensitivity
    - Category
    - Size (Bytes)
    - Custom date field
    - Custom number field
    - Custom string field
    - Sender and Recipient department based policy condition

Condition groups

You can group a set of conditions and nest grouped conditions within other grouped conditions. The group operator that you choose determines whether an item must meet all, some, or none of the conditions in the group to be considered a match. The following group operators are available:

- All of. An item must meet all the specified conditions.
- Any of. An item must meet at least one of the specified conditions.
- None of. An item must not meet any of the specified conditions.

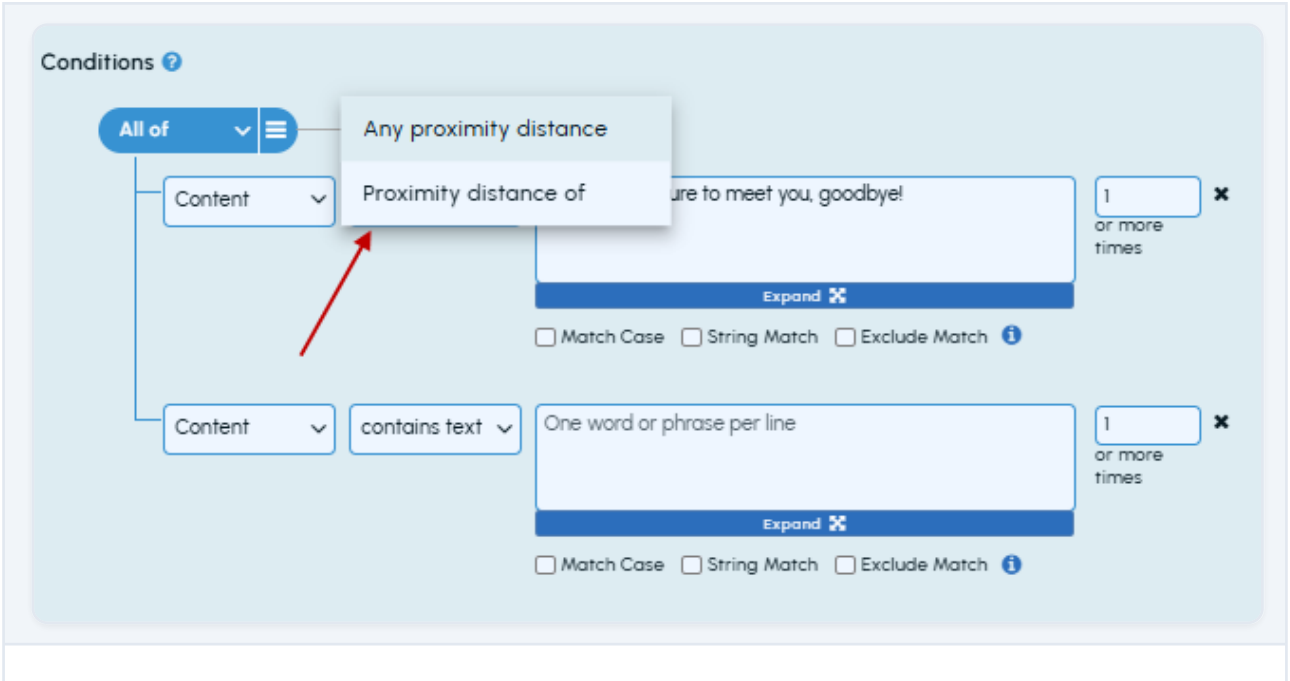
> “ ”

Note: You can nest aNone ofgroup within anAll ofgroup to look for certain condition matches while also excluding others. For example, to achieve the effect of "(condition X AND condition Y) BUT NOT condition Z", you would include the X and Y conditions in anAll ofgroup and the Z condition in a nestedNone ofgroup.

> “ ”

- *n* or more of. An item must meet the specified number of conditions.

For an All of group only, you can choose to look for instances where the conditions occur within any proximity or specific/specified number of characters from the first condition. For example, the following condition group looks for instances where the word *Goodbye*appears within 20 characters of the word*Hello* :



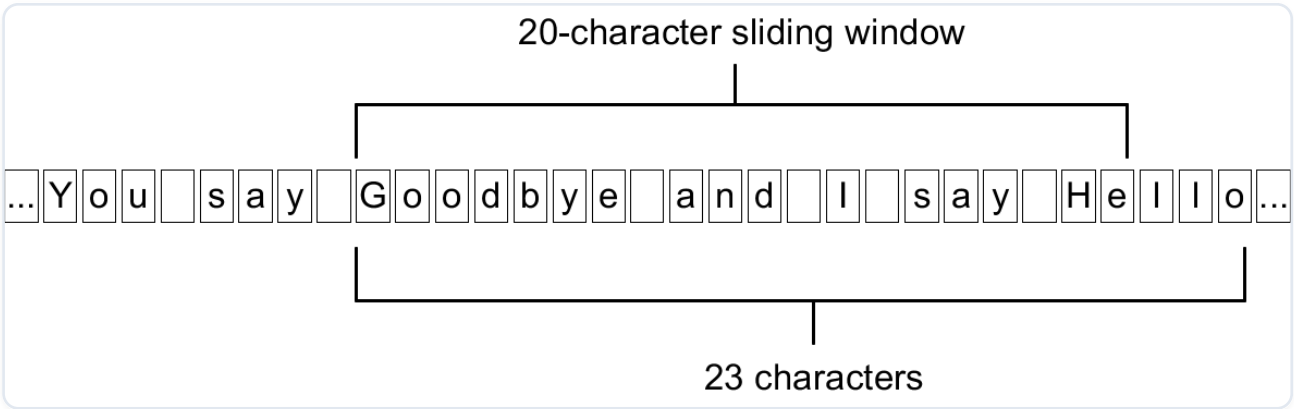
The text string "You say Goodbye and I say Hello" matches these conditions because there are fewer than 20 characters between the first character of *Hello*and the first character of*Goodbye*. Similarly, the string "You say Hello and I say Goodbye" also matches because there are fewer than 20 characters between the ends of the two words. In each case, the spaces count as characters.

> “ ”

Note: When you conductwithinnncharactersproximity searches, take care not to duplicate the same search terms across multiple conditions. For example, suppose that you define one condition to look for the namesFred,Sue, andBob, and a second to look forJoe,Bob, andSarah. An item that contains a single instance ofBobwould match these conditions.

> “ ”

Rather than choose the from the first condition option, you can choose in a sliding window. This option looks for instances where the conditions occur within any sequence of characters of the specified number. For example, a condition group that looks for instances where the word *Goodbye*appears within a 20-character sliding window of the word*Hello*does not match "You say Goodbye and I say Hello". There are 23 characters between the start of the word*Goodbye*and the end of the word*Hello*.



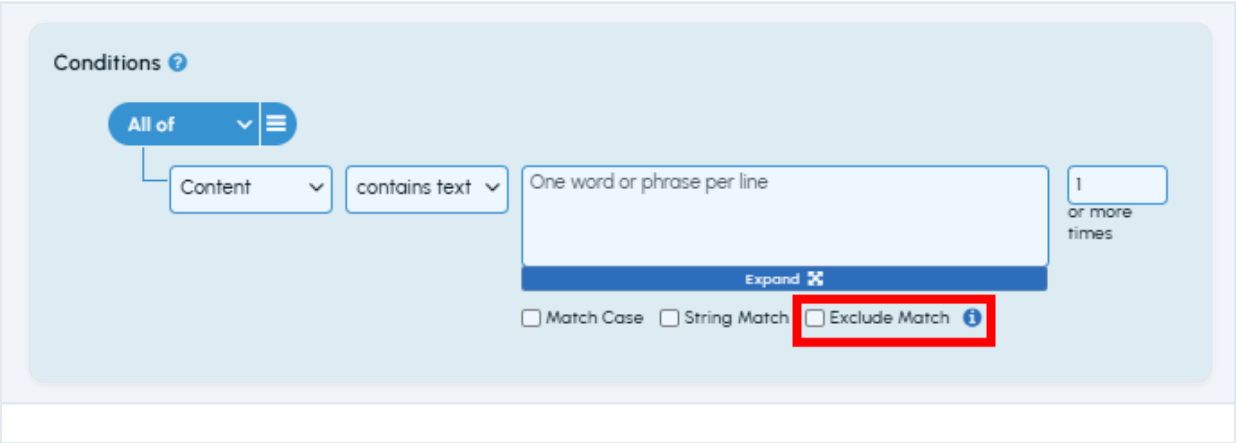
More Information

“To create a policy using an Exact Data Match pattern”

Using a keywords-based exclusion policy condition

To use a keywords-based exclusion policy condition

1. Click the **Exclude Match** check box.



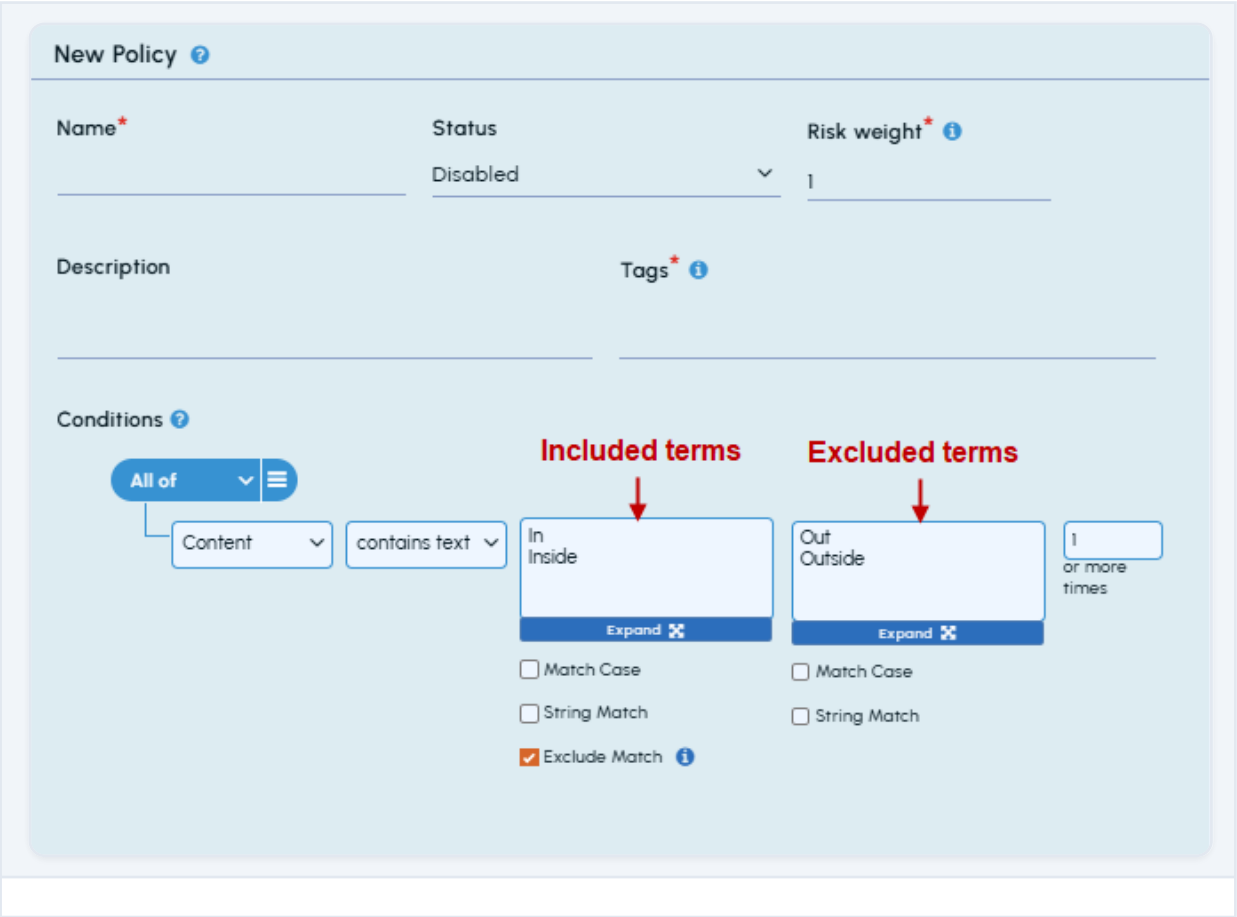
The new input field for defining exclusion terms is displayed.

1. In the inclusion terms field, enter the keywords, phrases, or regular expression that you want to include while evaluating the policy condition criteria. In the exclusion terms field, define the keywords or phrases that you want to exclude while evaluating the policy condition criteria.

   > “ ”

   Note: Use only single word, phrase, or regular expression per line in both the fields as shown in the sample image below.

   > “ ”



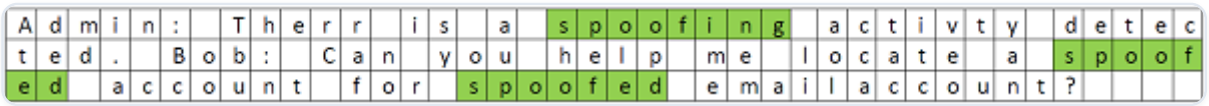
For more information about supported types of regular expression syntax, See Regular expression syntax.

1. If required, select the **Match Case**and/or**String Match** check boxes for both input fields.
2. Specify the minimum count value in the last input field that applies to the number of matches in the document, and not to the number of excluded matches.

Below is a scenario to understand the concept in detail.

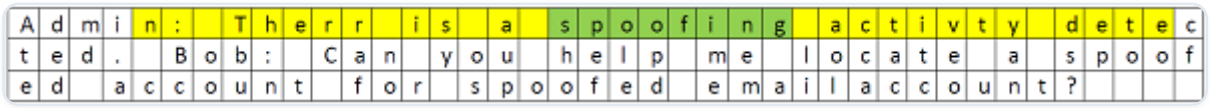
Test data - "Admin: There is a spoofing activty detec ted. Bob: Can you help me locate a spoof ed account for spoofed email account?"

If you do not use the Exclude Match option, the application provides a hit on spoofing and spoofed (two occurrences).

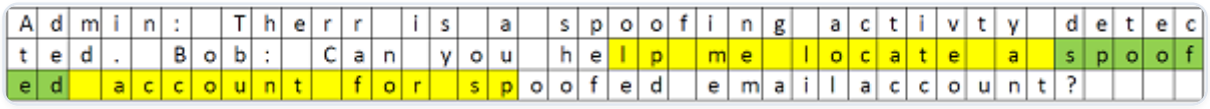


However, if you use the Exclude Match option, let's understand how the terms get excluded.

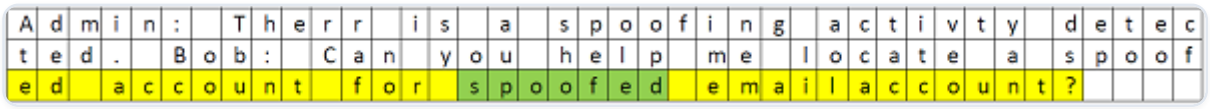
Step 1: The first hit which is on *spoofing*. The application looks in both the directions (forward and backward) by 22 characters proximity. Application can not find any exclusion term, therefore, the terms*spoofing* will be a match.



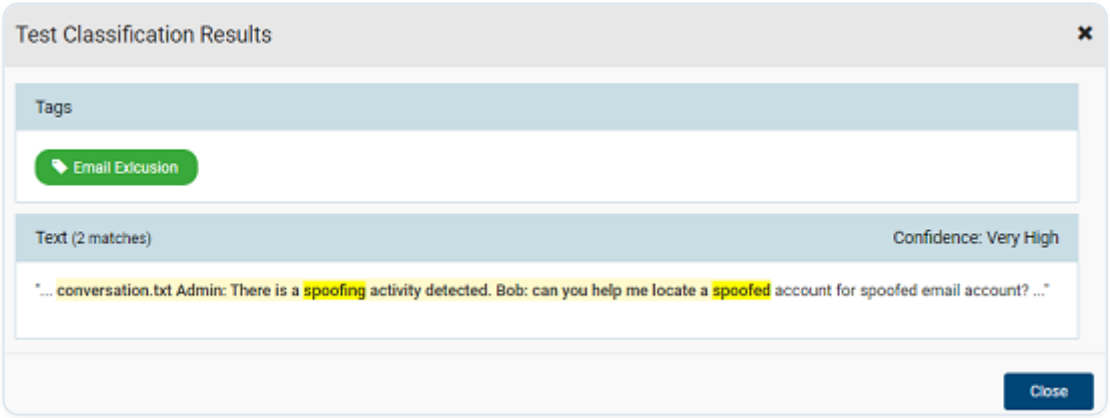
Step 2: The next hit will be the term *spoofed*. The application looks for exclusion terms in both the directions from the terms*spoofed*by 22 characters proximity. Application can not find any exclusion term, therefore, the terms*spoofed* will be a match.



Step 3: The next hit will again be the term *spoofed*. The application looks for exclusion terms in both directions from the terms*spoofed*by 22 characters proximity. The application finds an exclusion terms*spoofed email account* in the forward direction. Therefore, application ignores the term and do not consider it as a match.



In the Test Classification Results section, you get only one match for the term *spoofing*and another match for the terms*spoofed*terms. However, the match for the terms*spoofed email account* is excluded. The match result appears as shown in the sample image below:



Regular expression syntax

This topic includes the following:

- About regular expressions
- Wildcards
- Anchors

- Marked subexpressions
- Non-marking groupings
- Repeats
- Back references
- Alternation
- Character sets
- Escapes
- Perl-specific extensions
- Operator precedence

About regular expressions

The Arctera Insight Classification supports a regular expression syntax that is based on the syntax in the Perl programming language. In Perl regular expressions, all characters match themselves except for the following special characters:

```
.[]{}()\*+?|^$
```

For more information on the Perl syntax, see the following webpage:

https://perldoc.perl.org/perlre.html

You may find it helpful to build and test your regular expressions using the free online tool at https://regex101.com. This tool displays an explanation of your regular expression as you type it, and also lists all matches between the regular expression and a test string of your choice. The default regular expression flavor, pcre (php), is compatible with the Arctera Insight Classification.

> “ ”

Note: Looking for regular expression matches is considerably slower than looking for matches for specific words or phrases. You can greatly improve performance and accuracy by looking for instances where both types of matches occur in proximity to each other. To do this, set up anAll ofcondition group that contains both a regular expression condition and acontains textcondition for finding specific words and phrases, and specify the required distance within which matches must occur. The Arctera Insight Classification first evaluates thecontains textcondition and only then looks for a regular expression match.

> “ ”

Wildcards

The `.` (period) character matches any single character, when it is used outside of a character set.

Anchors

The `^` (caret) character matches the start of a line. The `$` (dollar) character matches the end of a line.

Marked subexpressions

A section that is surrounded with the characters `(` and `)` acts as a marked subexpression. The matching algorithms captures whatever matches the subexpression. Marked subexpressions can be repeated, or a back-reference can refer to them.

Non-marking groupings

A marked subexpression is useful for lexically grouping part of a regular expression, but it has the side-effect of additional overhead. As an alternative, you can lexically group part of a regular expression without generating a marked subexpression by using `(?:` and `)`. For example, `(?:ab)+` repeats `ab` without splitting out any separate subexpressions.

Repeats

You can repeat any atom (single character, marked or non-marked subexpression, or character class) with the operators `*`, `+`, `?`, and `{}`.

Table: Repeat operators

OPERATOR	DESCRIPTION
`*`	Matches the preceding atom zero or more times. For example, `a*b` matches any of the following\:
	lang=txt
	b
	ab

OPERATOR	DESCRIPTION
	aaaaaaaab
`+`	Matches the preceding atom one or more times. For example, `a+b` matches either of the following\:
	lang=txt
	ab
	aaaaaaaab
	However, it does not match `b` .
`?`	Matches the preceding atom zero or one times. For example, `ca?b` matches either of the following\:
	lang=txt
	cb
	cab
	However, it does not match `caab` .
`{}`	Repeats the preceding atom with a bounded repeat.
	`a{n}` matches `a` repeated exactly `n` times.
	`a{n,}` matches `a` repeated `n` or more times.
	`a{n,m}` matches `a` repeated between `n` and `m` times inclusive.

OPERATOR	DESCRIPTION
	For example, `^a{2,3}$` matches either of the following:
	lang=txt
	aa
	aaa
	However, it does not match `a`
	or `aaaa` .

These operators are "greedy"; they consume as much input as possible. However, non-greedy versions are available that consume as little input as possible while still producing a match. By following the repeat operators `*`, `+`, `?`, and `{}` with the `?` character, the repeats become non-greedy.

By default, when a repeated pattern does not match, the Arctera Insight Classification backtracks until it finds a match. This behavior can sometimes be undesirable for matchmaking or performance reasons, so there are also "possessive" repeats. These match as much as possible and do not then allow backtracking if the rest of the expression fails to match.

Back references

An escape character that is followed by a digit `n`, where `n` is in the range `1` through `9`, matches the same string that the subexpression `n` matched. For example, consider the following expression:

```
^(a{2,3}).*\1$
```

This matches `aaabbaaa`, but it does not match `aaabba`.

Alternation

The `|` operator matches either of its arguments. For example, `abc|def` matches both `abc` and `def`.

You can use parentheses to group alternations. For example, `ab(?:d|ef)` matches both `abd` and `abef`.

Character sets

A character set is a bracket expression that is enclosed within the characters `[` and `]`. It defines a set of characters, and matches any single character that is a member of the set.

A bracket expression can contain any combination of the following:

- Single characters. For example, `[abc]` matches any of the characters `a`, `b`, or `c`.
- Character ranges. For example, `[a-c]` matches any single character in the range `a` through `c`. By default, for Perl regular expressions, a character `x` is within the range `y` to `z`, if the code point of the character lies within the code points of the endpoints of the range.
- Negation. If the bracket expression begins with the `^` character, it matches the complement of the characters that it contains. For example, `[^a-c]` matches any character that is not in the range `a` through `c`.
- Character classes. An expression of the form `[[:name:]]` matches the named character class `name`. For example, `[[:lower:]]` matches any lowercase character. The supported character classes are as follows:

ALNUM	ANY ALPHANUMERIC CHARACTER.	PUNCT	ANY PUNCTUATION CHARACTER.
alpha	Any alphabetic character.	s	Any whitespace character.
blank	Any whitespace character that is not a line separator.	space	Any whitespace character.
cntrl	Any control character.	unicode	Any extended character whose code point is above 255 in value.
d	Any decimal digit.	u	Any uppercase character.

ALNUM	ANY ALPHANUMERIC CHARACTER.	PUNCT	ANY PUNCTUATION CHARACTER.
digit	Any decimal digit.	upper	Any uppercase character.
graph	Any graphical character.	w	Any word character (alphanumeric characters plus the underscore).
l	Any lowercase character.	word	Any word character (alphanumeric characters plus the underscore).
lower	Any lowercase character.	xdigit	Any hexadecimal digit character.
print	Any printable character.	-	-

- Escaped characters. All the escape sequences that match a single character or character class are permitted within a character class definition. For example, `[[]]` matches both `[` and `]`, whereas `[\W\d]` matches any character that is either a digit or not a word character.
- Combinations. You can combine one or more of the above in a character set declaration. For example, `[a-cmnx-y\d]`.

Escapes

Any special character that is preceded by an escape matches itself.

Table: Escape sequences that are synonyms for single characters

ESCAPE	CHARACTER
\a	\a
\e	0x1B
\f	\f

ESCAPE	CHARACTER
\n	\n
\r	\r
\t	\t
\v	\v
\b	\b (but only inside a character class declaration).
\c X	An ASCII escape sequence: the character whose code point is X % 32.
\x XX	A hexadecimal escape sequence: matches the single character whose code point is 0x XX .
\x{ XXXX }	A hexadecimal escape sequence: matches the single character whose code point is 0x XXXX .
\0 ddd	An octal escape sequence: matches the single character whose code point is 0 ddd .
\N{ name }	Matches the single character that has the symbolic name name . For example, \N{newline} matches the single character \n.

"Single character" character classes

When `x` is the name of a character class, the escaped character `x` matches any character that is a member of the class. Conversely, `X` matches any character that is not a member of the `x` class.

Table: Escape sequences for "single character" character classes

ESCAPE	EQUIVALENT TO	ESCAPE	EQUIVALENT TO
\d	\[\[:digit:\]]	\D	\[^\[:digit:\]]

ESCAPE	EQUIVALENT TO	ESCAPE	EQUIVALENT TO
\l	\[\[:lower:\\]]	\L	\[^\[:lower:\\]]
\s	\[\[:space:\\]]	\S	\[^\[:space:\\]]
\u	\[\[:upper:\\]]	\U	\[^\[:upper:\\]]
\w	\[\[:word:\\]]	\W	\[^\[:word:\\]]
\h	Horizontal whitespace	\H	Not horizontal whitespace
\v	Vertical whitespace	\V	Not vertical whitespace

Word boundaries

The following escape sequences match boundaries of words.

Table: Escape sequences for word boundaries

ESCAPE	DESCRIPTION
\<	Matches the start of a word.
\>	Matches the end of a word.
\b	Matches a word boundary (the start or end of a word).
\B	Matches only when not at a word boundary.

Line endings

The following escape sequences match line endings.

Table: Escape sequences for line endings

ESCAPE	DESCRIPTION
\n	Newline.

ESCAPE	DESCRIPTION
\r	CR.
\R	Any line-ending character sequence. This is identical to the following expression\:
	`(?>\x0D\x0A?|\[\x0A-\x0C\x85\x{2028}\x{2029}\])`

Other escapes

Except for the following characters, any escape sequence matches the character that is escaped:

```
' ` A C E G K Q X z Z

```txt



For example,``\@``matches a literal``@``.



Perl-specific extensions



All Perl-specific extensions to the regular expression syntax start
with``(?``.



Named subexpressions



You can create a named subexpression as follows:```txt

(?\<NAME\>expression)
```
```

```txt

You can then refer to the subexpression by the name``NAME``.
Alternatively, you can delimit the name, as in the following:```txt

(?'NAME'expression)

```txt



You can then refer to the subexpression in a backreference using
either``\g{NAME}``or``\k<NAME>``.



Comments``(?# ... )``is treated as a comment. Its contents are ignored.



Modifiers``(?imsx-imsx ... )``alters which of the Perl modifiers are in
effect within the pattern. Changes take effect from the point that the
block is first seen and extend to any enclosing``)``. Letters before
a``'-'``turn this Perl modifier on, and those after the``'-'``turn it
off.``(?imsx-imsx:pattern)``applies the specified modifiers
to``pattern``only.



Non-marking groups``(?:pattern)``lexically groups``pattern``, without
generating an additional subexpression.



Lookahead``(?=pattern)``consumes zero characters, but only
if``pattern``matches.``(?!pattern)``consumes zero characters, but only
if``pattern``does not match.



You typically use lookahead to create the logical AND of two regular
expressions. For example, if a password must contain a lowercase letter,
uppercase letter, and punctuation symbol, and it must be at least six
```
```

characters long, then you can use the following expression to validate the password:```txt

(?=.*[[:lower:]])(?=.*[[:upper:]])(?=.*[[:punct:]]).{6,}

```txt

Lookbehind``(?<=pattern)``consumes zero characters, but only if``pattern``can be matched against the characters that precede the current position (``pattern``must be of fixed length).``(?<!pattern)``consumes zero characters, but only if``pattern``cannot be matched against the characters that precede the current position (``pattern``must be of fixed length).

Independent subexpressions``(?>pattern)``matches``pattern``independently of the surrounding patterns. The expression never backtracks into``pattern``.

Conditional expressions``(?(condition)yes-pattern|no-pattern)``tries to match``yes-pattern``if the condition is``true``, and otherwise tries to match``no-pattern``.``(?(condition)yes-pattern)``tries to match``yes-pattern``if the condition is``true``, and otherwise matches the``NULL``string.``condition``may be one of the following:

- A forward lookahead assert.

- The index of a marked subexpression (the condition becomes true if the subexpression has been matched).

Here is a summary of the possible predicates:
```

|``(?(?=assert)yes-pattern\|no-pattern)``| Executes``yes-pattern``if the forward look-ahead assert matches, and otherwise executes``no-pattern``. |

| --- | --- |

|``(?(?!assert)yes-pattern\|no-pattern)``| Executes``yes-pattern``if the forward look-ahead assert does not match, and otherwise executes``no-pattern``. |

|``(?(N)yes-pattern\|no-pattern)``| Executes``yes-pattern``if subexpression N has been matched, and otherwise executes``no-pattern``. |

|``(?()yes-pattern\|no-pattern)``| Executes``yes-pattern``if named subexpression``name``has been matched, and otherwise executes``no-pattern``. |

|``(?('name')yes-pattern\|no-pattern)``| Executes``yes-pattern``if named subexpression``name``has been matched, and otherwise executes``no-pattern``. |

Operator precedence

The order of precedence for the operators is as follows:

- Escaped characters``\``- Character set (bracket expression)``[]``- Grouping``()``- Single-character-ERE duplication``* + ? {m,n}``- Concatenation

- Anchoring``^$``- Alternation``|``

Enabling or disabling policies

By default, all the policies are disabled. You must enable a policy if you want the Arctera Insight Classification to check for and tag the items that match the policy.

> “ ”

Note: Enabling a lot of policies can affect performance. In addition, policies with complex conditions take longer to process than those with simple conditions.

> “ ”

To enable or disable policies

1. In the left navigation pane, click **Policies**.
2. To enable or disable a single policy, search for and select that policy, and click **Edit**.

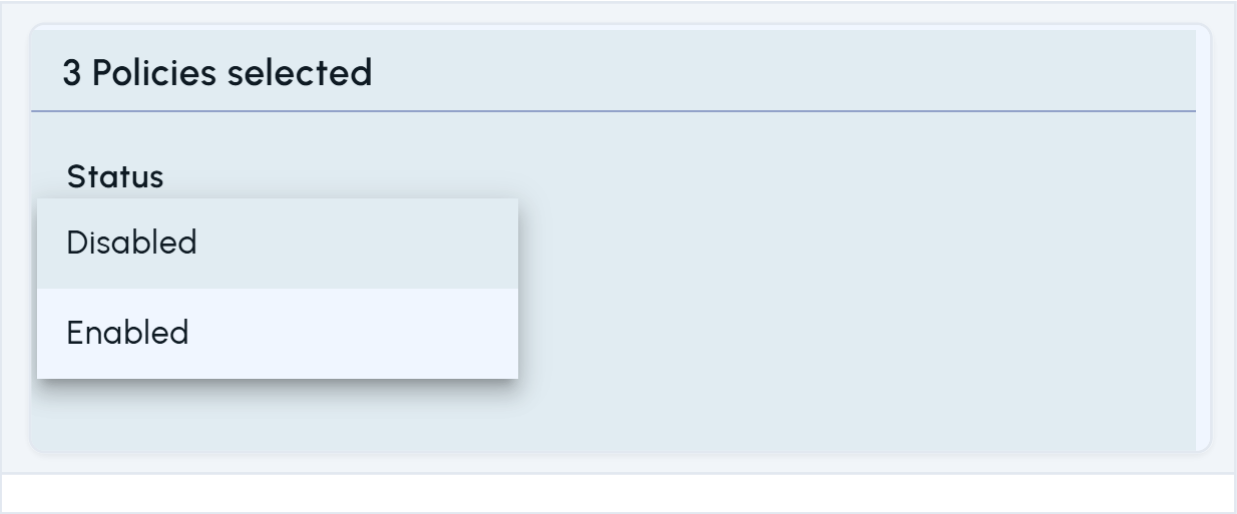
The **Edit \<*Policy Name* \>** dialog box appears.

In the **Status**field, select**Enabled**or**Disabled** for corresponding action.

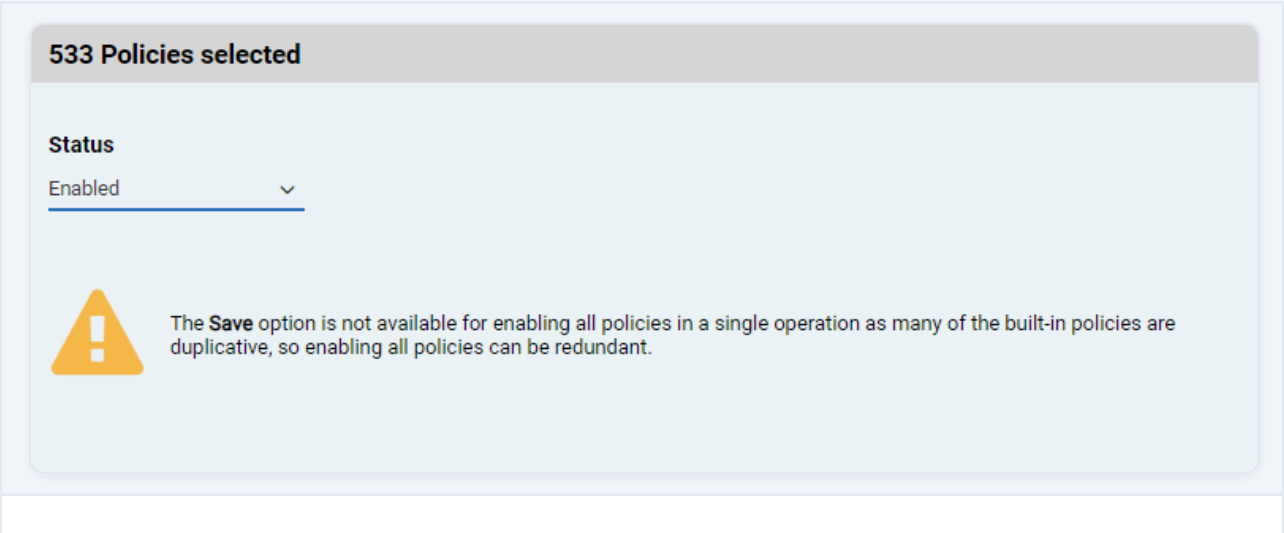
1. To enable or disable multiple policies simultaneously, either search for specific policies or select policies.

In the right pane, the application displays the number of policies selected. Click **Edit**.

1. In the **Status**field, select**Enabled**or**Disabled** for corresponding action.



Enabling all policies in a single operation is prohibited as there are multiple factors that you need to consider before enabling a policy. If you select all policies and try to enable, following error message will appear on the screen, prohibiting the operation.



1. Click **Save**.

Editing policies

If required, you can edit the policy to add or remove conditions and tags.

To edit a policy

1. In the left navigation pane, click **Policies**.
2. Search for and select the policy you want to edit.
3. Click **Edit**.

The **Edit \<*Policy Name* \>** dialog box appears.

1. Modify the following details as required.

NAME	SPECIFIES THE POLICY NAME. THE NAME MUST BE UNIQUE, AND IT CAN
	CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
Status	Enables or disables the policy. You must enable the policy if you want the

NAME	SPECIFIES THE POLICY NAME. THE NAME MUST BE UNIQUE, AND IT CAN
	CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
	Arctera Insight Classification to check for and tag the items that match the policy.
Description	(Optional.) Provides a short description of the policy for display
	in the Arctera Insight Classification.
Risk weight	Specify the risk weight for the policy. This is a mandatory field.
	By default, the risk weight value of all the custom policies and most of the built-in policies is configured as 1. Users can modify the risk weight value in the range of 0 to 10.
Tags	Nominates one or more tags that you want to apply to the items that match the policy conditions. Click the Tags field to choose from a list of the available tags.
Conditions	Specifies one or more conditions that an item must meet for the Arctera Insight Classification to consider it a match.
	Click + Condition to add a new condition for this policy.
	Click + Group to add a new group of conditions for this policy.
	See About policy conditions .

2. Test the policy. For more information on testing the policy,
3. After testing the policy, click **Save**.

Exporting or importing policies

The Arctera Insight Classification is available for use with multiple Arctera products. If you have several of these products and want to distribute the same policies across them all, you can export the policies from one instance of the Arctera Insight Classification and then import them into the others. The format in which the Arctera Insight Classification exports the policies is JavaScript Object Notation (JSON), an industry-standard format for exchanging data in a readable form.

You cannot export or import the built-in policies, but you can export or import any custom policies that you have created.

> “ ”

Note: Export and import functionalities are not supported for the Exact Data Match based policies and Sender and Recipient department based policy condition.

> “ ”

To export a policy

1. At the left of the Arctera Insight Classification, click **Policies**.
2. Select one or more policies that you want to export and then click **Export**.

Any custom patterns and tags that you have associated with the policies will automatically be exported as well.

1. Save the exported JSON file.

To import a policy

1. At the left of the Arctera Insight Classification, click **Policies**.
2. Click **Import**.
3. Select the JSON file that you want to import.

Resetting policies

If you make a mistake when you edit a built-in policy, you can reset it to its original settings. However, you cannot reset any custom policies that you have created.

To reset a policy

1. At the left of the Arctera Insight Classification, click **Policies**.
2. Select the policy that you want to reset and then click **Reset**.
3. Click **Yes** to confirm that you want to reset the policy.

Deleting policies

You cannot delete the built-in policies, but you can delete any custom policies that you have created.

To delete a policy

1. At the left of the Arctera Insight Classification, click **Policies**.
2. Select the policy that you want to delete and then click **Delete**.
3. Click **Yes** to confirm that you want to delete the policy.

Transparent policies

Arctera Insight Classification introduces new category of policies called Transparent policies. Transparent policies provide visibility into the policy keywords and logic. These policies are fully transparent allowing users to view the exact conditions that trigger a hit during classification. Thus, they provide full control and defensibility over classification.

Users can granularly modify, add, or remove classification criteria within a transparent policy by simply creating a copy and then editing the copy.

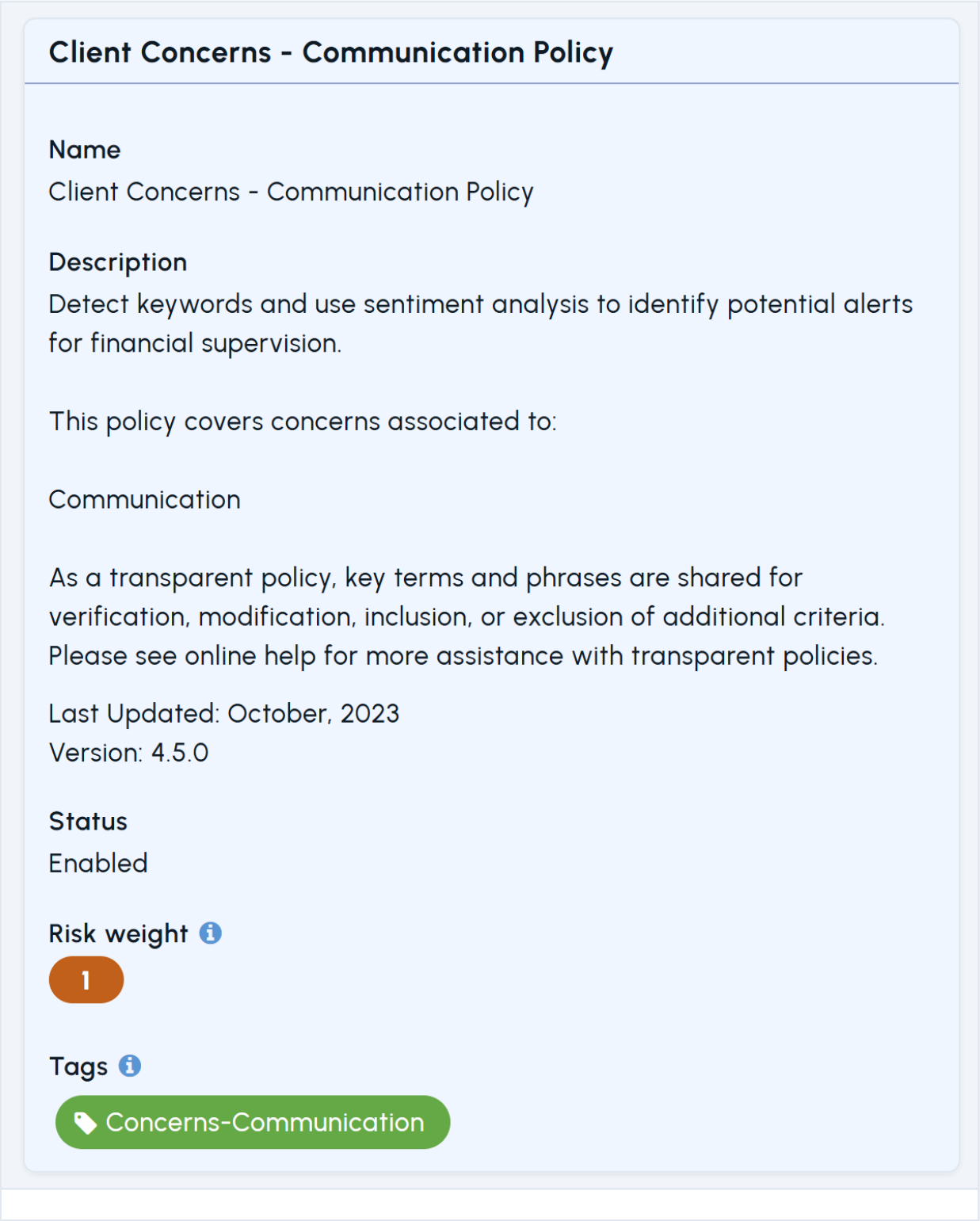
Users can use this editable copy to:

- Add nested conditions with \<Any of\>, \<None of\>, or \<All of\> criteria
- Remove keywords that trigger too many false positives
- Add keywords to reduce false negatives
- Add new criteria such as sentiment score, patterns, language, entity (if supported by the application) or regular expression
- Modify criteria to include min/max counts, change proximity criteria, specify string match, or case match

For more information on use cases via copied editable policies, See Creating policies.

For built-in policies belonging to the Transparent category, users can get following details about when the transparent policy logic was last updated by Arctera. Refer to the sample image below.

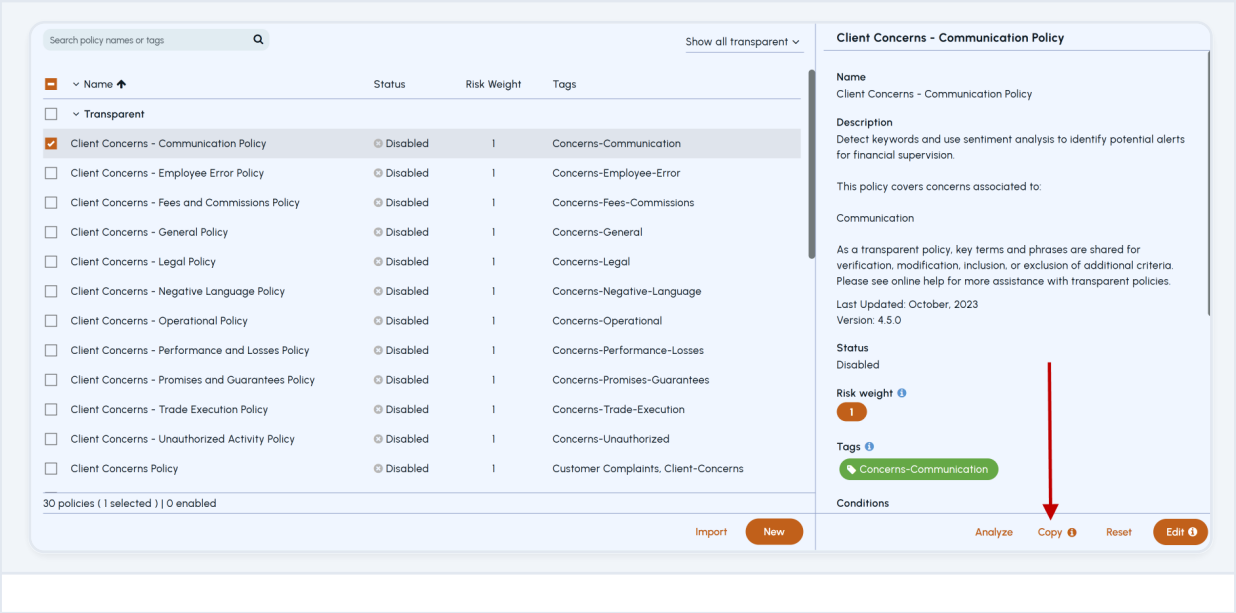
- Last Updated: Month and year of the last update to the policy.
- Version: A version of Arctera Insight Classification in which the policy is updated.



Creating a customized copy of transparent policies

To create a customized copy of transparent policy

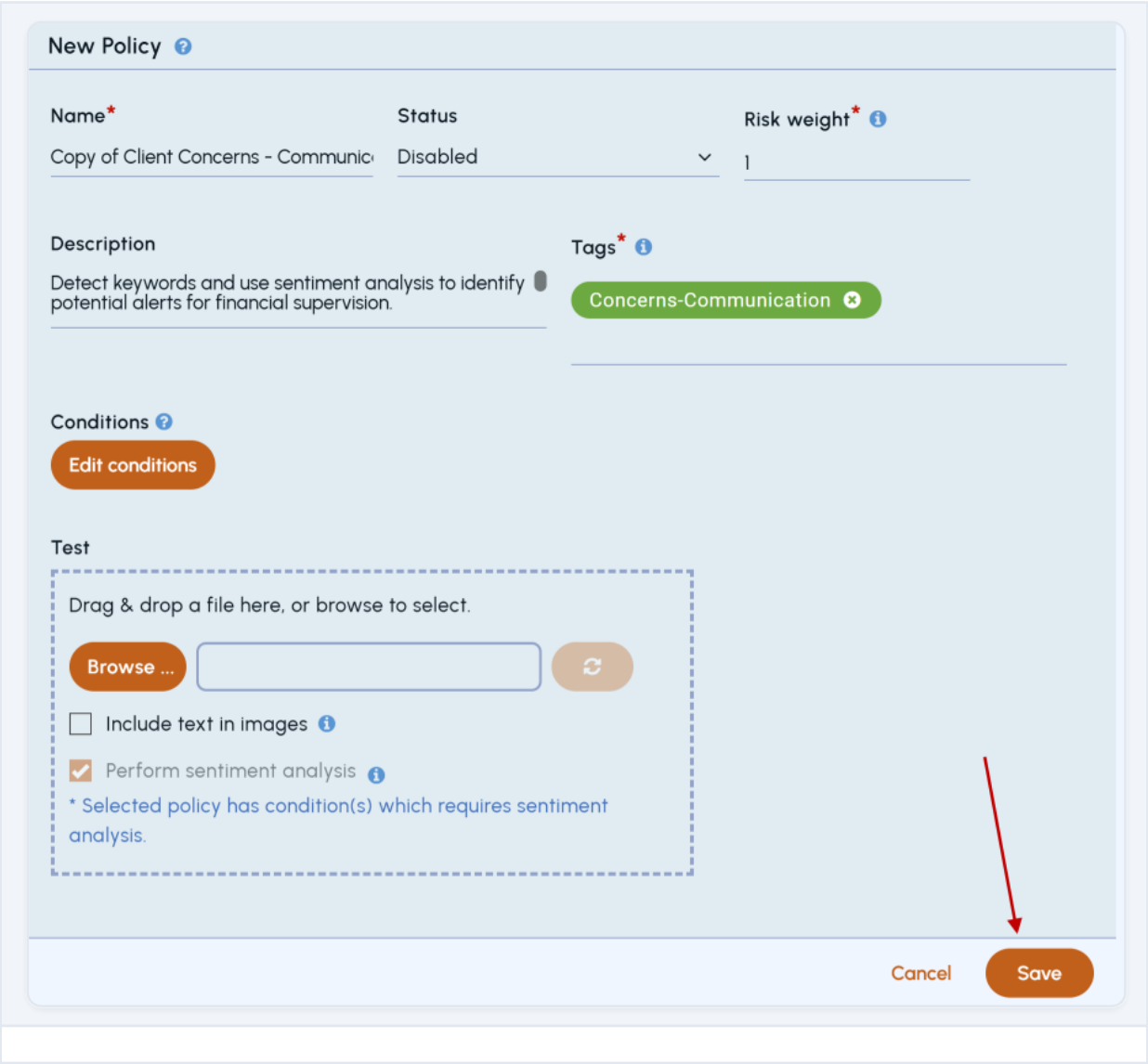
1. In the left navigation pane, click **Policies**.
2. In the top-right drop-down menu, select **Show all transparent**to view all the policies available under the**Transparent** category.
3. Search for and select a transparent policy that you want to copy, and click Copy.



4. In the New Policy dialog, set the fields as follows:

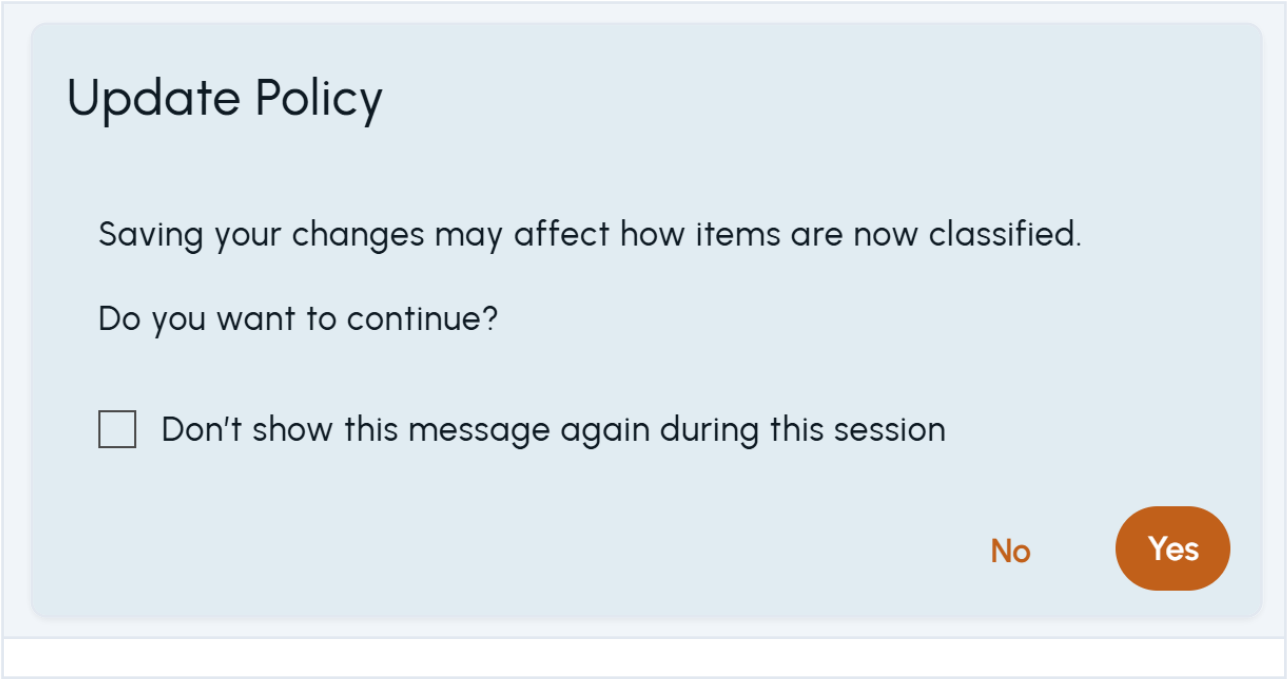
NAME	DISPLAYS THE NAME AS COPY OF THE POLICY YOU HAVE SELECTED. IF REQUIRED , YOU CAN CHANGE THE NAME.
Status	Displays the status same as the status of the policy you have selected. If required , you can change the status.
Description	Displays the description same as the description of the policy you have selected. If required , you can change the description. If you do not change the description, the last updated field do not show correctly.
Risk weight	This is a mandatory field.
	Displays the same risk weight for the copied policy. The by default risk weight value of the language and subscriptions policies is 1.

NAME	DISPLAYS THE NAME AS COPY OF THE POLICY YOU HAVE SELECTED. IF REQUIRED , YOU CAN CHANGE THE NAME.
	After you create a copy of a policy, the risk weight of the original policy gets propagated. You can change the risk weight value in the range of 0 to 10 if required.
Tags	Displays the tags same as the tags in the policy you have selected. If required , you can add or remove the tags. Add the tags to flag the results.
Conditions	Displays the conditions same as the conditions of the policy you have selected. If required , you can add or remove the tags.



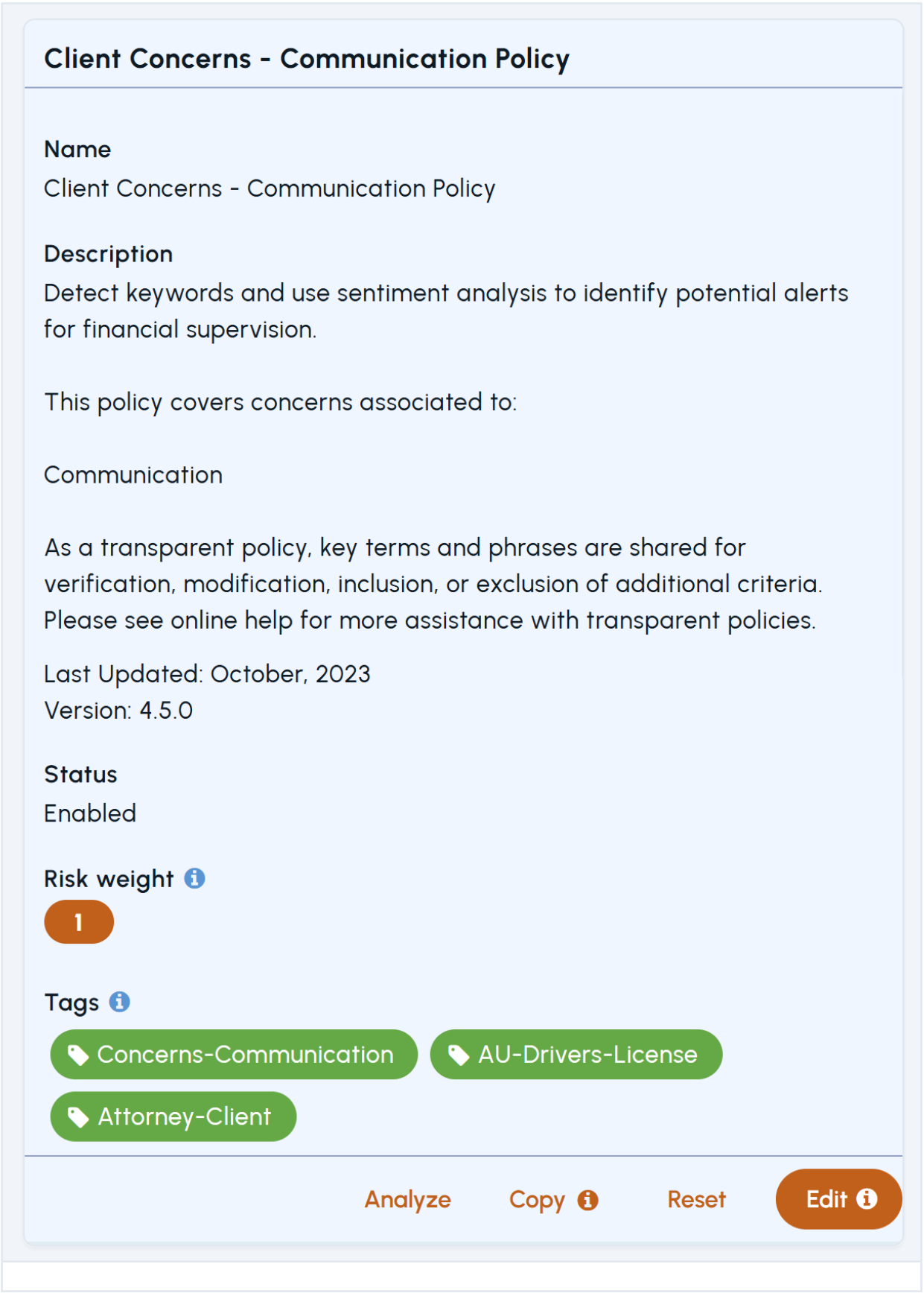
5. Click Save.

The copied policy gets created. However, sometimes if the following confirmation message appears, click **Yes**.



Following details (located on the Policy Details page) are provided only for custom policies that are created by copying transparent policies:

- Last updated information
- Origin details which shows a version and a link to the transparent policy from which the copy is created.



Origin policy information is useful if the original transparent policy logic is updated in subsequent releases. In such cases, this field will continue to show the version of transparent policy from which custom policy was created.

Microsoft Information Protection (MIP) Labels

Microsoft Information Protection (MIP) is a built in, intelligent, unified, and extensible solution designed to protect sensitive data. These are also referred to as Sensitive Labels. MIP technology integration allows adding labels to the documents. The labels might have a policy that restricts access to sensitive documents.

Configuring policies

After creating new tags, you need to associate those tags with a policy. To create new policy click New at the bottom of the Policies page. A new window appears.

- Add name of your choice
- Status should be Enabled.
- Add description (optional field).
- Specify the risk weight for this policy.
- Select the tags that you have created.
- Select custom string field from the Conditions drop-down. You can add multiple conditions as per your requirement.

- MIP labels are custom attributes of a file so you will have to provide information in specific format in the Text Box present in the Policy Manager. Each MIP Label name is associated to a specific GUID which must be included a policy condition.
    - For Office files with INTERNAL label, add the following condition:

```
custom:MSIP_Label_b43db875-2771-47af-9fc1-26bfde1b65ba_Name
"contains Text" INTERNAL
```

> “ ”

Note: You need to addcustom:before the label name and include GUID present in the label.

> “ ”

    - For PDF file with INTERNAL label, add the following condition:

```
pdf:docinfo:custom:MSIP_Label_b43db875-2771-47af-9fc1-26bfde1b65ba_N
ame

"contains Text" INTERNAL
```

> “ ”

Note: You need to addpdf:docinfo:custombefore the label name and include GUID present in the label.

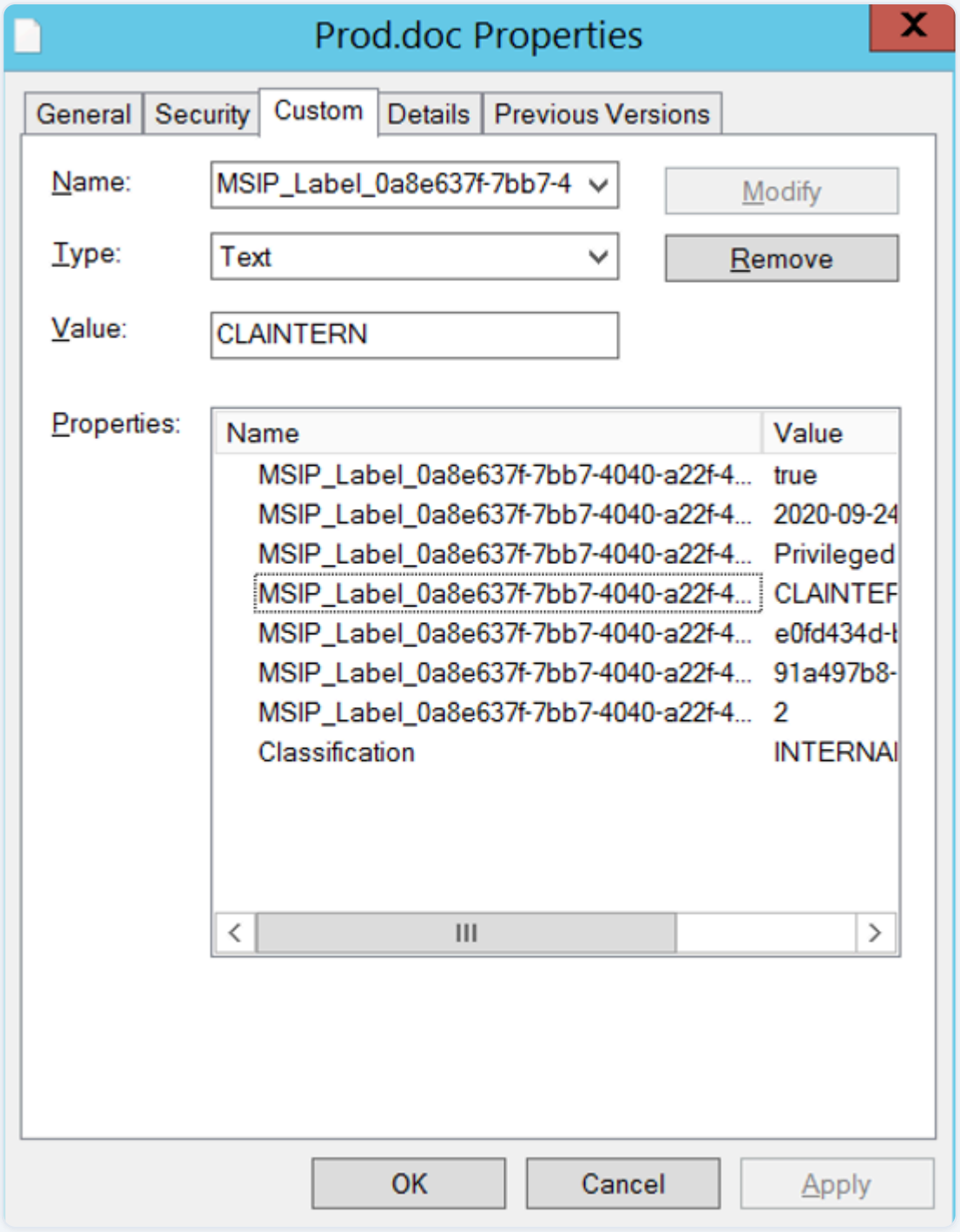
> “ ”

FILE TYPES	SAMPLE FORMAT	CONDITION	VALUE
Office Files	custom:MSIP_Label__Name	Contains Text	INTERNAL
PDF Files	pdf:docinfo:custom:MSIP_Label__Name	Contains Text	INTERNAL

Viewing MIP labels on a file system

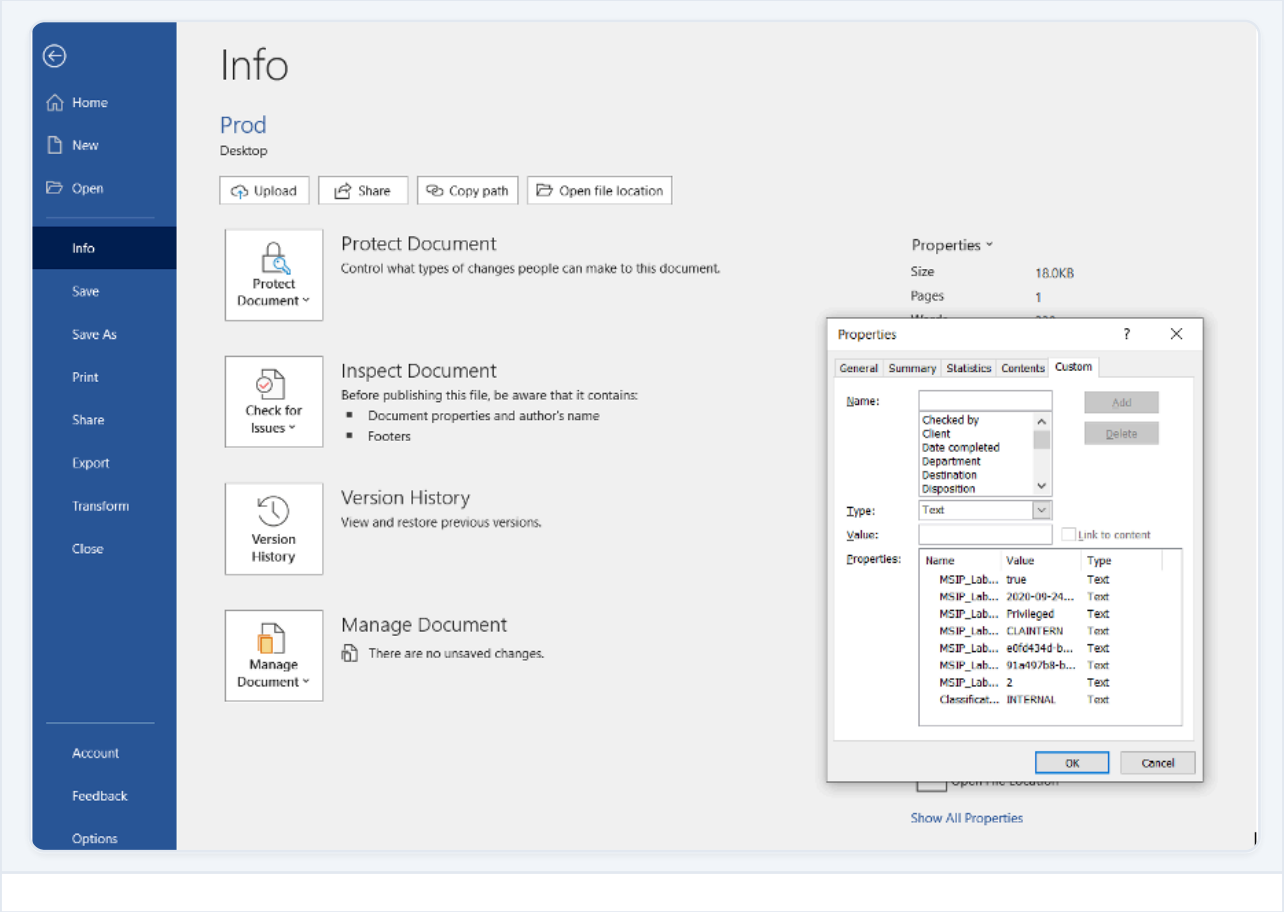
For legacy Microsoft Office (97-2003) files

To view labels for older legacy Microsoft Office files such as .doc, .xls, and .ppt, right-click on the file and then click View Properties, and then click the Custom tab.



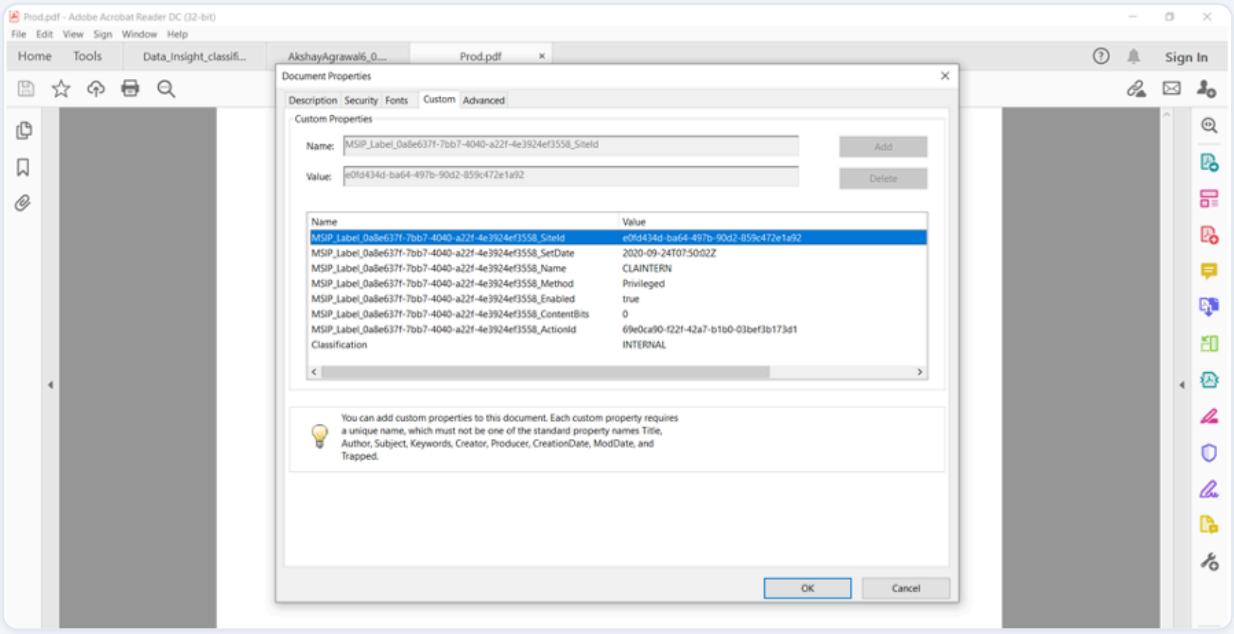
For Microsoft Office files in Open XML format

To view labels for newer Microsoft Office files such as .docx, .xlsx, and .pptx, open the file and then from the menu, click File > Info > Properties > Advanced Properties, and then click the Custom tab.



For PDF files

To view labels for Postscript Document Format (PDF) files, open the file using Adobe Reader and click File > Properties, and then click the Custom tab.


Prod.pdf - Adobe Acrobat Reader DC (32-bit)
File Edit View Sign Window Help
Home Tools
Prod.pdf
Sign In
Document Properties
Description Security Fonts Custom Advanced
Custom Properties
Name:
Value:
Add
Delete
Name
Value
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_SiteId
e0fd434d-ba64-497b-90d2-859c472e1a92
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_SetDate
2020-09-24T07:50:02Z
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_Name
CLAINTERN
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_Method
Privileged
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_Enabled
true
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_ContentBits
0
MSIP_Label_0a8e637f-7bb7-4040-a22f-4e3924ef3558_ActionId
69e0ca90-f22f-42a7-b1b0-03bef3b173d1
Classification
INTERNAL
You can add custom properties to this document. Each custom property requires a unique name, which must not be one of the standard property names Title, Author, Subject, Keywords, Creator, Producer, CreationDate, ModDate, and Trapped.
OK
Cancel

Patterns

This section includes the following topics:

About patterns

In the conditions of a policy, you can instruct the Arctera Insight Classification to look for one or more patterns in the items that it classifies. For example, here are the conditions for the built-in policy that is called "Ethics and Code of Conduct Policy":

Ethics & Code of Conduct

Name

Ethics & Code of Conduct

Description

Detect language which may be unethical or a violation of corporate code of conduct policy.

Supported Languages

English

Confidence

Level	Example
High	Curses, explicit or offensive language.
High	Insult keyword.
High	Bribery keyword.
High	Gift keyword AND inference of quid pro quo.
Medium	Curse keyword.
Medium	Words that can have non-violating meaning (e.g. proper anatomical names)

Policies

Used by 1 policy

Each condition looks for a match between the content of an item and an existing pattern: either "Individual Communication" or "Ethics & Code of Conduct". When an item matches both patterns, it meets the conditions of the policy.

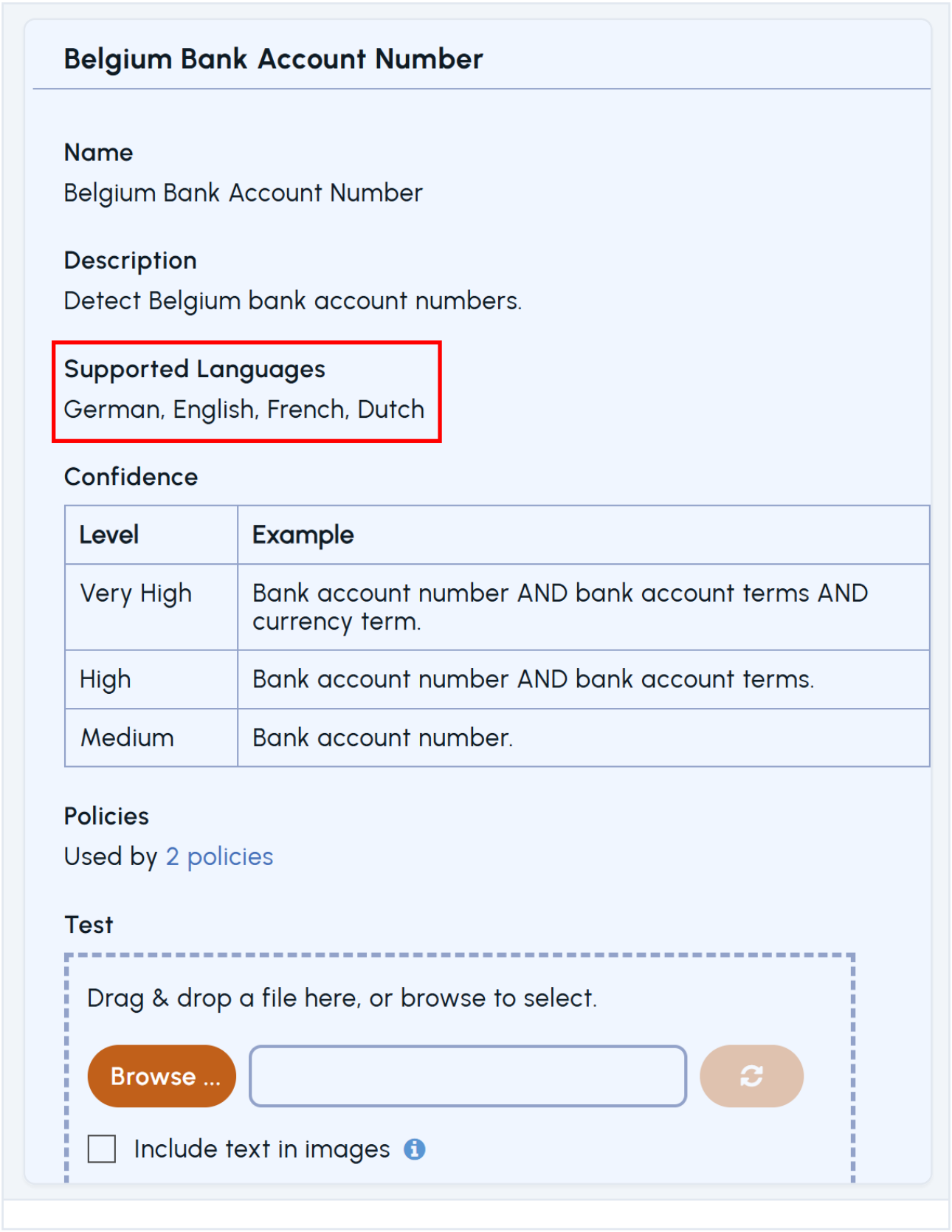
The built-in patterns that come with the Arctera Insight Classification use sophisticated algorithms to look for pattern matches and assign a confidence level. You can view the range of confidence levels for a pattern by selecting it in the item list. For example, the pattern "Credit/Debit Card Number" matches with low confidence if it finds a string of digits that conform to the format of a

credit card number, but it matches with high confidence if these digits are accompanied by credit-related keywords like "AMEX" and "Visa". When you create or edit policies, you can set the required confidence level for these pattern matches.

Each built-in pattern is used by at least one of the built-in policies, and you can incorporate the patterns in any custom policies that you create. You can also create custom patterns if the built-in ones do not meet your needs. However, it is important to note that these custom patterns are not as sophisticated as the built-in ones.

You can edit and delete custom patterns, but you cannot edit and delete built-in patterns.

Supported languages: starting with release 3.1.0, supported language information for built-in patterns is provided under Pattern Details page.



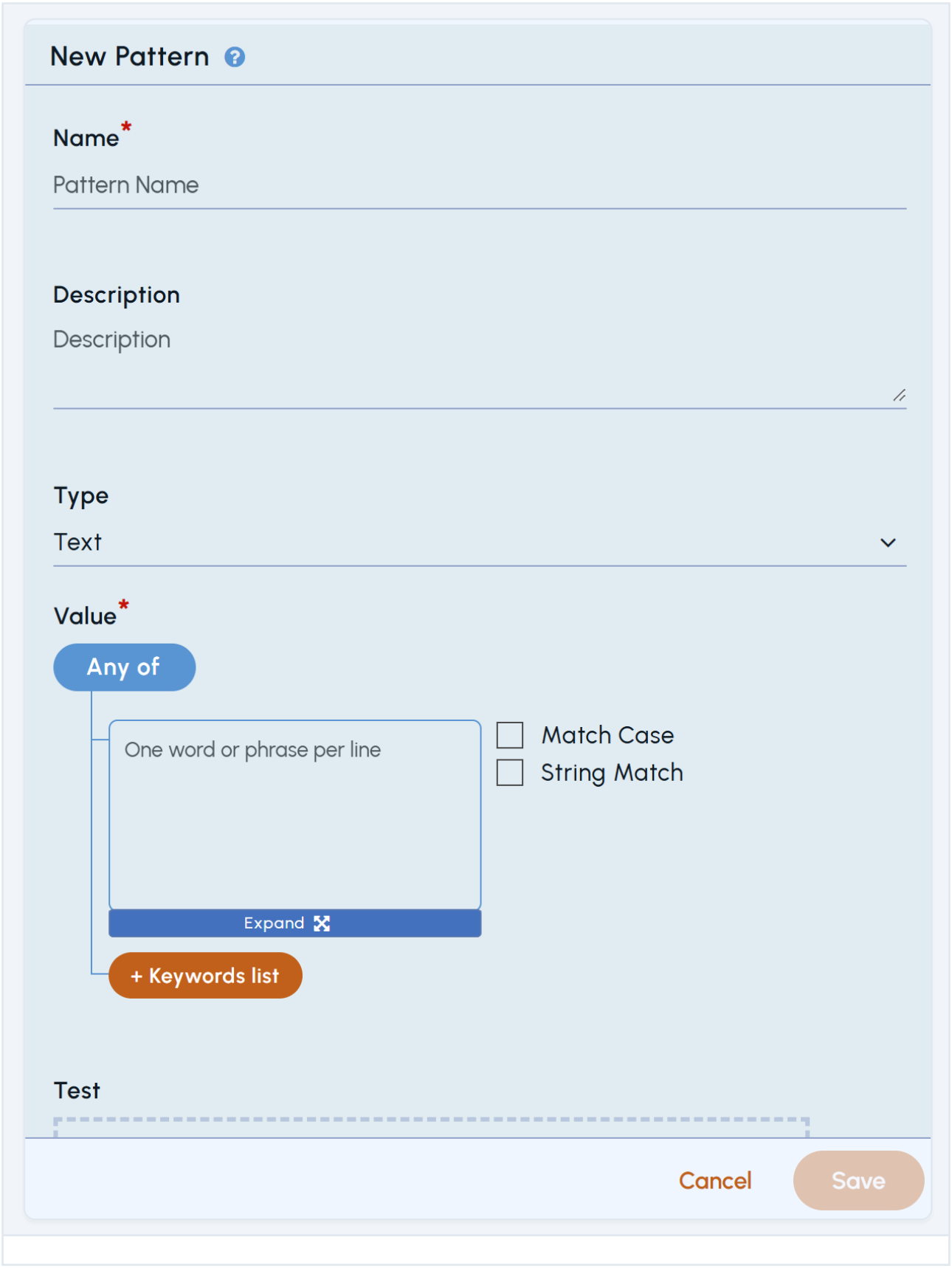
Creating or editing patterns

You cannot edit the built-in patterns, but you can edit any custom patterns that you have created.

To create or edit a pattern

1. At the left of the Arctera Insight Classification, click **Patterns**.
2. Do one of the following:
    - To create a pattern, click **New**.
    - To edit an existing pattern, select it and then click **Edit**.

The following diagram shows the **New Pattern**dialog with the pattern type as**Regular expression**.



Using Special Standalone Characters in Keyword-Based Patterns

When creating keyword-based conditions in Patterns, it is important to understand how the application handles special standalone characters like @, #, or $.

- Special standalone characters are not allowed by default in keyword conditions.
- If such characters are entered, the system will display a warning message.
- The pattern can still be saved, but any standalone special characters will be automatically removed in the back-end, unless the String Match option is explicitly selected.

To Retain Special Characters

If you want to retain standalone special characters in your keyword conditions, you must select the String Match option. This setting ensures that the characters are preserved exactly as entered and bypasses the automatic cleanup process.

Scenarios and Behavior

SCENARIOS	BEHAVIOR
Pattern with keyword condition containing special characters	Warning message is shown. Special characters will be removed unless String Match is selected.
Pattern with keyword condition and String Match selected	Warning message is not shown. Special characters are retained.

1. Set the fields as follows:

NAME	SPECIFIES THE PATTERN NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
Description	(Optional.) Provides a short description of the pattern for display in the Arctera Insight Classification.
Type	Specifies the pattern type.
	For a Text or Regular expression pattern, you must specify the value for which to look. The same guidelines that you must observe when you enter these values in a policy condition apply when you enter them as a pattern value.

NAME	SPECIFIES THE PATTERN NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
	See About policy conditions .
	Choose Similar document to find items that resemble a supplied template. For example, you can find completed forms by submitting the blank form as a template. Unlike Text and Regular expression patterns, you can set the required confidence levels for Similar document patterns when you incorporate them in a policy condition.
	The document similarity feature can find instances where users have created variants of the template document by adding, removing, or reordering paragraphs, sentences, or words. It can also find instances where users have changed individual words. However, the more extensive these word changes are, the less likely the Arctera Insight Classification is to find a match.
	You must choose the required similarity mode: Full or Section . In Full mode, the Arctera Insight Classification compares the template document in its entirety with other documents in their entirety. This mode is useful when looking for instances where users have altered the template document in places without greatly affecting its overall size. In Section mode, the Arctera Insight Classification looks for instances where the content of the template document appears as one section within a larger document.
	To submit the template document, click Browse and then select the required document.

NAME	SPECIFIES THE PATTERN NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 100 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS.
	Choose Exact Data Match to find match of one or more specific values in an item. Exact Data Match (EDM) gives precise control over the data classification process by setting more granular level data match conditions and provides less false positives.
	With EDM you can create patterns using database records.
	See “To create an Exact Data Match based pattern” .

2. Test the pattern by clicking **Browse** and then choosing a document that ought to match it.

Select the **Perform sentiment analysis** checkbox for performing sentiment analysis on the selected item to determine whether the sentiment associated with the item is positive, negative, or neutral.

> “ ”

Note: If sentiment analysis fails, the classification process will continue without evaluating the sentiment analysis conditions within the policies. As a result, hits and matches based on the sentiment condition will be affected.

> “ ”

Select the **Include text in images** checkbox for extracting information from images and performing classification using Optical character recognition (OCR).

> “ ”

Note: TheInclude text in imagescheckbox appears only when the Tesseract software is installed on the system where Arctera Insight Classification is running.

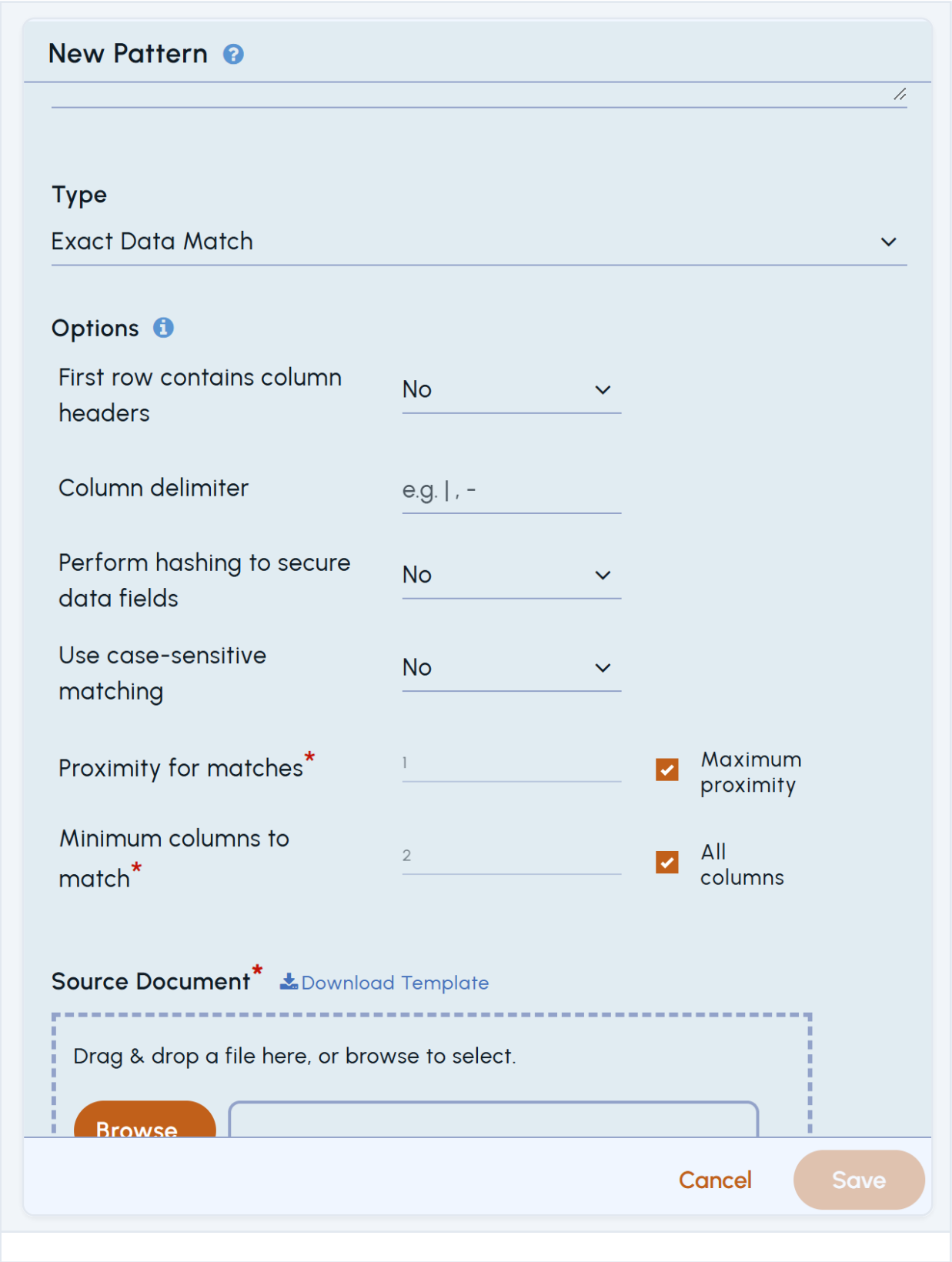
> “ ”

After a few moments, the Arctera Insight Classification indicates whether it has found a match. When this is the case, you can click **Show details** to see the matching text and confidence levels.

1. Click **Save**.

To create an Exact Data Match based pattern

1. Follow the initial steps for creating a pattern as described earlier.
2. In the Type box, click to select **Exact Data Match**.



3. Specify the following configuration options:

FIRST ROW CONTAINS COLUMN HEADERS	SELECT YES IF THE FIRST ROW IN THE SOURCE DOCUMENT CONTAINS THE NAMES OF EACH FIELD. IF SELECTED, CONTENT OF FIRST ROW FROM THE SOURCE DOCUMENT WILL NOT BE CONSIDERED FOR RULE GENERATION.
	SELECT NO IF THE FIRST ROW IN THE SOURCE DOCUMENT DO NOT CONTAIN THE NAMES OF EACH FIELD.
Column delimiter	This is an optional field. It specifies the delimiter character that separates each column/field in the data file.
	Note:
	- Delimiter can be any single special character. For example, a comma(,), pipe (|), a space, and so on.
	- If the source document contains only a single column/field, you can set any delimiter character that is not present in file.
	- Delimiter must be a single character value.
Perform hashing to secure data fields	Select Yes if the generated rule used for creating EDM pattern need to be hashed to protect the data. The data fields are hashed using hashing algorithm SHA256 when storing them in the generated classification rule.
	Note: Classification performance drops if hashing is used while creating Exact Data Match pattern.
Use case-sensitive matching	Select Yes if the match needs to be case sensitive.
Proximity for matches	

FIRST ROW CONTAINS COLUMN HEADERS	**SELECT YES IF THE FIRST ROW IN THE SOURCE DOCUMENT CONTAINS THE NAMES OF EACH FIELD. IF SELECTED, CONTENT OF FIRST ROW FROM THE SOURCE DOCUMENT WILL NOT BE CONSIDERED FOR RULE GENERATION.**
	SELECT NO IF THE FIRST ROW IN THE SOURCE DOCUMENT DO NOT CONTAIN THE NAMES OF EACH FIELD.
	Specifies the distance between two columns or fields in number of characters for a match to be considered valid. Valid values are greater than 0.
	Note:
	- If source document contains only a single column/field, proximity value should be set to 1.
	- The generateRulePack API that generates classification rule uses "From the first condition option" proximity option. "Sliding Window" proximity option is not supported for Exact Data Match.
	Example:
	With proximity = 20, if the CSV source document content is as follows,
	Goodbye, Hello
	and test document content is,
	… You say Goodbye and I say Hello …
	Here, between the two words "Goodbye" and "Hello" the proximity is 19 characters. The matched words are within the set range of proximity value, that is 20 characters.

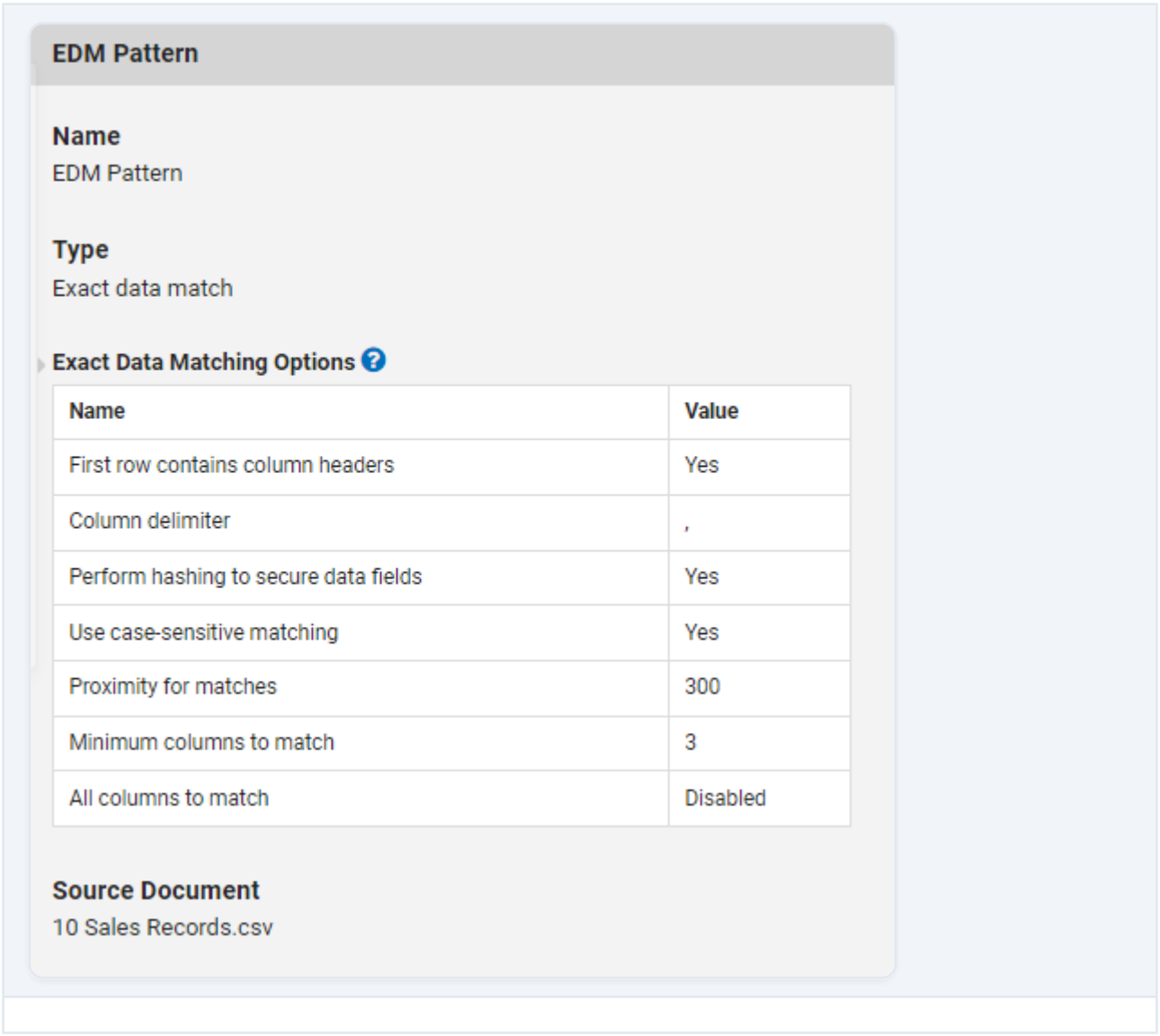
FIRST ROW CONTAINS COLUMN HEADERS	**SELECT YES IF THE FIRST ROW IN THE SOURCE DOCUMENT CONTAINS THE NAMES OF EACH FIELD. IF SELECTED, CONTENT OF FIRST ROW FROM THE SOURCE DOCUMENT WILL NOT BE CONSIDERED FOR RULE GENERATION.**
	SELECT NO IF THE FIRST ROW IN THE SOURCE DOCUMENT DO NOT CONTAIN THE NAMES OF EACH FIELD.
	Therefore, Arctera Arctera Insight Classification will show a match.
Minimum columns to match	Specifies the minimum number of columns that should match to trigger a result. Note that matching of the first column is compulsory regardless of the value specified in Minimum columns while creating EDM pattern.
	Note: Minimum columns field will be ignored if All columns checkbox is selected.
All columns	Select this checkbox if all columns/fields in source document need to match to trigger a result.

4. Under the **Source Document**section, browse to select the EDM source file based on which you want to create the classification rule.**Note** :

- EDM source document must be of type CSV or TXT (plain text only)
- Maximum document size is configurable. Recommended size is 5 MB.
- CSV document with fields quoted is not supported

1. Click **Save**.

The created EDM pattern shows the user configured exact data matching options. The source document name is retained for pattern, but its location or direct link is not provided. See the following image.



You can use the EDM pattern created to:

- Enhance an existing policy
- Create a new policy

For more information, See About policy conditions.

Known issue while editing EDM patterns

While editing EDM patterns, updating the pattern name or description may fail due to an internal system error. If you experience this issue, contact your system administrator or Arctera support.

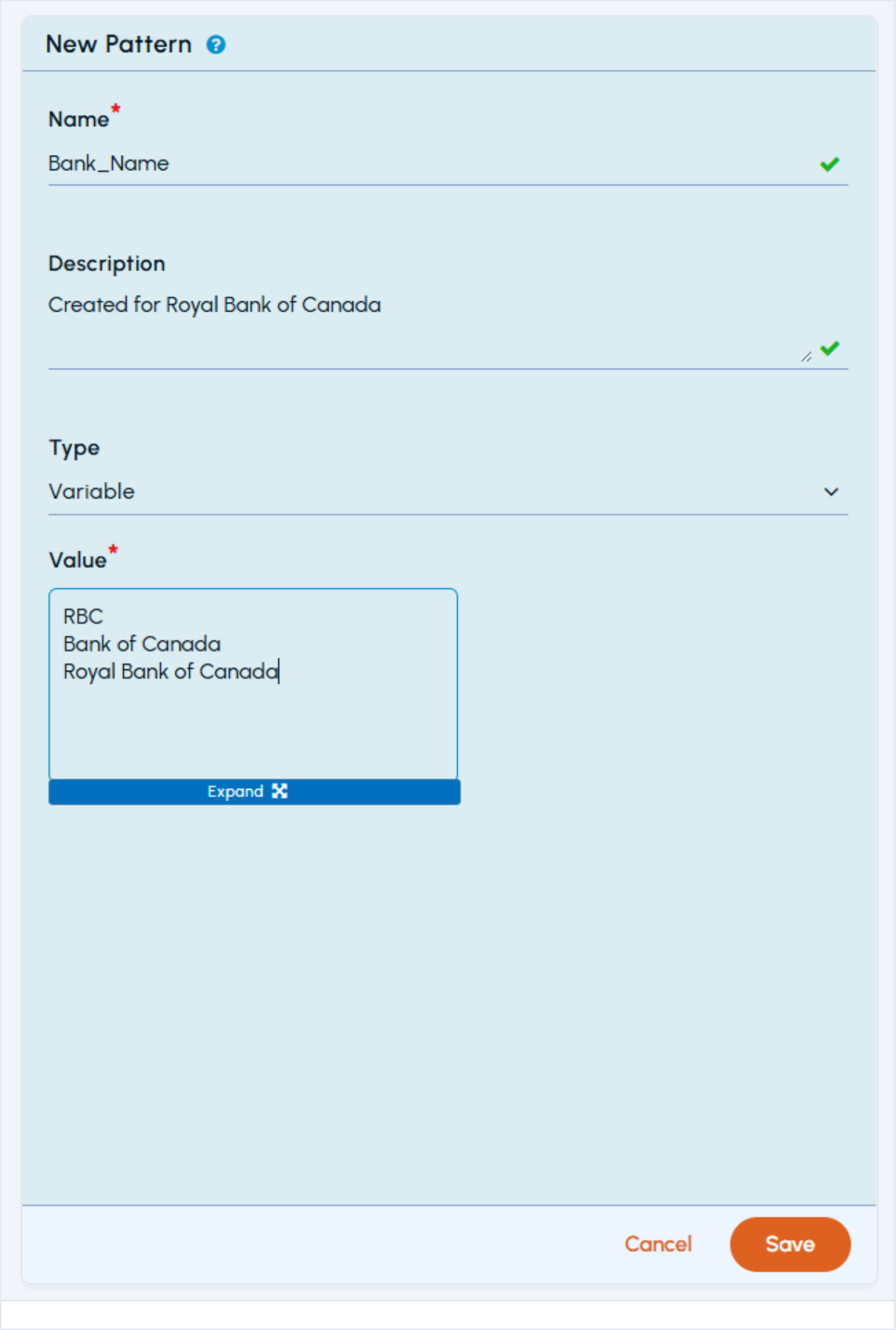
Variable

Variable Support allows you to insert dynamic values into text conditions across policies and patterns using simple placeholders such as . During classification, the system automatically evaluates the content against all defined values of the variable. As a result, any content containing

RBC, Bank of Canada, or Royal Bank of Canada will be classified, without the need to create separate conditions for each variation.

To create the basic variable,

1. Follow the initial steps for creating a pattern as described earlier.
2. In the Type box, click to select Variable.
3. Define the values as per your requirement. For example, if you want to define a new pattern for identifying a Royal Bank of Canada as a bank name, add Bank_Name in the Name field and add relevant description. In the Value field, add values like RBC, Bank Of Canada, Royal Bank of Canada or the values you want to get a match on.

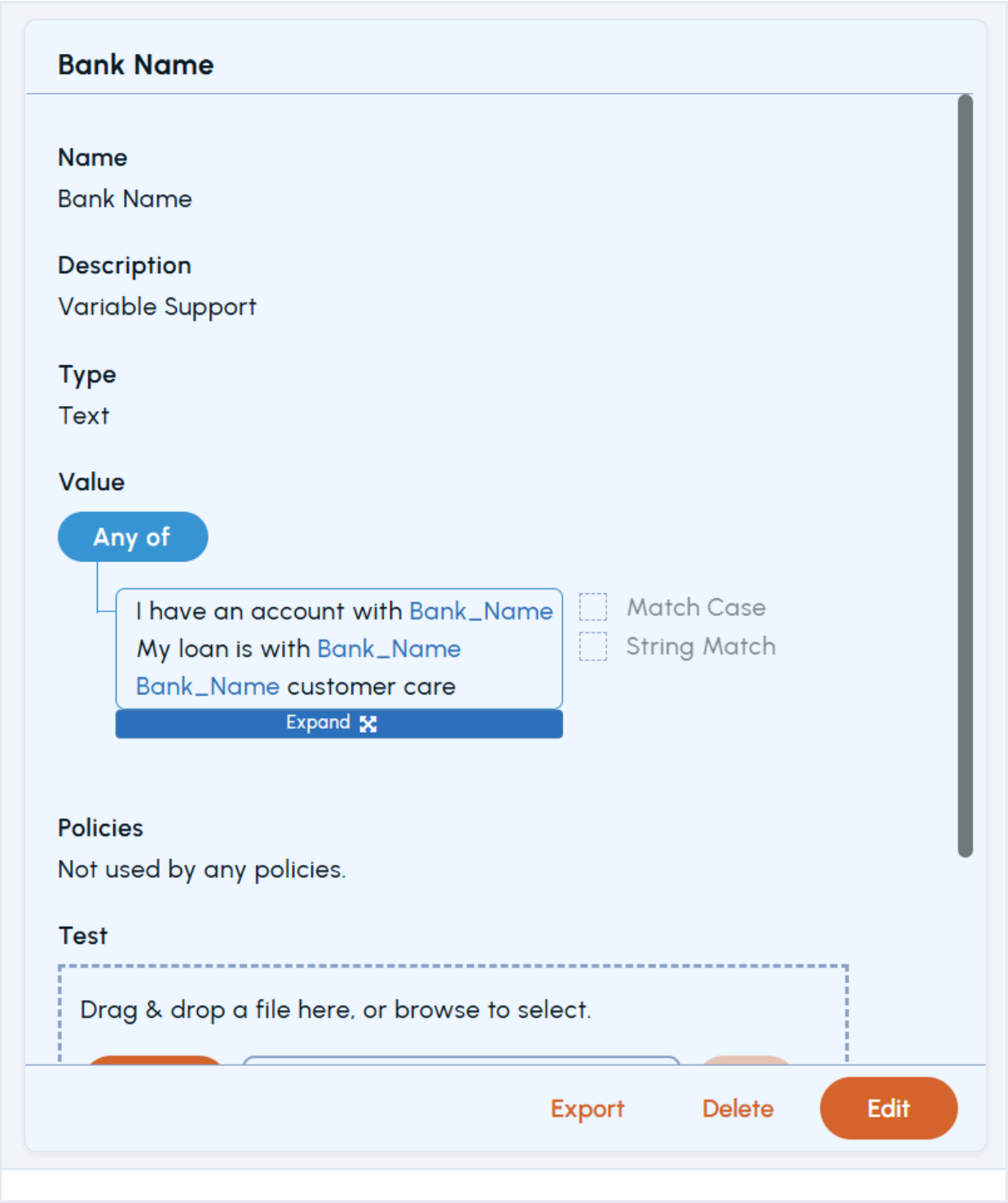


4. Click Save to add this variable.

After saving the pattern, the system checks the content during classification and looks for any occurrence of the values you added to the pattern. If the content contains one or more of these values, it is identified as a match and classified accordingly. You can reuse this approach to create additional patterns, such as Department Names, Designations, Agency Names, and more.

To use the created variable in a pattern,

1. Follow the initial steps for creating a pattern as described earlier.
2. In the Value box, type and list of all available variables appear on the page. All the values you have entered while creating the variable will be matched while classifying the content.
3. Click Save to add this pattern.
4. In View mode, variables in a text condition appear as links. You can click these link to navigate to the particular pattern.



You can use multiple variables in a condition; however, using a large number of variables may impact classification performance. It is recommended to use only the minimum number of variables required. This flexibility still allows you to define and reuse dynamic values across different text conditions.

Exporting or importing patterns

Just as you can export your custom policies and tags from one Arctera Arctera Insight Classification environment and then import them into another, you can also export and import custom patterns. You cannot export or import the built-in patterns.

> “ ”

Note: Export and Import functionality are not supported for the Exact Data Match based patterns.

> “ ”

To export a pattern

1. At the left of the Arctera Insight Classification, click **Patterns**.
2. Select one or more patterns that you want to export and then click **Export**.
3. Save the exported JSON file.

To import a pattern

1. At the left of the Arctera Insight Classification, click **Patterns**.
2. Click **Import**.
3. Select the JSON file that you want to import.

Deleting patterns

You can delete the custom patterns that you have created, provided that they are not in use in any policies. You cannot delete the built-in patterns.

To delete a pattern

1. At the left of the Arctera Insight Classification, click **Patterns**.
2. Select the pattern that you want to delete and then click **Delete**.
3. Click **Yes** to confirm that you want to delete the pattern.

> “ ”

Note: If you delete an EDM based pattern, its associated rule also gets deleted.

> “ ”

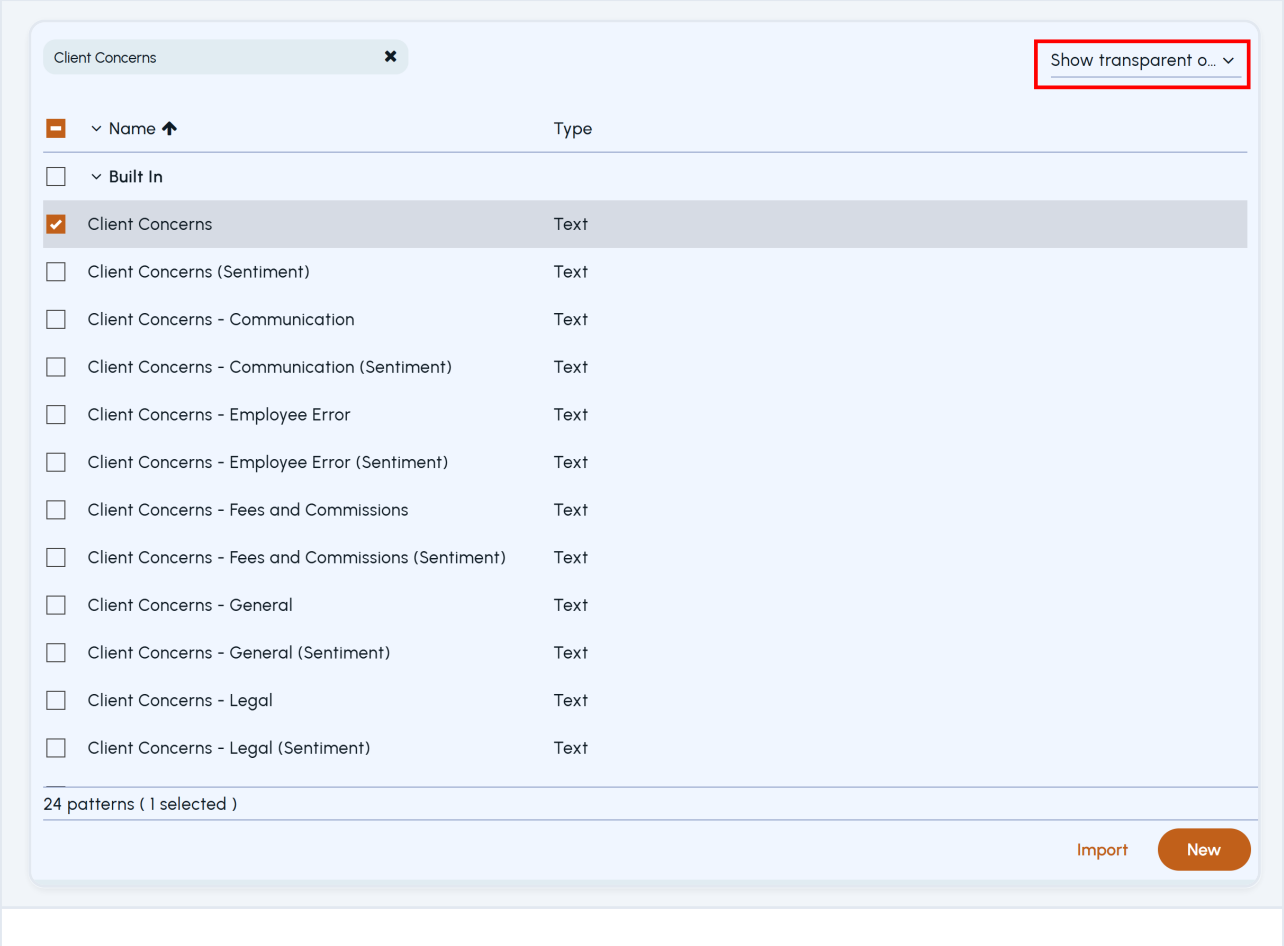
Transparent patterns

Transparent patterns are new built-in patterns that provide users with visibility into the internal logic of the pattern. However, customers cannot modify these patterns. Instead, they can create a copy of a pattern, and modify it as per requirement and save it as a new transparent pattern. Transparent patterns provide better control and defensibility over classification.

Transparent patterns display the following additional details:

- Last Updated: Month and year of the last update to the pattern.
- Version: A version of Arctera Insight Classification in which the pattern is updated.

Refer to the following sample screen of transparent pattern details.



For more information on these patterns, refer to *Arctera Insight Classification Policies and Patterns Guide*.

Creating a customized copy of transparent patterns

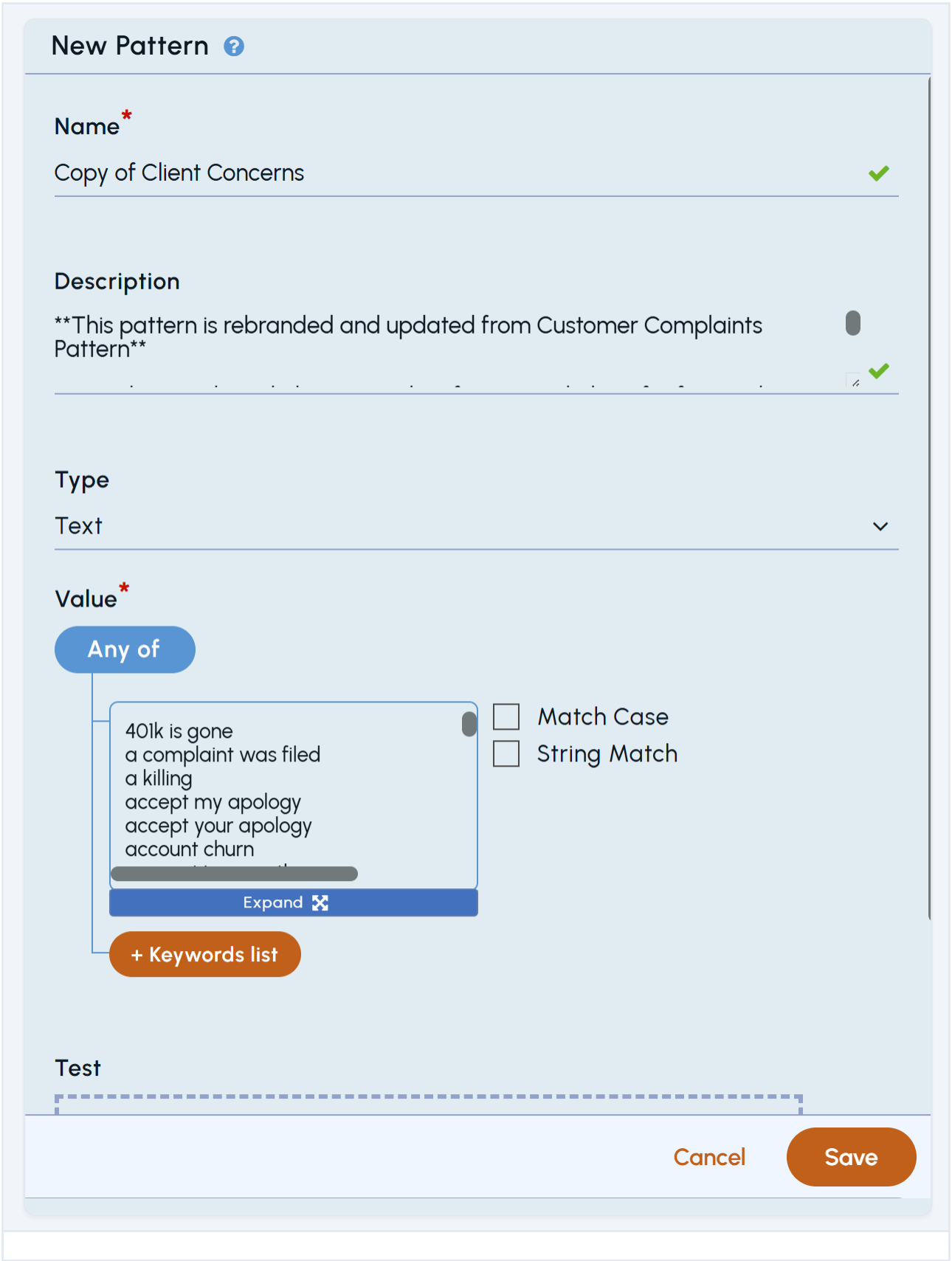
To create a customized copy of a transparent pattern

1. In the left navigation pane, click **Patterns**.

The application displays all the available patterns.

1. Search for and select the transparent pattern you want to copy.

Alternatively, in the top-right drop-down list, select **Show transparent only** to view all the available transparent patterns.



The application displays details of the selected transparent pattern in the right-side pane.

1. Click **Copy**.

The **New Pattern** dialog box appears.

1. Modify the required details (such as name, description, and internal logic of pattern), and click **Save**.
2. Ensure that the pattern is added to the transparent patterns list, and the following details appear. Copied pattern will not be a part of transparent patterns. Instead, it appears under custom patterns.
    - **Last Updated** : Month and year of creating a copy of a pattern.
    - Origin details that shows a version and a link to the transparent pattern from which the copy is created.

Tags

This section includes the following topics:

About tags

Every Arctera Insight Classification policy is associated with one or more tags. When an item that you have submitted for classification matches the conditions of a policy, the Arctera Insight Classification assigns the associated tags to the item. For example, the Arctera Insight Classification assigns the tag "Corporate-Ethics" to items that match the built-in policy "Ethics and Code of Conduct Policy".

Preset tags

Following preset tags for classification responses are available to identify failure types during classification:

- SENTIMENT-NOT-DETECTED

If sentiment analysis evaluation fails, classification will continue bypassing the sentiment failure and return *SENTIMENT-NOT-DETECTED*tag in POLICY tag-set. In addition,*SENTIMENT*tag with sentiment score*will NOT* be returned in response.

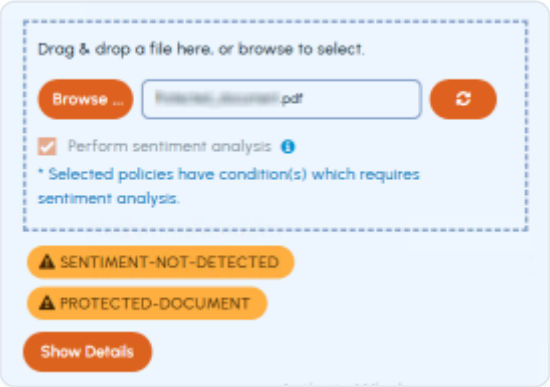
> “ ”

Note: Only evaluation against policy with sentiment conditions will be skipped.

> “ ”

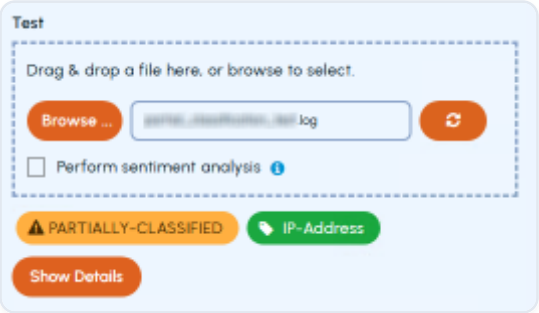
- PROTECTED-DOCUMENT

If document sent for classification is either password protected or encrypted, classification will be marked as Failed and *PROTECTED-DOCUMENT* tag will be returned in POLICY tag-set.



- PARTIALLY-CLASSIFIED

If partial classification is required in the classification request and the document is actually classified partially, classification will be successful and *PARTIALLY-CLASSIFIED* tag will be returned in POLICY tag-set.



Creating or editing tags

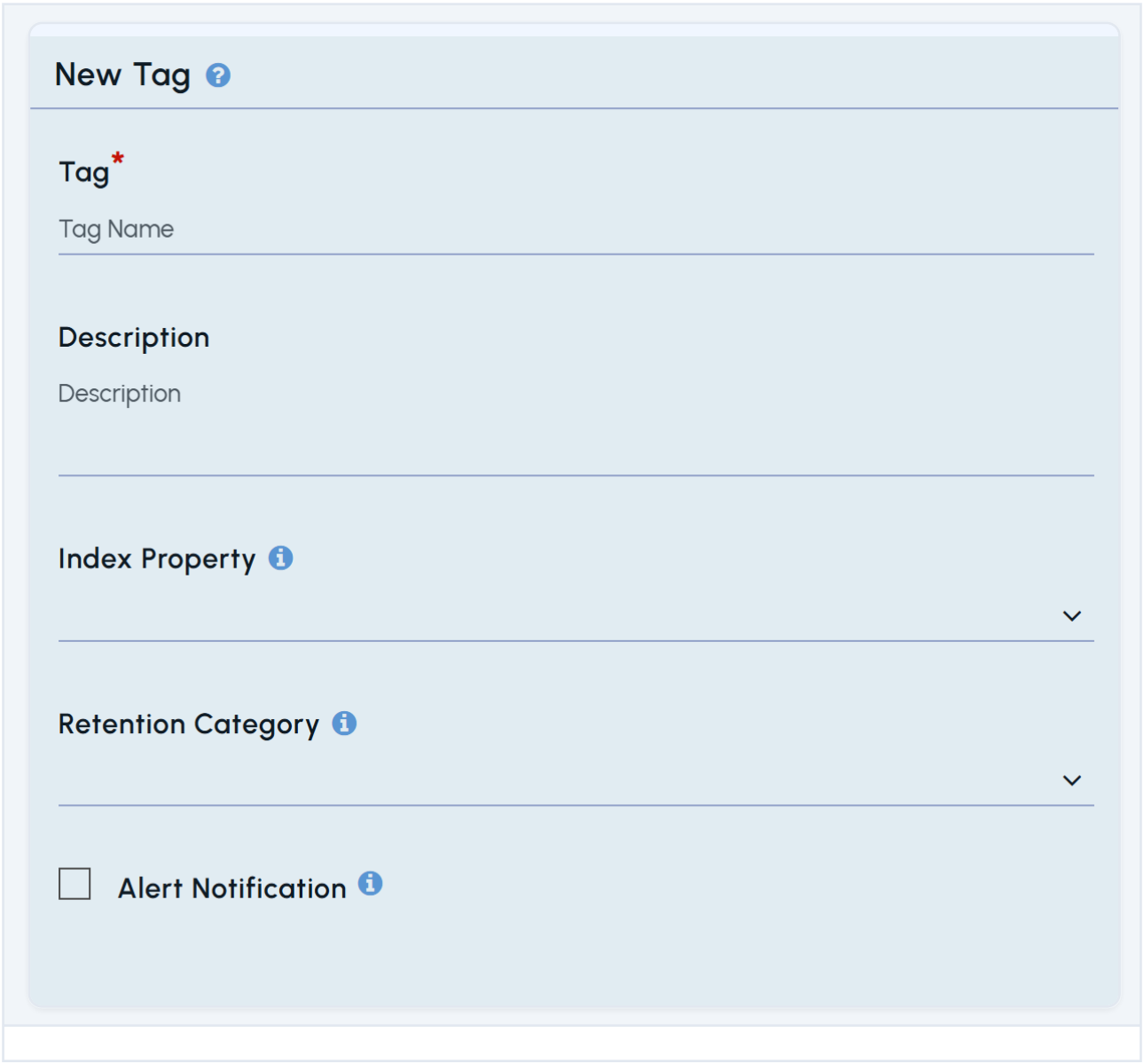
The Arctera Insight Classification comes with a large number of built-in tags, but you can create custom tags if the built-in ones do not meet your needs.

You cannot edit the built-in tags, but you can edit the descriptions of the custom tags.

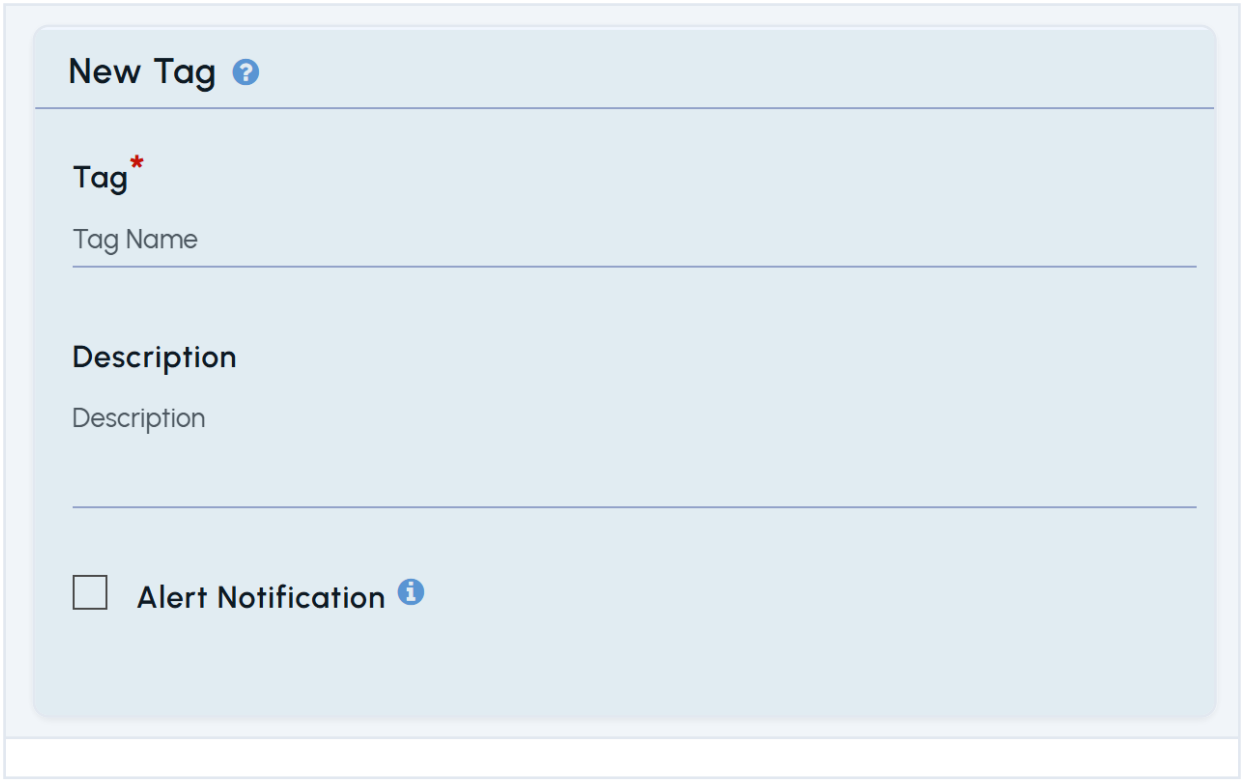
To create or edit a tag

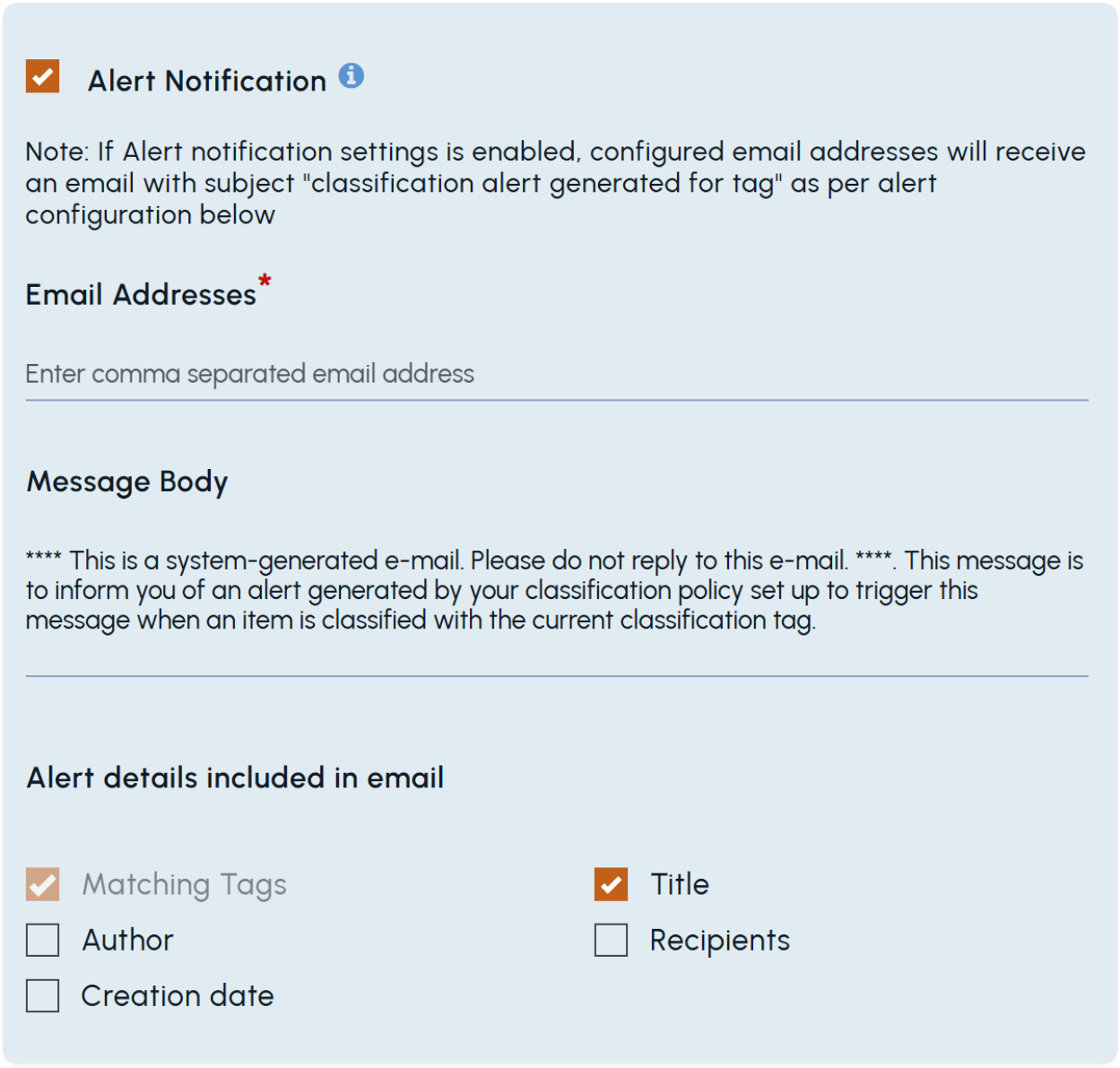
1. In the left pane of the Arctera Insight Classification , click Tags.
2. Do one of the following:
    - To create a tag, click New.
    - To edit an existing tag, select it and then click Edit.

Depending on the property definition value setup in yaml, specific to the product, one of the following pop-up will appear on the screen.



1. If you have opted for email alert notification through yaml and check the box on this screen, alert notification details will appear on the screen.


New Tag
Tag*
Tag Name
Description
Description
Alert Notification



2. Set the fields as follows:

TAG	SPECIFIES THE TAG NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 30 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS. HOWEVER, THE NAME MUST NOT INCLUDE THE FOLLOWING CHARACTERS\:
	& : / \ % + < > ?
	IF YOU ARE EDITING AN EXISTING TAG, YOU CANNOT CHANGE ITS NAME.
Description	(Optional.) Provides a short description of the tag for display in the Arctera Insight Classification.

TAG	SPECIFIES THE TAG NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 30 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS. HOWEVER, THE NAME MUST NOT INCLUDE THE FOLLOWING CHARACTERS\:
	& : / \ % + < > ?
	IF YOU ARE EDITING AN EXISTING TAG, YOU CANNOT CHANGE ITS NAME.
Index Property	(Optional.) Select one applicable option from the drop-down.
Retention Category	(Optional.) Select one applicable option from the drop-down. None option will be selected by default.
	The Retention Category
	is a managed tag
	associated with a
	retention policy.
	Managed tags can be
	created and managed in
	the Alta View under
	Configuration ->
	Managed Tags . This
	category helps in
	managing the retention
	of archived items based
	on predefined policies.

TAG	SPECIFIES THE TAG NAME. THE NAME MUST BE UNIQUE, AND IT CAN CONTAIN UP TO 30 ALPHANUMERIC, SPACE, AND SPECIAL CHARACTERS. HOWEVER, THE NAME MUST NOT INCLUDE THE FOLLOWING CHARACTERS\:
	& : / \ % + < > ?
	IF YOU ARE EDITING AN EXISTING TAG, YOU CANNOT CHANGE ITS NAME.
	By assigning a retention
	category to a tag, you
	ensure that tagged items
	comply with the retention
	rules linked to that
	category. This ensures
	consistency and
	compliance with
	organizational data
	retention requirements
Email Addresses	Add comma separated list of email addresses, who will receive email alerts when an item is classified with this tag.
Message Body	(Optional.) Update this field if you want to add specific instructions for the email recipients.
Alert details included in email	Alert details that will be included in the alert email. You can check or uncheck options except Matching Tags , which selected by default.

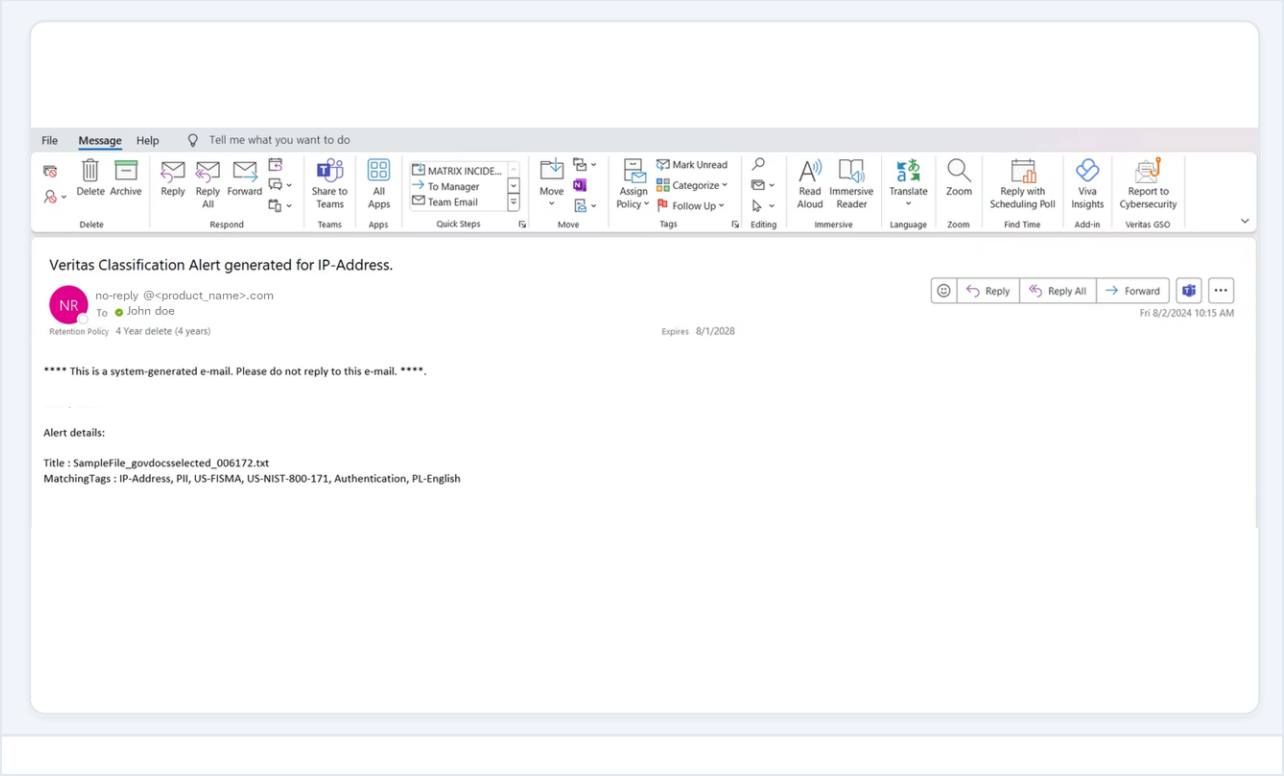
> “ ”

Note: Author, Creation date, Title and Recipients fields will be available for selection. If you want any additional fields, edit the yaml and set the values accordingly.

> “ ”

3. Click **Save** to complete the process.

If you have opted for email notification while creating or editing a tag, an email will be sent to the specified users whenever a tag is used for classification.



Exporting or importing tags

You can export your custom policies and patterns from one Arctera Insight Classification environment and then import them into another. You can export and import custom tags but you will not be able to export index property, retention category, and email alert details. In addition to that, you will not be able to export or import the built-in tags as well.

To export a tag

1. At the left of the Arctera Insight Classification, click **Tags**.

2. Select one or more tags that you want to export and then click **Export**.
3. Save the exported JSON file.

To import a tag

1. At the left of the Arctera Insight Classification, click **Tags**.
2. Click **Import**.
3. Select the JSON file that you want to import.

Deleting tags

You cannot delete the built-in tags, but you can delete any custom tags that you have created. However, you must first ensure that no policies use these tags.

To delete a tag

1. At the left of the Arctera Insight Classification, click **Tags**.
2. Select the tag that you want to delete and then click **Delete**.
3. Click **Yes** to confirm that you want to delete the tag.

Analyze

This section includes the following topics:

About analyze

Before you put any policies into effect, you may want to evaluate them against a sample set of your organization's content. Not only does the Analyze feature let you do this, but it also generates a detailed Risk and Compliance Analysis report of its findings. Use this report to determine whether the built-in policies meet your needs or whether extra, custom policies are also required.

Note that the Analyze feature skips certain types of files, such as image and audio files, because they do not have any textual content. It also skips any files that are larger than 10 MB.

Analyzing sample content for policy matches

To analyze sample content for policy matches

1. Do one of the following:
    - To find the content that matches one or more specific policies, select these policies in the policy list and then click **Analyze**.
    - To find the content that matches any of the available policies, or a group of related policies such as the GDPR and Personal Data policies, click **Analyze** at the left of the Arctera Insight Classification.
2. In the Policies page of the Analyze wizard, verify that you have selected the required option and then click **Next** to go to the Folder page.
3. In the Folder page, drag and drop the local or network folder containing the files that you want to analyze, and then click **Next** to go to the Analyze page.

A maximum of 20,000 files can be analyzed.

1. In the Analyze page, wait for the analysis to complete, and then click respective tabs to view more information on potentially sensitive files.

Click **Next** to go to the Summary page.

The analysis may take some time if there are a large number of files or they contain a lot of text.

1. In the Summary page, click **Download (CSV)** to download the Compliance Analysis report in comma-separated value (CSV) formats.

Analyze page

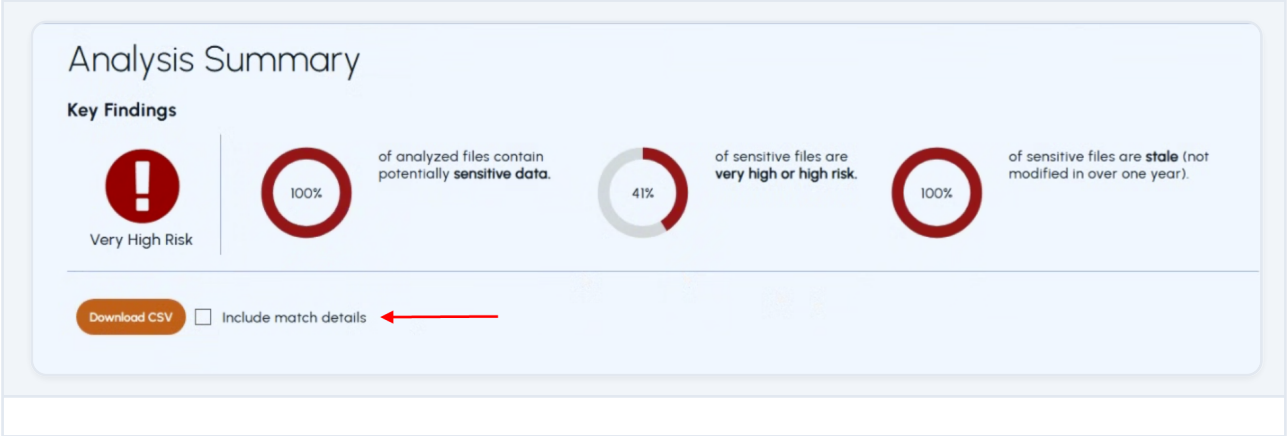
The Analyze page provides information on potentially sensitive files through different tabs as shown in following table:

TAB	DETAILS FOR ANALYZED SENSITIVE FILES
Overview	- A donut chart that shows sensitive and non-sensitive files
	- A donut chart that shows the risk level associated with all sensitive files, excluding 'No Risk' files, based on their risk level.
	- A bar chart that shows the most common file extensions within sensitive files
	- A bar chart that shows the top 10 patterns or policy conditions based on hits for sensitive files
	- A bar chart that shows the last modified date for sensitive files
	- A bar chart that shows the file sizes for analyzed sensitive files
	- A bar chart that shows sentiment score buckets on X-axis and number of files in each bucket on Y-axis

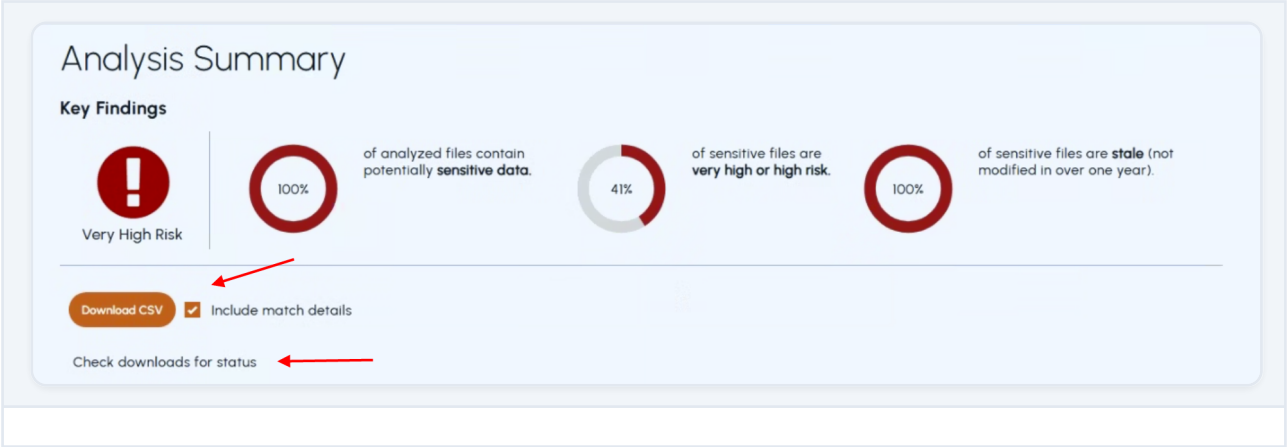
TAB	DETAILS FOR ANALYZED SENSITIVE FILES
Sensitive Files	List of sensitive files with their individual risk levels and risk scores
Duplicate Files	List of duplicate files with their individual risk levels and risk scores
Visualization	Word cloud of all entities, patterns, and policy conditions based on number of hits, along with People and locations.
Sentiment Score **	List of files with their sentiment scores
	** Sentiment Score tab is visible only if Analyze sentiment checkbox is selected for a particular analysis.

The following image shows the details displayed in the Overview tab of the Analyze page:

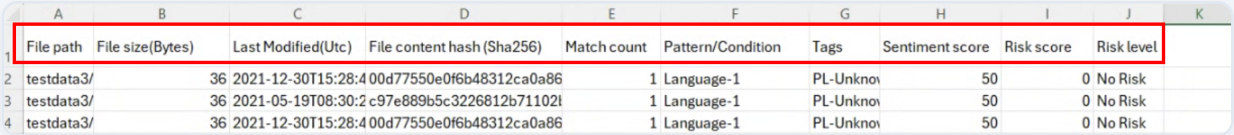
Following image shows the details displayed on the Summary page:



You can download the CSV without including match details. To include match details, check the box and click Download CSV.

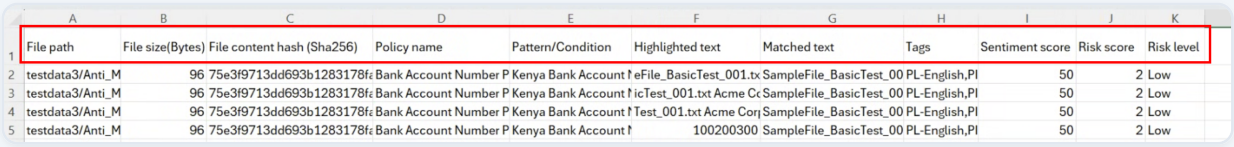


After downloading the CSV, check the download status from your browser and open the file. If you have not included match details, the CSV column headers will appear as shown in the following image:



File path	File size(Bytes)	Last Modified(Utc)	File content hash (Sha256)	Match count	Pattern/Condition	Tags	Sentiment score	Risk score	Risk level
testdata3/	36	2021-12-30T15:28:4	00d77550e0f6b48312ca0a86	1	Language-1	PL-Unknov	50	0	No Risk
testdata3/	36	2021-05-19T08:30:2	c97e889b5c3226812b71102l	1	Language-1	PL-Unknov	50	0	No Risk
testdata3/	36	2021-12-30T15:28:4	00d77550e0f6b48312ca0a86	1	Language-1	PL-Unknov	50	0	No Risk

If you have included match details, the CSV column headers will appear as shown in the following image:



File path	File size(Bytes)	File content hash (Sha256)	Policy name	Pattern/Condition	Highlighted text	Matched text	Tags	Sentiment score	Risk score	Risk level
testdata3/Anti_M	96	75e3f9713dd693b1283178fa	Bank Account Number P	Kenya Bank Account N	eFile_BasicTest_001.tx	SampleFile_BasicTest_00	PL-English,Pl	50	2	Low
testdata3/Anti_M	96	75e3f9713dd693b1283178fa	Bank Account Number P	Kenya Bank Account N	icTest_001.txt Acme Cc	SampleFile_BasicTest_00	PL-English,Pl	50	2	Low
testdata3/Anti_M	96	75e3f9713dd693b1283178fa	Bank Account Number P	Kenya Bank Account N	Test_001.txt Acme Cor	SampleFile_BasicTest_00	PL-English,Pl	50	2	Low
testdata3/Anti_M	96	75e3f9713dd693b1283178fa	Bank Account Number P	Kenya Bank Account N	100200300	SampleFile_BasicTest_00	PL-English,Pl	50	2	Low

The Summary page provides the analysis summary of potentially sensitive files. It also provides options to download reports for Compliance Analysis in CSV format.

Sentiment analysis support under Analyze page

You can perform sentiment analysis by using Analyze feature in UI. The Analyze sentiment checkbox appears on Analyze page when you drag and drop a folder to be analyzed. Analyze sentiment checkbox is selected by default if any of the selected policies have sentiment score related condition. If selected policies do not have sentiment condition, you can select or clear the Analyze sentiment checkbox.

Once the analysis of the selected content is complete, sentiment score is displayed on the Overview page at the bottom, in the form of a bar chart. Sentiment Score tab is visible only if Analyze sentiment checkbox is selected for a particular analysis. The Sentiment Score tab shows list of files and their sentiment scores. The sentiment analysis for all the files is also available through a downloadable CSV file.

You can capture the sentiment score against an entire test dataset to measure and understand what to expect from the sentiment analysis feature before using it in production.

The Sentiment Analysis feature can be enabled or disabled using YAML.

Risk level and risk score information on Analyze page

On Analyze page, the Overview tab provides donut chart that shows the number of sensitive files (any file which have got at least one policy hit) under different risk categories (no risk, low, medium, high, and very high). The Sensitive Files page shows the list of all sensitive files with corresponding risk level and risk score.

> “ ”

Note: Starting from release 4.1.0, the risk info (i.e. Risk Level and Risk Score) is available in the Analyze CSV report.

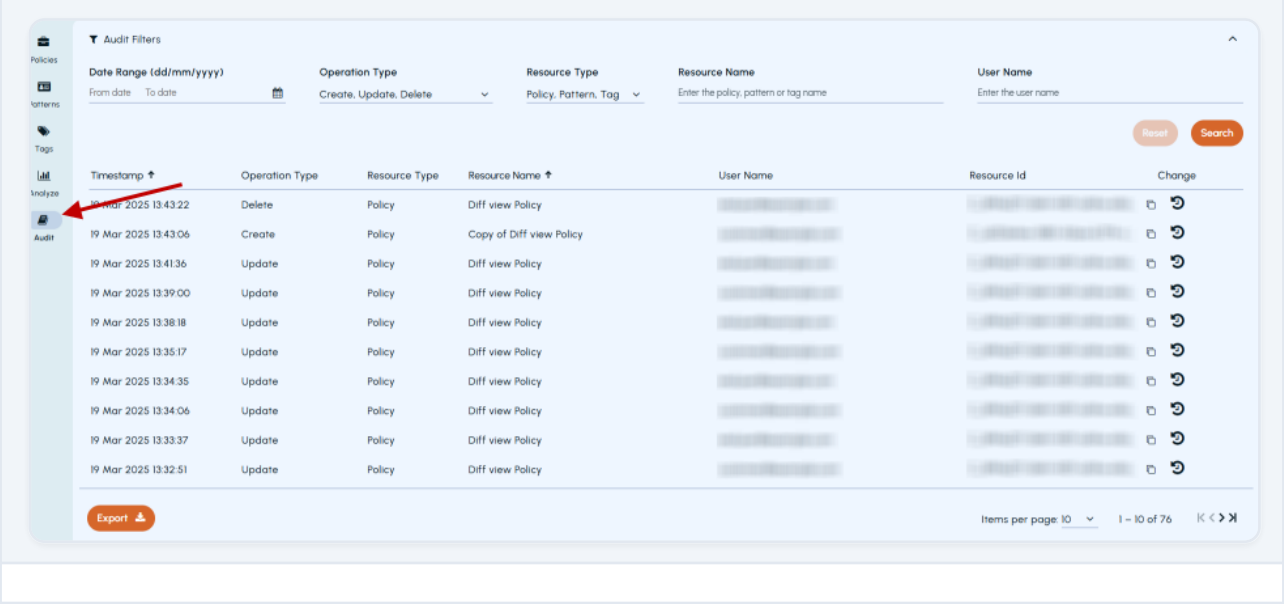
> “ ”

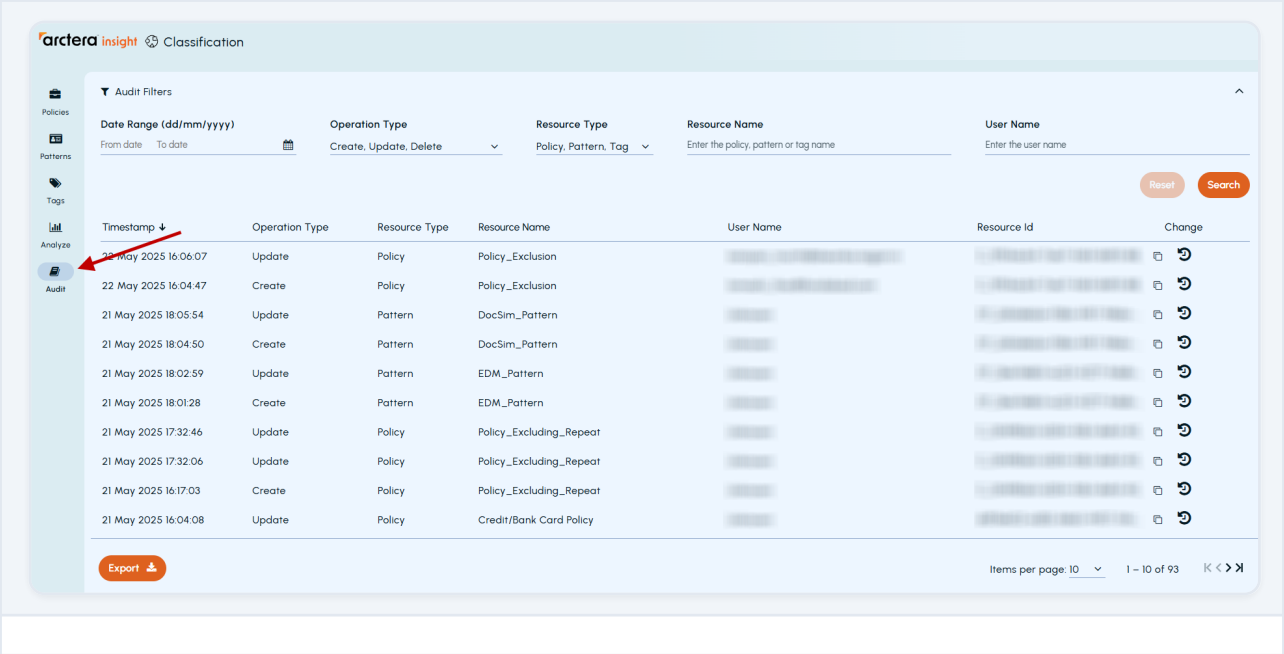
Audit

This section includes the following topics:

Overview

The Audit feature allows users to track changes made to policies. It provides a detailed history of modifications, including timestamp, user information, and a comparison of previous and updated values. This ensures transparency, accountability, and compliance.





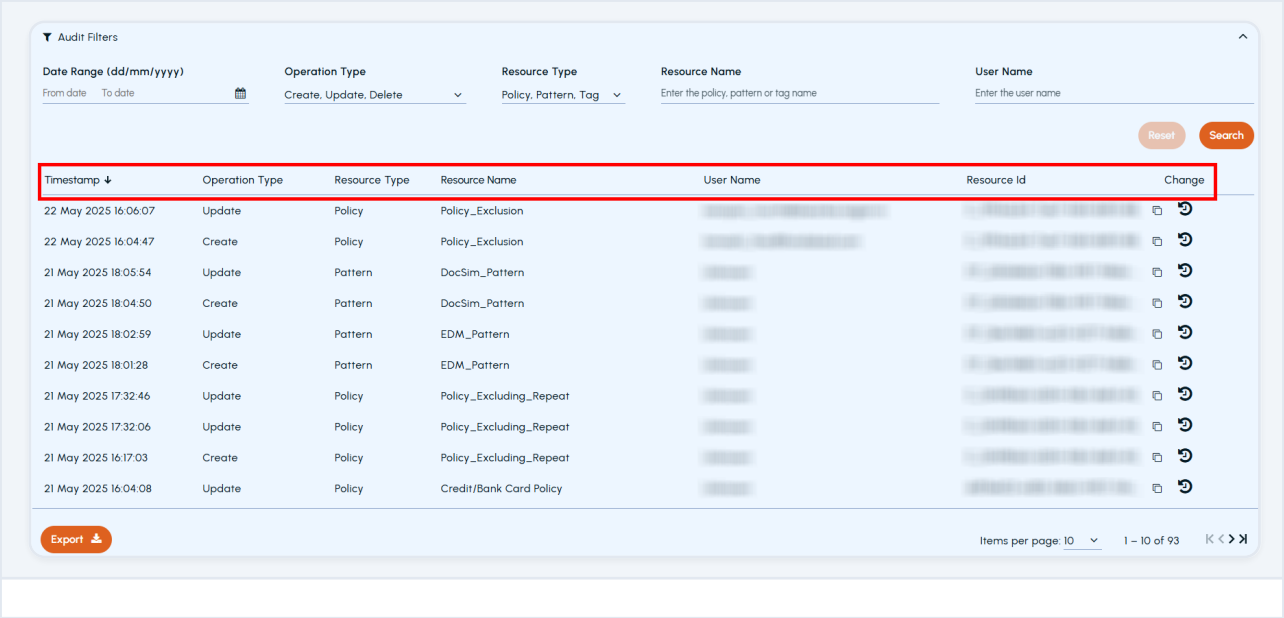
> “ ”

Note: Audit logs will only be captured starting from the date the Audit feature is introduced. Any changes made before this date will not be recorded in the audit history.As a result, older records and past modifications may not be available for review.

> “ ”

Understanding Audit Logs

Every action performed on a resource is recorded in the audit log, providing a detailed history of changes. Each entry in the audit log captures essential information to help users track modifications, identify who made the changes, and understand the impact.

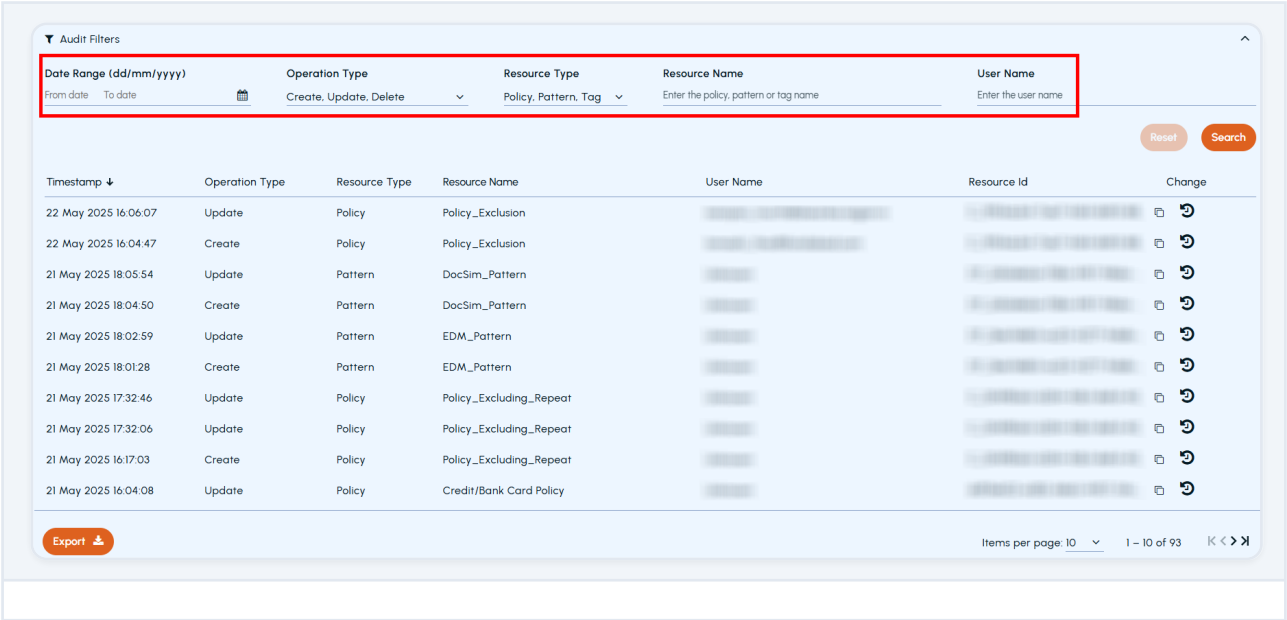


Each audit log entry includes the following:

- Timestamp \- The exact date and time when the change occurred.
- Operation Type \- Specifies whether the resource was Created, Updated, or Deleted.
- Resource Type \- Identifies the type of resource (Policy, Pattern, or Tag).
- Resource Name \- The name of the modified resource.
- User Name \- The user who performed the change.
- Resource ID \- A unique identifier for the resource.
- Change Details \- Highlights the differences between the previous and updated versions.

Filtering and Searching Audit Logs

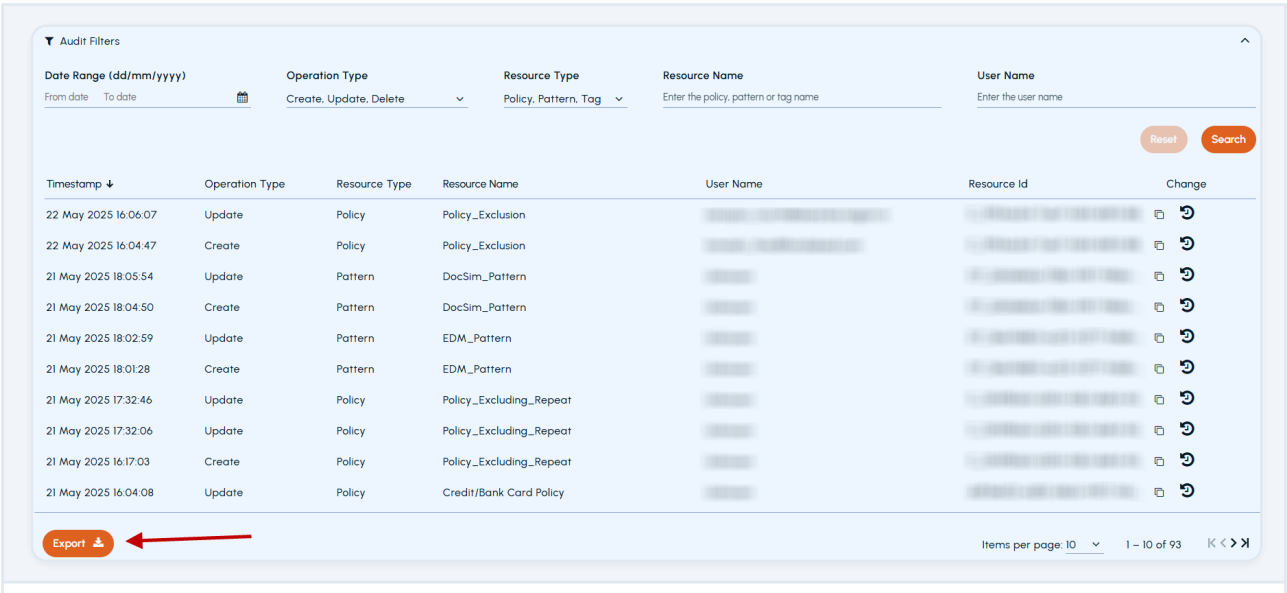
To quickly narrow down the audit logs result, you can use various filters and search options based on time, action type, resource details, or user activity.



- Date Range Filter \- View changes made within a specific time period.
- Operation Type Filter \- Filter logs based on change type (Created, Updated, Deleted).
- Resource Type Filter \- Narrow results by resource type (Policy, Pattern, Tag).
- Resource Name Filter \- Find changes related to a specific resource by name.
- User Name Search \- Locate modifications made by a particular user.

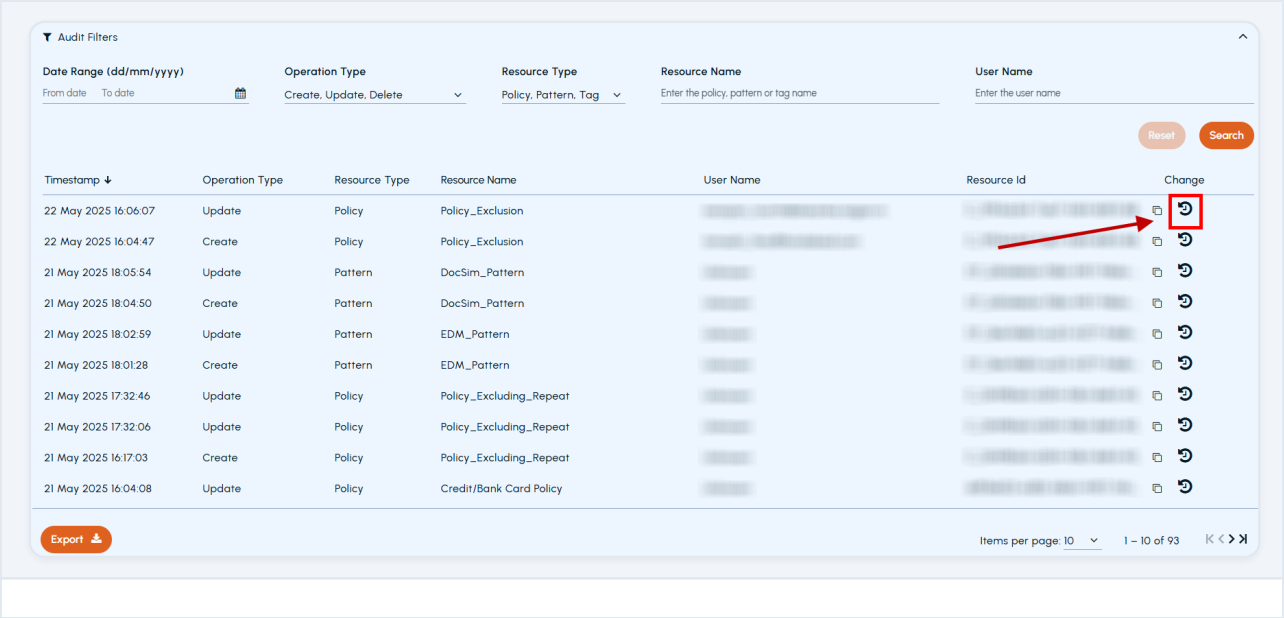
Exporting Audit Logs

You can download the searched audit logs in CSV format for offline analysis or compliance reporting by clicking Export and saving the file on your computer..



Using the Diff View

To use Diff View, select a specific audit entry and click the Load icon.



A side-by-side comparison window appears, displaying modified sections with color-coded highlights for easier analysis. The change highlights feature in the Diff View is currently available for comparisons. Each type of change is represented using a specific color, and you can refer to the legend located at the bottom right corner of the page for clarity. You can sort the records by Timestamp or Resource Type by clicking the corresponding column headers. The list updates automatically based on the selected sort option. Use the pagination controls at the bottom of the page to select the number of records per page. You can also navigate between pages using the arrow buttons next to the pagination controls.

Known limitations for Diff View

- The audit side-by-side Diff View may not render the correct changes when edit or reset operations are performed on a built-in policy.
- The highlighting is designed to maximize visibility of changes; however, it may not be 100% accurate in all cases.