LUX 2.6 Engine and Plugins Configuration Guide

1 Main Configuration Files

The configuration files for LUX Engine are found in `<engine_home>/EngineMain/data/conf`, and are a mix of properties files and xml files. Data files for plugins are found in `<engine_home>/EngineMain/data` and its subdirectories. Section 1 of this document provides an overview of the two main configuration files: `engine.properties` and `ae.xml`. These files load plugins, which may in turn load their own configuration files. Plugin configuration is explained in Section 2.

1.1 engine.properties

`engine.properties` controls many core Engine options, in addition to specifying which Ingest and Alerter plugins are loaded by the engine. Table 1 lists options that can be configured in `engine.properties`.

Property	Description	Data Type or Valid Values
engine.system.name	Name of the system for Admin Console	string
engine.realtime	Whether to run the engine against streaming (true) or historical (false) data	boolean
engine.system.support.address	If specified, rule disable emails will be copied to this email address	Any valid email address
matching.engine.threadcount	Number of engine threads. For values <=0 supported hardware threads count will be used.	0 - max system threads
matching.engine.accesscontrol.enabled	Whether the data to be processed has classification markings	boolean
matching.engine.classification.required	Optional, default true. If false, events without classifications specified will be treated as UNCLASSIFIED	boolean
accesscontrol.processor	Name of the Java class that should be used to process classification markings	`icg.engine.security.UnclassAccessControlProcessor`, `icg.engine.security.NullAccessControlProcess`

		`or`, `icg.engine.security.jblocks.JBlocksWrapper`
tc.max.alerts.per.rule	Maximum number of alerts a Time Correlated Rule will hold in memory	positive integer, default 1000
tc.max.alerts.per.bucket	Maximum number of alerts a Time Correlated Rule will hold in memory per bucket, for example per unique attribute	positive integer, default 50
tc.max.queue.size.per.worker	Number of events to buffer per processing thread	positive integer, default 100
engine.stats.update.frequency.ms	How frequently to update the core engine stats in the log and Admin Console, in milliseconds	positive long
alert.stats.update.frequency.ms	How frequently to update the alert stats in the log and Admin Console, in milliseconds	positive long
tc.stats.update.frequency.ms	How frequently to update the time correlated rule stats in the log and Admin Console, in milliseconds	positive long
metrics.update.frequency.ms	How often to write metrics to metrics.log, set to 0 to disable metrics logging	positive long
event.logger.enabled	The event logger will write events in their entirety to a log file	boolean
event.logger.streams	specify streams or leave commented out for all streams	Comma separated list of stream names
event.logger.xslt.file	Path to XSLT file used to transform events written by the event logger	Path to a valid XSLT file
circuitbreaker.global.maxaps	Global maximum alerts per second, per rule value. Will be used if maxaps is not set for an alerter or if the the alerts maxaps is greater than this value.	positive integer
circuitbreaker.checkfrequency	How often the circuit breaker checks to see if maximum alerts per second has been exceeded, in milliseconds	positive long
alerter.default.discard.policy	Determines how alerts exceeding the alerter queue size are handled (dropped entirely, pause pipeline and wait to catch up, cache to disk)	DROP, WAIT, CACHE
alerter.plugin.path	Path to alerter plugins. Relative to Engine working directory (EngineMain/data by default)	Valid file path, default plugins/alerter_plugins
matching.engine.max.rule.eval.duration.ms	How long a rule can take to process an event, in ms, before getting a strike	long

matching.engine.max.slow.rule.strikes.before.disable	How many strikes a rule can get before being disabled by the system for being too slow	int
matching.engine.streamanalyzer.enabled	Option to enable/disable the engine's StreamAnalyzer, which automatically finds attributes in events to populate the advanced rule forms in the UI. (default: true)	boolean
Alert plugin blocks (see 1.1.1)		
event.ingest.plugin.path	Path to event ingest plugins. Relative to Engine working directory (EngineMain/data by default)	Valid file path, default plugins/event_ingest_plugins
Event Ingest plugin blocks (1.1.2)		
rulesource.disabledrules.email.enabled	Whether to send emails when a rule becomes disabled	boolean
rulesource.rule.age.max.days	If enabled, engine will delete any rule whose lastModified date is older than this number of days	positive integer
rulesource.rule.age.warn.days	Engine will send a warning for any rule whose lastModified date is older than this number of days	positive integer <= rulesource.rule.age.max.days
rulesource.rule.age.enabled	Whether to monitor rule age for automatic warning and deletion	boolean
rulesource.plugin.path	Path to rule source plugins. Relative to Engine working directory (EngineMain/data by default)	Valid file path, default plugins/rule_source_plugins
Rule source blocks (see 1.1.3)		
naisource.plugin.path	Path to NAI source plugins. Relative to Engine working directory (EngineMain/data by default)	Valid file path, default plugins/nai_source_plugins
NAI source blocks (see 1.1.4)		
admin.console.web.enabled	turn the Admin Console webpage updates on or off	boolean
admin.console.update.frequency.ms	Update frequency for REST calls, in milliseconds	positive long
admin.console.rest.url	Base URL for admin console REST services	URL pointing to a running Admin Console
admin.console.rest.session.id	Session ID for REST communication with the Admin Console	default PleaseChangeMe

Table 1 - engine.properties options

1.1.1 Alerter Plugin Blocks

The format is as follows (# starts at 1 and increments from there)
alerter.#.classpath - java.classpath.to.alerter.Class
alerter.#.name - Name of the alerter. Will be used in the rules "alerters" list.
alerter.#.maxaps - Optional property to set the maximum alerts per second for a specific alerter.
alerter.#.maxalerts - Optional property to set the maximum number of alerts from a rule for a specific alerter.

Example:

```
alerter.1.classpath=icg.engine.alerter.jms.lux.LUXJsonConsoleAlerter
alerter.1.name=Console
alerter.1.maxaps=10
alerter.1.alert.format=XML
```

1.1.2 Event Ingest Plugin Blocks

The format is as follows (# starts and 1 and increments from there)
event.ingest.#.classpath - java.classpath.to.event.ingest.Class
event.ingest.#.name - Name of the ingest plugin.
event.ingest.#.confpath - Path to configuration file for plugin, will be passed in constructor
event.ingest.#.stream.name - Stream name for events that come from this plugin (Use a comma or semicolon separated list to duplicate to multiple streams)

Example:

```
event.ingest.1.stream.name=Netflow
event.ingest.1.classpath=icg.engine.event.generator.NetflowEventGenerator
event.ingest.1.name=NetflowEventGenerator
event.ingest.1.confpath=NetflowEventGenerator.properties
```

1.1.3 Rule Source Block

The format is as follows (# starts and 1 and increments from there)
#rulesource.#.classpath - java.classpath.to.rulesource.Class
#rulesource.#.name - Name of the rule source.

Example:

```
rulesource.1.classpath=icg.engine.rulesource.luxfile.LUXFileRuleSource
rulesource.1.name=LUXFile
```

1.1.4 NAI Source Block

The format is as follows (# starts and 1 and increments from there)

#naisource.#.classpath - java.classpath.to.naisource.Class
#naisource.#.name - Name of the alert source.

Example:
```
naisource.1.classpath=icg.engine.naisource.luxfile.LUXFileNAISource
naisource.1.name=LUXFileNAISource
```

1.2 ae.xml

`ae.xml` controls which Analytic, Enrichment, and Event Output plugins are loaded by the engine. Table 2 lists options that can be configured in `ae.xml`.

Property	Description	Data Type or Valid Values
java_analytics_path	Disk location of Analytic plugins	Valid folder path
java_processing_thread_count	Number of threads to assign to this group of plugins	positive integer
stat_reporting_interval	Frequency of stats reporting for this group of plugins, in milliseconds	positive long
java_enrichments_path	Disk location of Enrichment plugins	Valid folder path
plugin_path	Disk location of Event Output plugins	Valid folder path
stream_in	Specifies a data stream to send to a particular Analytic, Enrichment, or Event Output plugin	A stream name specified in an Ingest plugin's configuration, or a `stream_out` property of an analytic
stream_out	The name of the stream that an analytic will send its results on	string
name	The Java class name containing the plugin code	A class name on the runtime classpath
classification	Classification of an plugin's name and configuration	A valid classification string
description	Plugin description sent to User Interface	string
display_name	Plugin name for User Interface and Admin Console	string
overlay	Whether an Analytic plugin outputs a map layer for the User Interface	boolean

Table 2 - ae.xml options

2 Plugin Configuration

This section describes some the out-of-the-box Engine plugins, and how to configure them.

2.1 Analytic Plugins

Most Analytic plugins support either an EventFilter or a GeoAnalyticFilter. New Analytics should use GeoAnalyticFilter, which wraps EventFilter (and GeoFilter) and can perform a superset of EventFilter's capabilities. EventFilter and GeoAnalyticFilter are configured by the properties in tables 3 and 4, respectively. These parameters are optional.

Property	Description	Data Type or Valid Values
filter_xpath.x	XPath to an event attribute to filter on	XPath string
filter_regex.x	Regex with which to filter the corresponding event attribute. Attribute must match specified regex. Specify either a regex **OR** list for each xpath.	Regex string
filter_list.x	List file with which to filters the corresponding event attribute. Attribute must equal at least one value from file. List files are expected to contain one valid value per line. Specify either a regex **OR** list for each xpath.	File path
list.poll.interval.ms	The frequency with which to poll filter files, in milliseconds	long

Table 3 - EventFilter options

Property	Description	Data Type or Valid Values
<EventFilter properties>	<GeoAnalyticFilter contains an EventFilter>	
min_lat	Events with less than min_lat latitude will not be processed	double, -90 to 90
min_lon	Events wil less than min_lon longitude will not be processed	double, -180 to 180
max_lat	Events with more than max_lat latitude will not be processed	double, -90 to 90
max_lon	Events with more than max_lon longitude will not be processed	double, -180 to 180
geo_filter_file	Path to KML/KMZ file to use as a geospatial filter	Path to KML/KMZ file
geo_filter_poll_interval_s	How frequently to poll geo_filter_file for changes, in seconds. -1 = don't poll	integer
geo_filter_is_inside	Whether to process events that are inside the KML/KMZ shapes (true) or outside (false)	boolean
geo_required	Default true. If false, only use EventFilter and ignore geospatial filters.	boolean
max_distance_to_region_m	Max distance an event can be from	long

	a region, in meters. 0 (default) means the event geo overlaps or is contained by the region geo.	
apply_max_distance_to_points_only	If true, max_distance_to_region_m only applies to areas that are points. Areas that are not points will effectively have a 0 value for max_distance_to_region_m	boolean

Table 4 - GeoAnalyticFilter options

2.1.1 Abnormal Track Analytic

`icg.engine.analytic.track.analysis.AbnormalTrackAnalytic`

Detects when a unique ID has a track greater than a specified length, but the distance delta between the start and end points are less than a certain amount. Abnormal tracks are then sent through a classifiers to attempt to identify the activity represented by the track. Classifiers include PMML models, machine learning models, and computer vision models.

Property	Description	Data Type or Valid Values
aggregate_classifications	Whether to combine the classifications of all events in all paths for the output classification	boolean
clear_alerted_data	Whether to clear data from the window once it has been represented in output	boolean
event.prop.x.display.name	Display name of a property to be included in output	string
event.prop.x.key	Property key of a property to be included in output	string
event.prop.x.xpath	XPath to a property to be included in output	XPath
id_attribute_xpath	XPath to the ID attribute	XPath
id_display_name	Display name to use for the ID attribute	string
location_xpath	XPath to the event location	XPath
min_events	Minimum events needed to form a path	positive integer
min_travel_distance_meters	Minimum length of a path in meters	positive long
max_percent_distance_travelled	Maximum percentage of a path's actual length represented by the difference between the start and end points for which a path will be considered "abnormal"	1-100
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long

track.classifier.x	Optional, up to 3 names of TrackClassifiers to load.	“shape”, “icg”, or “kmeans”
<filter>	GeoAnalyticFilter	

Table 5 - Abnormal Track Analytic options

2.1.2 Association Analytic

`icg.engine.analytic.association.AssociationAnalytic`

Creates a directed multigraph with configurable node and edge definitions. Alerts when any 2 nodes have n or more edges between them in a configurable time window.

Property	Description	Data Type or Valid Values
description_format_string	Format string for output description with %s for source ID followed by %s for destination ID	String.format() format string
display_attribute_x_display_name	Display_attributes will be stored and output with results. Display name for attribute	string
display_attribute_x_name	XML/JSON name for attribute	string
display_attribute_x_xpath	XPath to attribute	XPath string
dst_id_xpath	XPath in the event where the destination ID is found	XPath string
graph_store_relationship_desc	Relationship description for the Graph Store, if write_to_graph_store is true	string
graph_store_relationship_desc_attribute	Relationship description attribute in the Graph Store, if write_to_graph_store is true	string
graph_store_relationship_output_property	Output property for Graph Store Relationship ID, if write_to_graph_store is true	string
graph_store_relationship_output_property_display_name	Display name for Relationship ID output property, if write_to_graph_store is true	string
id_filter_file	Optional file that contains IDs to process	Path to IDs file
num_associations	The number of times two IDs need to be associated before an alert	positive integer
src_id_xpath	XPath in the event where the source ID is found	XPath string
timeout_duration_seconds	Timeout before the same source/destination ID pair can be alerted on again	positive integer
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long

write_to_graph_store	Whether or not to write entities and relationships to the Entity Manager Graph Store	boolean
<filter>	EventFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
<RelationshipConfig properties>		

Table 6 - Association Analytic options

2.1.3 Co-Travel Analytic

`icg.engine.analytic.cotravel.CoTravelAnalytic`

When two unique attributes are within a specified distance from each other within a specified amount of time, an association is created. Alerts are generated when a minimum number of associations occur within a specified time window, and cover at least a minimum distance.

Property	Description	Data Type or Valid Values
aggregate_classifications	Whether to combine the classifications of all events in all paths for the output classification	boolean
description_format_string	Format string for output description with %s for first ID followed by %s for second ID	String.format() format string
dual_alerts	Whether or not to produce two alerts for each co-travelling pair (x->y and y->x)	boolean
location_xpath	XPath to the event location	XPath string
max_association_distance_m	Max distance between IDs to be associated, in meters	positive long
max_association_time_s	Max temporal delta between events for IDs to be associated, in seconds	positive integer
max_associations	Max associations to store for a pair of IDs	positive integer
min_associations	Minimum number of associations to declare two IDs are co-travelling	positive integer
min_travel_distance_m	Minimum distance a pair must travel to be co-travelling	positive long
required_attribute_display_name	Optional parameter to enforce the presence of an attribute, name for it	string
required_attribute_xpath	XPath to required attribute.	XPath string
attribute_display_name	Display name for required attribute	string
unique_attribute_xpath	XPath to ID attribute	XPath string
window_size_s	Amount of data to hold in memory, measured as seconds since current	positive long

	time	
<filter>	GeoFilter	

Table 7 - CoTravel Analytic options

2.1.4 Country Heatmap Visualization

`icg.engine.analytic.heatmaps.CountryHeatmapVisualization`

Visualization Analytic that heatmaps events by areas loaded from a CSV file.

Property	Description	Data Type or Valid Values
geometry_column	CSV column containing geometry	
include_descriptions	Whether to include description blocks in KML output	boolean
kml_output_file	Optional file to output KML into	Path to file
region_id_column	CSV column that contains the region ID	integer
region_xpath	XPath to region ID in event	XPath string
regions_csv_file	CSV file containing, at a minumum, a region ID column and a KML geometry column (which should be quoted)	Path to CSV file
window_size_s	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	EventFilter	

Table 8 - Country Heatmap Visualization options

2.1.5 Dead Reckoning Analytic

`icg.engine.analytic.deadreckoning.DeadReckoningAnalytic`

Provides a future predicted location of an attribute based on the last two positions observed.

prediction_size_s	How many seconds into the future to predict	positive integer

Property	Description	Data Type or Valid Values
date_format_string	Format string for event date, see SimpleDateFormat javadoc	Date format string
date_xpath	XPath to event date	XPath string
event.prop.x.key	Event property to store with output events	string
event.prop.x.display.name	Display name for event property	string
event.prop.x.xpath	XPath to event property	XPath string
id_attribute_name	Property name for ID attribute	string

id_attribute_xpath	XPath to ID attribute	XPath string
location_xpath	XPath to event location	XPath string
prediction_geo_filter_file	Optional KML/KMZ file containing geospatial regions to exclude from predictions	Path to KML/KMZ file
<filter>	GeoAnalyticFilter	

Table 9 - Dead Reckoning Analytic options

2.1.6 Dupe ID Analytic

`icg.engine.analytic.analytic.dupeid.DupeIDAnalytic`

Detects unique attributes that appear at two locations that could not be reached by traveling at a configurable maximum speed in a specified time delta between events.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location attribute	XPath string
id_attribute_name	Output display name for ID attribute	string
id_attribute_xpath	XPath to ID attribute	XPath string
max_speed_meters_per_second	The maximum speed entities in the data stream could move, measured in meters per second.	Positive long
min_distance_delta_m	Min distance in meters that the events can be apart to qualify for an analytic event	Positive long
<filter>	GeoAnalyticFilter	

Table 10 - Dupe ID Analytic options

2.1.7 Entity Geospatial Normalcy Analytic

`icg.engine.analytic.normalcy.entitygeo.EntityGeospatialNormalcyAnalytic`

Determines the geospatial normalcy for an event, based on the history of the entity identified by the ID_ATTRIBUTE. Optionally determines normalcy based on time of day/week/year.

Property	Description	Data Type or Valid Values
geom_query	XPath to event location. "geom_query" "lat_query" "lon_query" geom.name or alternately lat.name,lon.name are used to define a geo-location for the event.	XPath string
grid_size_m	Approximate width of single cell in geospatial grid	Postive long
id_attribute_regex	Regex ID attribute must pass to be processed	Regex string
id_attribute_xpath	XPath to ID attribute	XPath string

lat_query	XPath to latitude attribute	XPath string
lon_query	XPath to longitude attribute	XPath string
normalcy_threshold	Maximum normalcy value to alert on, measured 0 (abnormal) to 1 (normal)	Double 0 to 1
training_file	Optional CSV training file	File path
training_period_s	Training time in seconds, no alerts during this time	Positive integer
use_day_of_week	Use the day of the week in the time bucket (Is this normal for Monday?)	boolean
use_hour	Use hour in time bucket (Is this normal for 1am-2am?)	boolean
use_minute	Use minute in time bucket	boolean
use_month	Use month in time bucket	boolean
use_second	Use second in time bucket	boolean
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 11 - Entity Geospatial Normalcy Analytic options

2.1.8 Geo New Value Analytic

`icg.engine.analytic.newvalue.GeoNewValueAnalytic`

NewValueAnalytic (see below) with added geospatial filter.

Property	Description	Data Type or Valid Values
history_file_name	Optional file to save values in between system restarts	File path
history_file_save_interval	How often to save to history file, if enabled	positive long
location_xpath	XPath to event location	XPath string
identity_output_display_name	Display name for ID attribute	string
identity_output_property	Property name for ID attribute	string
identity_path	XPath to ID attribute	XPath string
training_time_s	(Optional) Amount of time to build history before alerting begins, in seconds	positive long
value_path	XPath to value attribute	XPath string
value_output_property	(Optional) Prop name for new value	String
value_output_display_name	(Optional) Display name for new value	String
last_value_output_property	(Optional) Prop name for last value	String
last_value_output_display_name	(Optional) Display name for last value	String
description_output_property	(Optional) Prop name for description	String

description_output_display_name	(Optional) Display name for description	String
description_format_string	(Optional) Format string for description, takes 2 strings: the new value and the last value	Format string
change_only	Flag that will make this analytic only alert if there was a previous value for an attribute (can't be null). Default: false	boolean
<filter>	GeoAnalyticFilter	

Table 12 - Geo New Value Analytic options

2.1.9 Geohash Clustering Analytic

`icg.engine.analytic.geohashclustering.GeoHashClusteringAnalytic`

Alerts when it detects a geo-spatial clustering of data. The data is grouped by whatever element is specified in configuration. Clusters must be of a minimum configurable size, must occur within a specified time window, and must be within a cluster bounded by a circle of no more than a specified size.

Property	Description	Data Type or Valid Values
aggregate_classifications	Whether to combine the classifications of all events in all paths for the output classification	boolean
clear_alerted_data	Whether to clear data from the window once it has been represented in output	boolean
area_diameter_meters	Maximum size of a cluster, in meters	positive long
common_attribute_display_name	Optional attribute all events in a cluster must share, display name	string
common_attribute_xpath	XPath to common attribute	XPath string
description_format_string	Format string for output description with %s for what is being clustered	String.format() format string
diameter_display_name	Display name for diameter of cluster	string
location_xpath	XPath to event location	XPath string
max_events	Max number of events in a cluster	positive integer
min_cluster_diameter_meters	Minimum size of a cluster, in meters	positive integer
min_cluster_time_seconds	Optional parameter to specify a minimum time window in seconds for all cluster events to occur	positive integer
min_events	Minimum events in a cluster	positive integer
unique_attribute_xpath	Optional attribute that must be unique among all events in a cluster, XPath	XPath string
window_size_seconds	Amount of data to hold in memory, measured as seconds since current	positive long

	time	
<filter>	GeoAnalyticFilter	

Table 13 - Geohash Clustering Analytic options

2.1.10 Geohash Proximity Analytic

`icg.engine.analytic.geohashproximity.GeoHashProximityAnalytic`

Alerts when two unique ID attributes are in events that occur within a specified proximity, within a configured amount of time. The analytic can optionally be configured with event filters, which will partition the data into two groups. If used, analytic will alert if a new event from one group is in proximity to a stored event in other group.

Property	Description	Data Type or Valid Values
description_format_string	Format string for output description with two %s for the names of the two things in proximity	String.format() format string
distance_display_name	Display name for the distance between the two things in proximity	string
distance_threshold_meters	Max distance between things to be in proximity	positive long
filter_attribute_regex_a	Optional way to break events into groups “A” and “B” based on an attribute, so proximity is only between items in different groups. Attribute regex for “A” group	Regex string
filter_attribute_xpath_a	Attribute XPath for “A” group	XPath string
filter_attribute_regex_b	Attribute regex for “B” group	Regex string
filter_attribute_xpath_b	Attribute XPath for “B” group	XPath string
id_display_name	Display name for ID attribute	string
id_xpath	XPath to ID attribute	XPath string
location_xpath	XPath to event location	XPath string
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	GeoFilter	

Table 14 - Geohash Proximity Analytic options

2.1.11 Graph Clusters Analytic

`icg.engine.analytic.graphclusters.GraphClustersAnalytic`

GraphClusters builds a network graph from a stream of data, where nodes in the graph are created through identified "source" and "destination" fields within the same event. The data stream feeding GraphClusters can be filtered by performing regex queries on any fields within

the event, such that only events which pass these filters are added to the graph. GraphClusters look for bi-directional "association" between nodes in the graph, where association is defined as a minimum number of edges existing between the nodes in a configured time window. Once associations are identified, they are added to a second graph, called the Association Graph. When a node is added to the AssociationGraph, a check is performed to see how many nodes in that graph can be reached by a certain degree of edge traversal, specified via configuration. If the number of nodes reachable by traversal is greater than or equal to a configured threshold, the analytic produces an event. The analytic event conveys to the user that the triggering node is a new member of a GraphCluster, as well as all constituent nodes of the GraphCluster. Once a GraphCluster has been identified, subsequent alerts on the same cluster can optionally be ignored for a configurable time window.

Property	Description	Data Type or Valid Values
attach_images	Whether to add a jpg of the graph cluster to the output (in ascii binary)	boolean
bidirectional_required	Whether bidirectional connection between nodes is required to form an association	boolean
cluster_degree	The number of “hops” in the graph to traverse to meet min_nodes_for_cluster nodes	positive integer
dst_id_xpath	XPath to the destination node ID	XPath string
edge_attribute_xpath	XPath to attribute to store with edge	XPath string
id_filter_file	Optional file to filter graph nodes	File path string
min_edges_for_association	Minimum number of edges between nodes before they are considered associated	positive integer
min_nodes_for_cluster	Minimum number of nodes	positive integer
src_id_xpath	XPath to ID of source node	XPath string
timeout_duration_seconds	Timeout before the same cluster can be alerted on again	positive integer
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	EventFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 15 - Graph Clusters Analytic options

2.1.12 Group DAOI Analytic

`icg.engine.analytic.groupdaoi.GroupDAOIAnalytic`

Correlates events between two streams, A and B. Creates dynamic areas of interest around unique attributes in stream A. When an event from stream B is received, creates an alert for

each attribute from stream A within a configurable distance threshold. Optionally, instead alerts if there are no attributes from stream A within the distance threshold.

Property	Description	Data Type or Valid Values
alert_on_no_hits	If true, alert when there are no events from stream B near an event from stream A within the specified timeframe	boolean
description	Description string for alert	string
distance_display_name	Display name for distance between events	string
distance_threshold_meters	Maximum distance between events	positive long
id_display_name	Display name for ID attribute	string
id_xpath_a	XPath to ID attribute	XPath string
location_xpath_a	XPath to event location in A stream	XPath string
location_xpath_b	XPath to event location in B stream	XPath string
stream_name_a	Stream name of A stream	string
stream_name_b	Stream name of B stream	string
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	GeoAnalyticFilter postfix “A” GeoAnalyticFilter postfix “B”	One for each stream

Table 16 - Group DAOI Analytic options

2.1.13 Heatmaps Visualization

`icg.engine.analytic.heatmaps.HeatmapVisualization`

Sends KML periodically which shows the density of attributes, within a configured time window, by coloring grid cells on the map. Cell colors are based on the normal distribution of attributes, compared to the number in the given cell. Can limit counting to unique occurrences of attribute values, or use aggregate occurrences.

Property	Description	Data Type or Valid Values
grid_size_m	Approximate width of a single cell of the geospatial grid for the heatmap	positive long
id_attribute_name	Output name for ID attribute	string
id_attribute_regex	Regex ID attribute must satisfy	Regex string
id_attribute_xpath	XPath to ID attribute	XPath string
kml_3d_enabled	Whether output KML should be in 3D	boolean
kml_show_descriptions	Whether output KML should include description blocks per cell, will increase size	boolean

kml_classification	Classification for KML output	Classification string
kml_data_name	Name of things being heatmapped, for KML description blocks	string
kml_include_lookat	Whether to include KML lookat, which positions the camera to look at the data	boolean
kml_output_file	Optional file to output KML to	File path string
kml_output_interval_s	How often to output KML, in seconds	positive long
kml_output_max_length	Maximum character length of KML output	positive long
location_xpath	XPath to event location	XPath string
only_count_uniques	If true, only unique IDs will be counted in heatmap, rather than all events	boolean
window_size_s	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 17 - Heatmaps Visualization options

2.1.14 ID Inactivity Analytic

`icg.engine.analytic.inactivity.IdInactivityAnalytic`

Monitors a stream tracking occurrences of a unique ID. After the first occurrence of a value, if that value is not seen again within **`window_size_s`** seconds, an analytic event is sent. If **`continuous_alerting`** is set, analytic events will be sent every **`check_interval_s`** seconds until the ID appears, or **`max_inactive_time_s`** seconds have passed since it was last seen.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location attribute	XPath string
value_xpath	XPath to ID attribute	XPath string
window_size_s	Min time before an ID is considered inactive, in seconds	positive long
<filter>	GeoAnalyticFilter	
field.x.name	Display name for attribute to save from event	String
field.x.path	XPath for attribute to save from event	XPath string
continuous_alerting	Whether to alert multiple times (at each interval) for the same ID, or to clear after the first, default false	Boolean
max_inactive_time_s	(Optional) Time before ID stops alerting and is reset	Positive long

check_interval_s	How often to check timeout window, in seconds	Positive long
missing_value_display_name	Display name for missing IDs	String

Table 18 - ID Inactivity Analytic options

2.1.15 Moving Analytic

`icg.engine.analytic.moving.MovingAnalytic`

Detects when something has moved x meters in y seconds.

Property	Description	Data Type or Valid Values
clear_alerted_data	Whether to clear data from window once it has been alerted on	boolean
location_xpath	XPath to event location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
id_display_name	Display name for ID attribute	string
min_events	Minimum events needed before distance is measured	positive integer
min_travel_distance_meters	Minimum distance needed to be considered moving	positive long
window_size_seconds	Amount of data to hold in memory, measured as seconds since current time	positive long
<filter>	GeoAnalyticFilter	

Table 19 - Moving Analytic options

2.1.16 Multi-Attribute Normalcy Analytic

`icg.engine.analytic.normalcy.multiattribute.MultiAttributeNormalcyAnalytic`

Computes normalcy scores based on event counts for anomaly detection. Events are grouped by a configurable set of attributes (including optionally geospatial gridding) and counted over a configurable time window. Time windows themselves can be optionally grouped (e.g. by hour of day, day of week, etc.), and the binned counts are passed to one of several algorithms for assigning ‘normalcy’ values based on the history of counts in that bin.

Property	Description	Data Type or Valid Values
attribute.x.filter.include_regex	(optional) Regex inclusion filter for values to accept; values not matching will be ignored.	Regex string
attribute.x.filter.exclude_regex	(optional) Regex exclusion filter for values to accept; values matching this regex will be ignored	Regex string
attribute.x.filter.include_list_from_file	(optional) Path of a csv file with a list of values to include. Values not in this	File path

	list will be ignored.	
attribute.x.filter.include_list_from_file..header_rows	(optional) Number of header rows to skip when reading the include list csv file. Default is 0.	Integer
attribute.x.filter.include_list_from_file.column	(optional) Index of column in the include list csv file to take as the list of values to include. Default is 0.	Integer
attribute.x.filter.exclude_list_from_file	(optional) Path of a csv file with a list of values to exclude. Values in this list will be ignored.	File path
attribute.x.filter.exclude_list_from_file..header_rows	(optional) Number of header rows to skip when reading the exclude list csv file. Default is 0.	Integer
attribute.x.filter.exclude_list_from_file.column	(optional) Index of column in the exclude list csv file to take as the list of values to exclude. Default is 0.	Integer
attribute.x.output_display_name	Display name for attribute	string
attribute.x.output_property	Output property for attribute	string
attribute.x.query	XPath to attribute	XPath string, or one of PROPERTY_NAME, PROPERTY_VALUE
attribute.x.type	Type of attribute. Default is TEXT.	TEXT, GEO
attribute.x.grid_size_m	Approximate width of single cell in geospatial grid. Required if attribute type is GEO, ignored otherwise	double
classification	Classification of output	Classification string
drop_rare_items_threshold	(optional) If specified, the analytic will forget it's ever seen a value in a time bucket if the average count for that time bucket drops lower than this threshold	double
accumulate_history_for_unseen_items	If false, the analytic won't begin 'training' on a particular value ina time bucket until it's seen it at least once. If true, the first time a value is encountered its history will be back-filled with counts of 0 since the analytic began. Default false.	boolean
normalcy_method	Which method to use to track and determine normalcy values. Default is SIMPLE_NORMAL_STAT.	SIMPLE_NORMAL_STAT, SLIDING_WINDOW_NORMAL_STAT, ROLLING_NORMAL_STAT, ROLLING_NORMAL_STAT_WITH_HISTORY, BINARY_MODE
window_size	Number of historical values to track for each time bucket. Required with SLIDING_WINDOW_NORMAL_STAT,	Integer >= 0

	ROLLING_NORMAL_STAT_WITH_HISTORY. Optional with BINARY_MODE. Ignored otherwise	
decay	Decay factor. Required with ROLLING_NORMAL_STAT, ROLLING_NORMAL_STAT_WITH_HISTORY. Higher values cause expectations to shift more slowly in response to new data.	Double from 0.0 to 1.0
generate_normalcy_graph	(optional) If true, the analytic will include a visual graph of the history of a value when it outputs an event. Only works with some normalcy_method types. Default false.	
history_file_name	Optional file to store history	File path
history_file_usage	How to use history file, valid values are READ_ONLY, WRITE_ONLY, READ_WRITE	Enum
max_eventlist_output	Max events in event output	Positive integer
normalcy_threshold	Maximum normalcy value to alert on, measured 0 (abnormal) to 1 (normal)	Double 0 to 1
num_training_values	Number of time periods to train. Warning: if your time bucket uses month and this value is 1 or more, you won't get alerts for years.	Integer >= 0
For preload_default_value	Optional, value to preload into time buckets	double
use_day_of_week	Include the day of the week in time bucket IDs	boolean
use_hour	Include the hour of day in time bucket IDs	boolean
use_minute	Include the minute of the hour in time bucket IDs	boolean
use_month	Include the month of the year in time bucket IDs	boolean
use_second	Include the second of the minute in time bucket IDs	boolean
window_type	(optional) The size of a time bucket. If unspecified, the smallest size consistent with the specified use_x parameters will be used.	SECOND, MINUTE, HOUR, DAY, WEEK, MONTH, YEAR
<filter>	EventFilter and GeoAnalyticFilter, depending on if there are any GEO properties specified. Needs to change.	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 20 - Multi-Attribute Normalcy Analytic options

2.1.17 New Value Analytic

`icg.engine.analytic.newvalue.NewValueAnalytic`

NewValueAnalytic alerts whenever a value for a specified attribute is seen that hasn't been seen before. Can also find when a unique value of an attribute is found with multiple values for a different attribute.

Property	Description	Data Type or Valid Values
history_file_name	Optional file to save values in between system restarts	File path
history_file_save_interval_s	How often to save to history file in seconds, if enabled	positive long
identity_output_display_name	Display name for ID attribute	string
identity_output_property	Property name for ID attribute	string
identity_xpath	XPath to ID attribute	XPath string
training_time_s	Optional, amount of time to build history before alerting begins, in seconds	positive long
value_xpath	XPath to value attribute	XPath string
value_output_property	(Optional) Prop name for new value	String
value_output_display_name	(Optional) Display name for new value	String
change_only	Flag that will make this analytic only alert if there was a previous value for an attribute (can't be null). Default: false	boolean
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 21 - New Value Analytic options

2.1.19 Path Projection Analytic

`icg.engine.analytic.path.projection.PathProjectionAnalytic`

Dead reckons a position along a path a configurable number of points into the future at a configurable time interval. Optionally, the path can include stop boxes. If the projected path hits a stop box, it will end at the center of the box.

Property	Description	Data Type or Valid Values
date_format_string	SimpleDateFormat format string to parse event date	Date format string
date_xpath	Path to date in event	XPath string

id_attribute_regex	Regex for ID attribute to pass	Regex string
id_attribute_xpath	XPath to ID attribute	XPath string
location_xpath	XPath to event location	XPath string
path_lat_lon	Series of lat/lon coordinate pairs separated by spaces	Coordinate string
prediction_count	How many points into the future to project	Positive integer
prediction_size_s	How many seconds apart each prediction should be	Positive long
stop_box_lat_lon	Series of lat/lon coordinate pairs, which represent the corners of stop boxes.	Coordinate string
<filter>	GeoAnalyticFilter	

Table 23 - Path Projection Analytic options

2.1.20 Pattern of Life Normalcy Analytic

`icg.engine.analytic.normalcy.pol.POLNormalcyAnalytic`

Determines normal levels of number of occurrences of an event over time, and detects deviations from those levels. Normal operation will build Normalcy curves to determine normalcy, reporting area under the curve represented by the current value. Binary operation mode will detect when a value is the first zero or non-zero occurrence in a time bucket. In binary mode normalcy values will be reported as 0 or 1, if the value is the first zero/non-zero or not, respectively.

Property	Description	Data Type or Valid Values
binary_mode	Binary mode will alert for the first zero or non-zero number of occurrences in a time bucket	boolean
classification	Classification of output	Classification string
count_unique	If true, will only count unique IDs toward normalcy	boolean
history_file_name	Optional file to save history	File path
id_attribute_regex	Regex ID attribute must pass to be processed	Regex string
id_attribute_xpath	XPath to ID attribute	XPath string
ids_file	Known IDs, counts will start for these IDs when the analytic starts, even if no events with them have been seen	File path
ids_filter_file	List of IDs to filter data on	File path
normalcy_threshold	Maximum normalcy value to alert on, measured 0 (abnormal) to 1 (normal)	Double 0 to 1
num_training_values	Number of time periods to train. Warning: if your time bucket uses	Integer >= 0

	month and this value is 1 or more, you won't get alerts for years.	
use_day_of_week	Use the day of the week in the time bucket (Is this normal for Monday?)	boolean
use_hour	Use hour in time bucket (Is this normal for 1am-2am?)	boolean
use_minute	Use minute in time bucket	boolean
use_month	Use month in time bucket	boolean
use_second	Use second in time bucket	boolean
<filter>	EventFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 24 - Pattern of Life Normalcy Analytic options

2.1.21 Stream Inactivity Analytic

`icg.engine.analytic.inactivity.StreamInactivityAnalytic`

Monitors all event streams, sends an alert if there are no events on the data stream in the window_size. Then sends a follow up event when the data stream has new data.

Property	Description	Data Type or Valid Values
window_size	Time window size in seconds	Positive long
<filter>	none	

Table 25 - Stream Inactivity Analytic options

2.1.22 Term Frequency Analytic

`icg.engine.analytic.term.frequency.TermFrequencyAnalytic`

Produces term frequency counts, which are essentially dynamic word clouds, at a specified interval. Attempts to predict future frequencies based on historical trends.

Property	Description	Data Type or Valid Values
data_interval_seconds	Calculation and prediction interval	Positive long
filter_single_terms	Comma separated list of terms to ignore only for terms of length 1	Comma separated string
filter_terms	Comma separated list of terms to ignore globally	Comma separated string
id_attribute_xpath	XPath to ID attribute	XPath string
id_filter_file	Optional file containing list of IDs to include in this analytic	File path
max_term_size	Calculate for terms of size 1 to x	Positive integer
num_subintervals	How many times per DATA_INTERVAL_SECONDS to	Positive integer

	output sub-results	
results_per_term_size	Output the top x terms for each term size	Positive integer
state.file	Saves the analytic's state for a restart	File path
sync.start.to.minute	Synchronize time to minute boundary	boolean
text_attribute_xpath	XPath to text	XPath string
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 26 - Term Frequency Analytic options

2.1.23 Term Trend Analytic

`icg.engine.analytic.term.trend.TermTrendAnalytic`

Produces term frequency counts at a specified interval. Shows increase/decrease from previous interval.

Property	Description	Data Type or Valid Values
data_interval_seconds	Calculation and prediction interval	Positive long
filter_single_terms	Comma separated list of terms to ignore only for terms of length 1	Comma separated string
filter_terms	Comma separated list of terms to ignore globally	Comma separated string
id_attribute_xpath	XPath to ID attribute	XPath string
id_filter_file	Optional file containing list of IDs to include in this analytic	File path
min_unigram_size	Minimum length of a unigram (others will be filtered)	Positive integer
num_subintervals	How many times per DATA_INTERVAL_SECONDS to output sub-results	Positive integer
results_per_term_size	Output the top x terms for each term size	Positive integer
term_sizes	Comma separated list of term sizes to track	Comma separated string
text_attribute_xpath	XPath to text	XPath string
<filter>	GeoAnalyticFilter	

Table 27 - Term Trend Analytic options

2.1.23 Multi Area Association Analytic

`icg.engine.analytic.multiareaassociation.MultiAreaAssociationAnalytic`

Monitors a set of areas loaded from a KML file. When a unique ID is found within `max_distance_to_area` meters of `min_locations` number of them, each at least `min_observations_per_location` times, an analytic event is created. Optionally saves history to disk.

Property	Description	Data Type or Valid Values
id_attribute_xpath	XPath to unique ID attribute	XPath string
id_attribute_name	Name of unique ID for output	String
location_xpath	XPath to location	XPath string
min_locations	The minimum number of locations an ID needs to be seen at to create an analytic event.	long
min_observations_per_location	The minimum number of observations per location needed.	integer
max_distance_to_area_m	Max distance an event can be from an area to count, in meters. 0 means the event geo overlaps or is contained by the area geo.	long
apply_max_distance_to_points_only	If true, MAX_DISTANCE_TO_AREA only applies to areas that are points. Areas that are not points will effectively have a 0 value for MAX_DISTANCE_TO_AREA	boolean
grid_size_m	Used to create geohash grid, size of grid square in meters.	long
kml_file	Path to KML or KMZ areas file	File path
history_file	File to store observation history	File path
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 28 - Multi Area Association Analytic options

2.1.24 Track Shape Similarity Analytic

`icg.engine.analytic.track.analysis.TrackShapeSimilarityAnalytic`

Constructs tracks from data streams, compares the shape of those tracks against image files containing black and white drawings of track shapes. Comparison is done by converting track shapes into binary matrices. Alerts if the similarity is greater than min_match_threshold.

Property	Description	Data Type or Valid Values
window_size_seconds	How much track history to keep, in seconds	long
min_events	Minimum track length before track is analyzed	int

location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to unique ID attribute	XPath string
min_travel_distance_meters	Min length of the track in meters before it is analyzed	long
id_display_name	Display name for ID attribute	String
clear_alerted_data	Whether to clear data for a track after it is alerted on	boolean
event.prop.x.display.name	Display name of a property to be included in output	string
event.prop.x.key	Property key of a property to be included in output	string
event.prop.x.xpath	XPath to a property to be included in output	XPath
aggregate_classifications	Whether to aggregate classifications for all events that go into a track for the analytic event classification	boolean
shape.x.name	Name for shape represented in file x	string
shape.x.file	File path to input file x	File path string
min_match_threshold	Min max percent, 0-100, to report on	double
add_rotations	If true, analytic will match the track against 35 rotations of the input shape, in addition to the original input shape.	boolean
<filter>	GeoAnalyticFilter	

Table 29 - Track Shape Similarity Analytic options

2.1.25 Area Warning Analytic

`icg.engine.analytic.track.areawarning.AreaWarningAnalytic`

Monitors one or more areas specified in a KML/KMZ file for incoming entities. If an entity is projected to be in one of the areas within **`warning_threshold_s`** second, an analytic event is created.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
warning_threshold_s	Time to project tracks into the future for check against warning areas	long
course_degrees_xpath	XPath to course in degrees attribute	XPath string
speed_knots_xpath	XPath to speed in knots attribute	XPath string
GeoFilter properties with "_areawarning" appended, for specifying warning regions	E.g. **`geo_filter_file_areawarning`**	
<filter>	GeoAnalyticFilter	

Table 30 - Area Warning Analytic options

2.1.26 Dead Reckoning Course Speed Analytic

`icg.engine.analytic.deadreckoning.DeadReckoningCourseSpeedAnalytic`

Uses dead reckoning to predict future points on course and speed attributes.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
prediction_seconds_x	Time to project tracks into the future (use 1,2,3… for x)	long
course_degrees_xpath	XPath to course in degrees attribute	XPath string
speed_knots_xpath	XPath to speed in knots attribute	XPath string
<filter>	GeoAnalyticFilter	
prediction_geo_filter_file	KML/KMZ file to filter predictions, predictions that fall within the shapes in the file will not be sent	KML/KMZ file path

Table 31 - Dead Reckoning Course Speed Analytic options

2.1.27 Area Pattern of Life Analytic

`icg.engine.analytic.areapol.AreaPOLAnalytic`

Produces periodic reports with various statistics about a areas specified by a KML/KMZ file.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
report_interval_s	Number of seconds between sending statistics as analytic events	Positive long
history_file_name	Name of file to store statistics	File path string
top_percent_to_display	Percentage of ID stats to report individually	Integer 0-100
GeoFilter properties with "_areapol" appended, for specifying warning regions	E.g. "geo_filter_file_areapol"	
missing_id_evict_time_s	Optional. Amount of time an entity can be considered still in the area without being seen, in seconds.	Long
<filter>	GeoAnalyticFilter	
top_percent_limit	Optional, puts a hard limit on the number of entries that can be produced by "top entities" lists in the alert, which are otherwise controlled by `top_percent_to_display`	Integer
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean

Table 32 - Area Pattern of Life Analytic options

2.1.28 Geospatial Graph Analytic

`icg.engine.analytic.geospatial.graph.GeospatialGraphAnalytic`

Constructs a network graph in memory from geospatial associations between IDs. Periodically writes the graph to disk in GraphML format.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
output_interval_s	How often to write graph to disk in seconds	Long
history_file_name	File to save/load geospatial map	File path string
graphml_state_file_name	File to save GraphML to	File path string
inactive_expire_time_s	Time in seconds before an inactive ID is removed from the geospatial map, preventing future associations until it's seen again	Long
distance_threshold_meters	Maximum distance in meters IDs can be from each other and still be associated	Long
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
kml_output_file	Optional, file to output KML view of graph	File path string
association_output_threshold	Optional, min number of associations a node pair can have and be output to files	integer

Table 33 - Geospatial Graph Analytic options

2.1.29 Geospatial Normalcy Analytic

`icg.engine.analytic.normalcy.geospatial.GeospatialNormalcyAnalytic`

Creates geospatial normal distributions for occurrences of attributes that match a given regex. Wakes up every windowSize seconds and takes a measurement. After numTraining measurements, it will alert when a value falls far enough away from the normal dist curve for that cell. Far enough away is determined by normalcyThreshold, and is related to cumulative probability under the distribution curve.

Property	Description	Data Type or Valid Values
location_xpath	XPath to event location.	XPath string
grid_size_m	Approximate width of single cell in geospatial grid	Postive long
id_attribute_regex	Regex ID attribute must pass to be processed	Regex string

id_attribute_xpath	XPath to ID attribute	XPath string
history_file_name	File path to store history	File path string
count_unique	If only unique IDs should be counted in an area (true), or every event (false), default `false`	Boolean
normalcy_threshold	Maximum normalcy value to alert on, measured 0 (abnormal) to 1 (normal)	Double 0 to 1
binary_mode	Whether to only consider the first zero or non-zero results for an area significant, default `false`	Boolean
num_training_values	Training time in intervals, no alerts during this time	Positive integer
use_day_of_week	Use the day of the week in the time bucket (Is this normal for Monday?)	boolean
use_hour	Use hour in time bucket (Is this normal for 1am-2am?)	boolean
use_minute	Use minute in time bucket	boolean
use_month	Use month in time bucket	boolean
use_second	Use second in time bucket	boolean
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
ids_file	Known IDs, counts will start for these IDs when the analytic starts, even if no events with them have been seen	File path

Table 34 - Entity Geospatial Normalcy Analytic options

2.1.29 Paths Visualization

`icg.engine.analytic.paths.PathsVisualization`

Stores lists of position/time pairs for each ID. Sends out KML periodically.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_xpath	XPath to ID attribute	XPath string
output_interval_s	How often to output KML	Long
kml_output_file	File to save KML to	File path string
distance_threshold_meters	Min distance between points in a path in meters, intermediate points will be discarded	long
min_path_points	Min number of points to display a path	int
id_regex	Regex to filter ID	Regex string
<filter>	GeoAnalyticFilter	

aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
id_display_name	Display name for unique attribute	String
max_path_points	Max number of points to keep per path	int
gc_after_purge	Whether to manually run garbage collector after data purge	Boolean
placemark_titles_enabled	Whether to output titles for placemarks	Boolean
kml_include_points	Whether to include track points	Boolean
kml_point_color	Color for track points	KML color
kml_point_scale	Scale for track points	KML scale
kml_point_icon_url	Icon URL for track points	Icon URL
kml_output_max_length	Max length KML can be and still be sent to UI, in bytes	Long
color_category_attribute_xpath	Optional, xpath to an attribute in the event that will determine the color of the path. All paths with the same value for this attribute will be assigned the same (random) color	Xpath string

Table 34 - Paths Visualization options

2.1.30 Geo Grid Track Forecast Analytic

`icg.engine.analytic.track.forecast.GeoGridTrackForecastAnalytic`

Uses geohash network associations from historical tracks to predict future points on a track. Historical tracks are built from all events that pass the eventFilter. Predictions are made for events that pass the predictionFilter and have tracks of at least min_geohash_track_size_for_prediction geohashes. Tracks that have temporal gaps of more than track_gap_expire_time_s seconds will be removed.

Property	Description	Data Type or Valid Values
location_xpath	XPath to location	XPath string
id_attribute_xpath	XPath to ID attribute	XPath string
grid_size_m	Size of geohash grid in meters	Long
track_gap_expire_time_s	Max amount of time allowed between track updates before it is considered expired and removed, in seconds	Long
max_geohash_chain_size	Max number of consecutive geohashes to associate in graph. Default 4. Increasing this will make predictions more accurate, increase training time, and increase memory requirements.	int

min_geohash_track_size_for_prediction	Minimum number of geohashes present in a track for a prediction to be made, default 2	int
<GeoAnalyticFilter Options> `_prediction`	Optional, GeoAnalytic filter for predictions. An additional filter that events will have to pass before a prediction is created. Non-passing events will still be added to the model.	
<filter>	GeoAnalyticFilter	
aggregate_classifications	Whether to accumulate and combine classifications and data groups	boolean
min_geohash_prediction_size	Optional, specifies the minimum size of the prediction in geohashes. Shorter predictions will not be created.	Positive integer
output_to_disk	If true, forecasts will be written to `forecasts` directory in KML files	Boolean
training_time_s	Optional, time in seconds to wait before making forecasts. Default 0	long
history_file_name	Optional, file to save learned history to	File path string

Table 35 - Geo Grid Track Forecast Analytic options

2.2 Enrichment Plugins

2.2.1 Age Prediction Enrichment

`icg.engine.enrichment.ageprediction.AgePredictionEnrichment`

Uses name and language analysis to predict the age of content authors.

Property	Description	Data Type or Valid Values
name_file	CSV file containing name->avg age mappings	File path
name_path	XPath to name attribute in event	XPath string
output_age_bracket	Whether to output the age bracket that an age falls into (e.g. 18-24)	boolean
output_property	Property name for age	string
output_property_display_name	Display name for age	string
phrase_file	CSV file mapping phrases to age	File path
text_path	XPath to text in event	XPath string
text_starts_with_filter_string	Optional filter to eliminate text that starts with a certain pattern (e.g. “RT”)	string
use_name_on_filtered_text	Whether to continue processing filtered text events for age	boolean

Table 28 - Age Prediction Enrichment options

2.2.2 Gender Prediction Enrichment

`icg.engine.enrichment.genderprediction.GenderPredictionEnrichment`

Use name and phrase analysis to predict the gender authors based on their name and/or text blocks.

Property	Description	Data Type or Valid Values
name_file	File wi	
name_path	CSV file containing name->gender mappings	File path
output_property	Property name for gender	string
output_property_display_name	Display name for gender	string
phrase_file	CSV file mapping phrases to gender	File path
text_path	XPath to text in event	XPath string

Table 29 - Gender Prediction Enrichment options

2.2.3 Geo Proximity Enrichment

`icg.engine.enrichment.geoproximity.GeoProximityEnrichment`

Parse IDs and locations from a CSV file. Add IDs that are within distanceThresholdM meters to the event, as well as a count of the total IDs in proximity. CSV format: id (string), lat (decimal degrees), lon (decimal degrees).

Property	Description	Data Type or Valid Values
count_output_property	Property name for hit count	string
count_output_property_display_name	Display name for hit count	string
data_file	CSV file containing IDs and locations	File path
data_file_poll_interval_s	How frequently to poll the data file for changes, in seconds	Positive integer
distance_threshold_meters	Max distance an event can be from a location to be added	Positive long
hit_output_property	Property name for a location hit	string
hit_output_property_display_name	Display name for a location hit	string
location_path	XPath to event location	XPath string

Table 30 - Geo Proximity Enrichment options

2.2.4 Geo Tagging Enrichment

`icg.engine.enrichment.geotagging.GeoTaggingEnrichment`

Parse names and locations from a kml file. Add name to an event if it is within the geometry of a Placemark.

Property	Description	Data Type or Valid Values
location_path	XPath to event location	XPath string
data_file	KML file containing placemarks	File path
data_file_poll_interval_s	How frequently to poll the data file for changes, in seconds	Positive integer
hit_output_property	Property name for a location hit	string
hit_output_property_display_name	Display name for a location hit	string

Table 31 - Geo Tagging Enrichment options

2.2.5 Group Membership Enrichment

`icg.engine.enrichment.groupmembership.GroupMembershipEnrichment`

Determines which group a given ID has the most associations with. Groups, group members, and associations are loaded through config.

Property	Description	Data Type or Valid Values
connection_output_property	Property name for group membership	string
connection_output_property_display_name	Display name for group membership	string
group.x.name	Name for group x	string
group.x.member.y.file	Followers file for group x, member y	File path
group.x.member.y.name	Name for group x, member y	string
group_loader_type	“Twitter” is the only supported type	string
group_output_property	Property name for group attribute	string
group_output_property_display_name	Display name for group attribute	string
id_path	XPath to ID attribute in event	XPath string

Table 32 - Group Membership Enrichment options

2.2.6 HBOS Anomaly Enrichment

`icg.engine.enrichment.hbos.HBOSAnomalyEnrichment`

Histogram-based Outlier Score (HBOS) algorithm implemented in an analytic. Does unsupervised anomaly detection on numerical and/or categorical text feature sets, scores events from 0 (normal) to 1 (max anomaly). **Configured classification and data groups for this enrichment must represent the highest possible output; data groups and classifications from events are not aggregated automatically in enrichments.**

Property	Description	Data Type or Valid Values
anomaly_threshold	Min anomaly score to output	double 0 to 1
feature_path.x	XPath to feature in event	XPath string
feature_precision.x	Precision of this numerical feature (1, .1, .01, .001, etc)	1, .1, .01, .001, etc
feature_type.x	"STRING" or "NUMBER"	"STRING" or "NUMBER"
history_file	Optional - file to save state to / load state from.	File path
num_training_events	Number of events before we start scoring anomalies	Positive integer
save_interval_ms	Optional - frequency with which to save histogram state, in ms. Default - 180000	Positive long

Table 33 - HBOS Anomaly Enrichment options

2.2.7 IP Address Geoservice Enrichment

`icg.engine.enrichment.ipaddressgeoservice.IpAddressGeoServiceEnrichment`

Add geospatial information (lat, lon, city) for IP addresses found in an event. Geo information can comes from a web service.

Property	Description	Data Type or Valid Values
cache_size	Maximum number of mappings to cache, default 2^20	Positive integer
city_output_display_name	Display name for city property	string
city_output_property	Name for city property	string
geo_output_display_name	Display name for geo property	string
geo_output_property	Name for geo property	string
ip_path	XPath to IP address(es) in event	XPath string

Table 34 - IP Address Geoservice Enrichment options

2.2.8 IP to Geo Enrichment

`icg.engine.enrichment.iptogeo.IPToGeoEnrichment`

Add geospatial information (lat, lon, city, state, country) for IP addresses found in an event. Geo information can comes from a maxmind.geoip2 database file.

Property	Description	Data Type or Valid Values
db_file	Path to maxmind.geoip2 database file	File path
ip_path	XPath to IP address(es) in event	XPath string

Table 35 - IP to Geo Enrichment options

2.2.9 Language Detection Enrichment

`icg.engine.enrichment.languagedetection.LanguageDetectionEnrichment`

Detects language used in a text block.

Property	Description	Data Type or Valid Values
text_path	XPath to text attribute	XPath string

Table 36 - Language Detection Enrichment options

2.2.10 Last Observation Enrichment

`icg.engine.enrichment.last.observation.LastObservationEnrichment`

Adds a property to the event for the amount of seconds since the id attribute of the event was observed in this stream.

Property	Description	Data Type or Valid Values
id_path	XPath to ID attribute in event	XPath string
id_regex	Regex ID must satisfy	Regex string

Table 37 - Last Observation Enrichment options

2.2.11 Link Fetcher Enrichment

`icg.engine.enrichment.linkfetcher.LinkFetcherEnrichment`

Scrapes URLs found in events for articles and adds the contents as properties to events.

Property	Description	Data Type or Valid Values
article_url_query	XPath to article URL in event	XPath string
article_max_chars	Max characters to fetch from URL, default 16384	Positive integer
poll_host_interval	Timeout between HTTP connections	Postive long
max_connections	Max HTTP connections for connection pool, default 16	Positive integer
poll_timeout	HTTP connection timeout in milliseconds, default 10000	Positive long
title_output_property	Property name for article title	string
title_output_display_name	Display name for article title	string
text_output_property	Property name for article text	string
text_output_display_name	Display name for article text	string

Table 38 - Link Fetcher Enrichment options

2.2.12 Matching Enrichment

`icg.engine.enrichment.matching.MatchingEnrichment`

Match multiple attributes using regex, less-than, and/or greater-than operators, then add a property.

Property	Description	Data Type or Valid Values
attribute_arg_x	Argument for matching operation	Regex string or numeric string
attribute_op_x	Operation to perform against attribute	“regex”, “lt”, or “gt”
attribute_path_x	XPath to attribute to match against	XPath string
id_file	Optional list of known IDs to pre-load, also file to store new IDs if maintain_id_list is true	File path
id_path	XPath to ID attribute in event	XPath string
maintain_id_list	Whether to store found IDs in id_file	boolean
output_property_display_name	Display name for output property	string
output_property_value	Value for output property	string

Table 39 - Matching Enrichment options

2.2.13 NLP Enrichment

`icg.engine.enrichment.nlp.NLPEnrichment`

Uses U Illinois NER lib to do entity extraction on text.

Property	Description	Data Type or Valid Values
config_file_path	Path to UofI NLP config file	File path
value_path	XPath to text in event to process	XPath string

Table 40 - NLP Enrichment options

2.2.14 Political Party Enrichment

`icg.engine.enrichment.politicalparty.PoliticalPartyEnrichment`

Use text analysis to predict the political party a block of text is written about.

Property	Description	Data Type or Valid Values
output_property	Property name for party attribute	string
output_property_display_name	Display name for party attribute	string
phrases_file	CSV file with phrases mapped to party affiliation	File path
sites_file	CSV data file with domains mapped to party affiliation	File path
text_path	XPath to text attribute in event	Xpath string

Table 41 - Political Party Enrichment options

2.2.15 Regex Capturing Group Enrichment

`icg.engine.enrichment.regex.RegexCapturingGroupEnrichment`

Use regex against an event attribute to select a group to place the event in to. Group is added as an event property.

Property	Description	Data Type or Valid Values
group_name	Name of capturing group (see Pattern javadoc) to add to event	string
group	Number of capturing group (see Pattern javadoc) to add to event	Numeric string
output_display_name	Display name for output property	string
output_property	Name for output property	string
regex	Regex to match value against	Regex string
value_path	XPath to value attribute in event	XPath string

Table 42 - Regex Capturing Group Enrichment options

2.2.16 Regex Replacement Enrichment

`icg.engine.enrichment.regex.RegexReplacementEnrichment`

Match a series of regex expressions against a text attribute. For each match, replace the matches with the corresponding replacement string.

Property	Description	Data Type or Valid Values
text_path	XPath to text attribute	XPath string
output_property	Property name for output text	string
output_display_name	Display name for output text	string
regex.x.pattern	Regex pattern to find	Regex string
regex.x.replacement	Replacement for regex matches	string

Table 43 - Regex Replacement Enrichment options

2.2.17 Regex Substring Enrichment

`icg.engine.enrichment.regex.RegexSubstringEnrichment`

Use a series of regex expressions to match against an attribute. Return each match.

Property	Description	Data Type or Valid Values
text_path	XPath to text attribute	XPath string
regex.x	Regex to run against text	Regex string

Table 44 - Regex Substring Enrichment options

2.2.18 Shared Data Enrichment

`cg.engine.enrichment.shareddata.SharedDataEnrichment`

Enrichment to add properties to events when Shared Data updates (from Analytic plugins) are found that match an attribute in the event. **Configured classification and data groups for this enrichment must represent the highest possible output; data groups and classifications from events are not aggregated automatically in enrichments.**

Property	Description	Data Type or Valid Values
attribute_path	XPath to the event attribute to match against the Shared Data set	XPath string
data_set	The name of the shared data set from the Analytic plugin to match against	string
output_property_display_name	Display name for the output property	string
output_property_name	Name for the output property	string
output_property_value	Value for the output property	string

Table 45 - Shared Data Enrichment options

2.2.19 Splitter Enrichment

`icg.engine.enrichment.splitter.JavaSplitterEnrichment`

Splits a property using a regex expression and outputs the tokens as properties.

Property	Description	Data Type or Valid Values
output_display_name	Display name for output properties	string
output_property	Property name for output properties	string
regex_expression	Regex to split event attribute with	Regex string
streamname_output_display_name	Option, output the stream name with this display name	string
streamname_output_property	Option, output the stream name with this property name	string
value_path	XPath to value in event to be split	XPath string

Table 46 - Splitter Enrichment options

2.2.20 Stock Symbol Enrichment

`icg.engine.enrichment.stocksymbol.StockSymbolEnrichment`

The enrichment first reads a SYNONYM_FILE_PATH file containing synonyms for abbreviations like inc. and corp. The enrichment expands these abbreviations, when found in company names or search text, to the full form. Then it constructs a Directed Graph (DG) from the stock symbols and company names on STOCK_SYMBOL_FILE_PATH. Each vertex in the DG contains a word

from these symbols and names. In addition, the vertex is terminal if it matches the last word in the symbol or name. (A terminal vertex also contains the stock symbol value.) If the word in the vertex actually is a stock symbol (e.g., not a company name), it is marked exactMatch. So a word in text only matches a stock symbol if it is spelled and capitalized identically. Otherwise, an enrichment wide MAX_CHARS_TO_MATCH value governs how many characters, at most, must match for a word in text to be considered identical to a word in a vertex. If MAX_CHARS_TO_MATCH is 0, all characters (case insensitive) must match. The vertices are connected by directed edges. The vertex containing the first word in a company name is connected to the vertex that contains next word in a company name and so forth. Two vertices that contain the same word compare equal if they have the same number position in the phrase (e.g., if they are the second word in two different company names). A Map of Maps of Sets of Vertices is constructed that maps a word to the map that maps that word to the set of vertices for that word in the n'th position in a phrase. The enrichment works by comparing the input text, one word at a time to all the first rank vertices in the DG. If the DG contains the input word as the first word in a company name or symbol, the enrichment constructs a Searcher for that word and adds it to an existing list of searchers. The enrichment then iterates the searchers. Except for the first time, each searcher contains a list of previous vertices that it has "reached". If one of these vertices can reach the new word, then the searcher is in the MATCHING state. If the vertex that it has reached is terminal, then the searcher is in the MATCH state. If neither condition is true, the searcher is TERMINAL and removed from the searcher list. The first time, the searcher finds the set of vertices that contain the word. If one of these is terminal (e.g., a stock symbol) the state is set to MATCH. Otherwise the state is set to MATCHING. If a searcher reaches MATCHING state, then it has found the word or words it represents in the incoming text, and the symbol associated with the MATCH vertex found is assigned as an event property to the event. If a vertex is marked exact (stock symbols are marked exact) then the symbol and the word found must match exactly in case and length. If the enrichment is running with MAX_CHARS_TO_MATCH > 0, then that number of characters must match, case insensitive, to be a match. If the enrichment is running with SYNONYM_OPTIONAL then the last word is optional, if it originally was a synonym. (e.g., Coca Cola instead of Coca Cola Inc.).

Property	Description	Data Type or Valid Values
allow_duplicates	Whether to allow multiple of the same stock symbol in output	boolean
company_name_column_index	CSV column containing company name	integer
enable_ticker_symbol_searching	Enable searching by symbol in addition to company name	boolean
header_row_count	Number of header rows in the CSV file	integer
max_chars_to_match	A stock symbol name will match if at least this many characters match	integer
max_row_count	Max number of rows to load from file	integer
min_chars_to_match	If > 0, specifies the maximum number of characters, starting from the	integer

	beginning of the word, that must match to consider two words equal.	
output_display_name	Display name for output property	string
output_property	Name for output property	string
stock_symbol_file_path	Points to a CSV file that contains stock symbols in the first column and company names in the second column.	File path
symbol_column_index	CSV column containing stock symbol	integer
synonym_exclusion_file_path	File path to CSV file with synonym exclusions	File path
synonym_file_path	Points to a CSV file that contains abbreviations in the first column (e.g., INC) and fully spelled out synonyms in the second column (e.g., Incorporated).	File path
synonym_optional	Makes company name ending (inc or incorporated, etc) optional	boolean
text_query	XPath to text attribute	XPath string

Table 47 - Stock Symbol Enrichment options

2.2.21 Translation Service Enrichment

`icg.engine.enrichment.translationenrichment.GlpTranslationServiceEnrichment`

Translates a text property from events. Either set lang (LANGUAGE_CODE_PROP), set langQuery (LANGUAGE_QUERY_PROP) or set neither. Setting lang says this IS the language, setting langQuery says here's how to find the language from a field in the input, setting neither says use the local detector to figure out the language and if that fails, use the glp web service detector.

Property	Description	Data Type or Valid Values
common_text_display_name	Display name for default text	string
common_text	Default text to add if there’s none in text_query	string
language_code	Optional, says this is the language	string
language_query	Optional, says here's how to find the language from a field in the input	XPath string
language_code_translation_table	CSV file containing mapping for language codes	File path
text_query	XPath to text attribute	XPath string

Table 48 - Translation Service Enrichment options

2.2.22 Tweet Extractor Enrichment

`icg.engine.enrichment.tweetextractorenrichment.TweetExtractorEnrichment`

Uses Twitter Extractor to extract mentions, hashtags, cashtags, reply screennames, and URLs.

Property	Description	Data Type or Valid Values
cashtag_display_name	Display name for cashtag property	string
cashtag_property	Name for cashtag property	string
date_format_string	Format string to parse created date	Format string
followers_path	XPath to followers attribute	XPath string
following_path	XPath to following attribute	XPath string
hashtag_display_name	Display name for hashtag property	string
hashtag_property	Name for hashtag property	string
mentioned_display_name	Display name for mentions property	string
mentioned_property	Name for mentions property	string
profile_created_date_path	XPath to profile created date attribute	XPath string
reply_screenname_display_name	Display name for reply screenname property	string
reply_screenname_property	Name for reply screenname property	string
text_path	XPath to text attribute	XPath string
url_display_name	Display name for url property	string
url_property	Name for url property	string

Table 49 - Tweet Extractor Enrichment options

2.2.23 Twitter Influence Enrichment

`icg.engine.enrichment.twitterinfluence.TwitterInfluenceEnrichment`

Assign scores to Twitter users based on their followers, avg. retweets, and tweet interval. Also categorize authors as Individuals, News Organizations, or Company / Other.

Property	Description	Data Type or Valid Values
followers_path	XPath to followers attribute	XPath string
history_file_name	Optional file to store user/tweet history	File path
min_followers	Minimum number of followers a user can have to rate them	Positive integer
names_file_name	Path to data file containing names	File path
num_tweets_per_user	How many tweets per user to store per user for analysis. The latest tweets are stored.	Positive integer
orig_retweet_count_path	XPath to original tweet retweet count	XPath string
orig_tweet_id_path	XPath to original tweet ID	XPath string
text_path	XPath to tweet text	XPath string

user_id_path	XPath to user ID	XPath string
user_name_path	XPath to user name	XPath string

Table 50 - Twitter Influence Enrichment options

2.2.24 Twitter Main Subject Enrichment

`icg.engine.enrichment.twitter.mainsubject.TwitterMainSubjectEnrichment`

Extract the main mention of the tweet

Property	Description	Data Type or Valid Values
orig_text_xpath	XPath to original tweet text	XPath string
text_xpath	XPath to tweet text	XPath string

Table 51 - Twitter Main Subject Enrichment options

2.2.25 UofI Ethnicity Enrichment

`icg.engine.enrichment.uoiethnicity.UoIEthnicityEnrichment`

Use University of Illinois Ethnicity Classifier (and the result of a previous gender prediction from the GenderPredictionEnrichment) to predict the ethnicity of a name.

Property	Description	Data Type or Valid Values
attribute_filter_negative_match_value	Filter value for attribute_filter_query attribute to ignore if equal	string
attribute_filter_query	XPath to filter attribute (does not use EventFilter)	XPath string
config_file_path	Path to UofI config file	File path
contains_any_negative_filter	Filter value for primary_text attribute to ignore if it contains this value	string
main_ethnicity_display_name	Display name for ethnicity attribute	string
main_ethnicity	Name for ethnicity attribute	string
gender_prediction_query	XPath to gender property in event	XPath string
primary_text_query	XPath to text attribute in event	XPath string

Table 52 - UofI Ethnicity Enrichment options

2.2.26 UofI Sentiment Enrichment

`icg.engine.enrichment.uoisentiment.UoISentimentEnrichment`

Uses University of Illinois sentiment classifier to enrich events with a sentiment. VSSentimentEnrichment should be used instead.

Property	Description	Data Type or Valid Values
fallback_text_query	Optional, XPath to fallback text	XPath string

main_sentiment	Name for sentiment property	string
main_sentiment_display_name	Display name for sentiment property	string
primary_text_query	XPath to primary text	XPath string

Table 53 - UofI Sentiment Enrichment options

2.2.27 URL Enrichment

`icg.engine.enrichment.url.URLEnrichment`

Resolves shortened URLs and optionally rates them according to bias and validity.

Property	Description	Data Type or Valid Values
expanded_urls_display_name	Display name for expanded URL	string
expanded_urls_property	Property name for expanded URL	string
url_bias_file	Optional, CSV file with domain names mapped to bias ratings	File path
url_path	XPath to URL attribute in event	XPath string

Table 54 - URL Enrichment options

2.2.28 URL Splitter Enrichment

`icg.engine.enrichment.splitter.URLSplitterEnrichment`

Splits a URL into its parts and optionally adds each one as a property.

Property	Description	Data Type or Valid Values
include_host	Whether to output the host as a property	boolean
include_path	Whether to output the path as a property	boolean
include_port	Whether to output the port as a property	boolean
include_protocol	Whether to output the protocol as a property	boolean
include_query	Whether to output the query as a property	boolean
url_path	XPath to URL attribute in event	XPath string

Table 55 - URL Splitter Enrichment options

2.2.29 Value Map Enrichment

`icg.engine.enrichment.valuemap.ValueMapEnrichment`

Adds a property to an event if an attribute matches one of the configured values, default value also optionally supported. Values can be configured in `ae.xml`, and/or specified in a CSV file.

Property	Description	Data Type or Valid Values
default_value	Value to add if the value in the event isn't found in the match-groups	string
match-group.x.key	Key value for this match group, must match exactly	string
match-group.x.value	Value to add if this match group is matched with the value in the event	string
match-group.csv.file	Optional, path to CSV file with mappings	File path
output_display_name	Display name for output property	string
output_property	Name for output property	string
source.poll.interval	How frequently to poll the CSV file for updates, in milliseconds	Positive long
value_path	XPath to value attribute in event	XPath string

Table 56 - Value Map Enrichment options

2.2.30 Value Range Enrichment

`icg.engine.enrichment.valuerange.ValueRangeEnrichment`

Use a series of numeric ranges (inclusive) to place an attribute into a group

Property	Description	Data Type or Valid Values
output_property	Name of property to output	string
output_property_display_name	Display name of property to output	string
range.x.max	Max value for group x (inclusive)	Double string
range.x.min	Min value for group x (inclusive)	Double string
range.x.name	Group name for group x	string
value_path	XPath to value attribute in event	XPath string

Table 57 - Value Range Enrichment options

2.2.31 VS Sentiment Enrichment

`icg.engine.enrichment.sentiment.VSSentimentEnrichment`

Detect sentiment and emotion in text attributes using a word list of varying +/- scores for each word.

Property	Description	Data Type or Valid Values
contains_at_usernames	Whether the text attribute contains twitter-style username	boolean
contains_hashtags	Whether the text attribute contains hashtags	boolean
emotion_output_property	Emotion output property name	string

emotion_output_property_display_name	Emotion output property display name	string
emotion_words_file	CSV file that maps words to emotions	File path
output_emotion	Whether or not to evaluate emotion	boolean
output_property	Sentiment output property name	string
output_property_display_name	Sentiment output property display name	string
text_path	XPath to text value in event	XPath string
words_file	CSV file that maps words to sentiment scores between -5 and 5	File path

Table 58 - VS Sentiment Enrichment options

2.2.32 Word2Vec Enrichment

`icg.engine.enrichment.word2vec.Word2VecEnrichment`

Use Word2Vec neural net to categorize text using similarity to one or more subjects specified in the config.

Property	Description	Data Type or Valid Values
model_file	Word2Vec model file path	File path
subject.x	Subjects to run similarity against	string
text_path	XPath to text attribute in event	XPath string

Table 59 - Word2Vec Enrichment options

2.2.33 OCR Enrichment

`icg.engine.enrichment.ocr.OCREnrichment`

Performs OCR on images and adds text as a property.

Property	Description	Data Type or Valid Values
image_url_path	XPath to event property containing image URL	XPath string
output_property_name	Name of property to output, default **`ocr_text`**	string
output_property_display_name	Display name of property to output, default **`OCR Text`**	string
tesseract_data_path	Path to directory containing **`tessdata`** directory	File path

Table 60 - OCR Enrichment options

2.2.34 Course Speed Projection Enrichment

`icg.engine.enrichment.coursespeed.CourseSpeedProjectionEnrichment`

Projects a declared course and speed into the future by a number of configurable intervals, adds geos representing those future points.

Property	Description	Data Type or Valid Values
course_degrees_xpath	XPath to event property containing image URL	XPath string
speed_knots_xpath	Name of property to output, default `ocr_text`	XPath string
location_xpath	Display name of property to output, default `OCR Text`	XPath string
prediction_seconds_x	Path to directory containing `tessdata` directory	long
prediction_geo_filter_file	Path to KML/KMZ filter file for predictions	File path string
<filter>	GeoAnalyticFilter	
prediction_point_style_name	Style name for points, default `predicted_point`	String

Table 61 - Course Speed Projection Enrichment options

2.3 Alerter Plugins

2.3.1 Cloudant Alerter

`icg.engine.alerter.cloudant.CloudantAlerter`

Outputs alerts as JSON documents to a Cloudant database.

Property	Description	Data Type or Valid Values
lux.alert.format.title	Optional, alert title format string for alerts, default: `${alert.classification} ${alert.title}`	Alert title format string
cloudant.alert.database	Cloudant database name in which to store alerts	string
cloudant.alert.key	Key for Cloudant instance	string
cloudant.alert.password	Password for Cloudant instance	string
cloudant.alert.uri	URI to Cloudant instance	URI string
alert.priority	Optional, alert priority 1-5, default is 1	“1”,”2”,”3”,”4”,or “5”

Table 62 - Cloudant Alerter options

2.3.2 Console Alerter

`icg.engine.alerter.ConsoleAlerter`

Sends alert to stdout.

2.3.3 DNAI Alerter

`icg.engine.alerter.dnai.DNAIAlerter`

An Alerter for DNAI rules, sends updated NAIs to NAI manager

2.3.4 Email Alerter

`icg.engine.alerter.email.EmailAlerter`

Sends GeoRSS XML alerts via email, optionally transformed by XSLT. KML will be attached if it was specified in the alert (option comes from rule forms).

Property	Description	Data Type or Valid Values
mail.debug.kml	Optinal, whether to save KML to a file for debug, default false	boolean
mail.alert.xslt	Optional, XSLT file to process alert XML with	File path
mail.subject	Optional, subject line for email alerts, default “LUX Alert!”	string
mail.send.email	Whether to send emails, default true	boolean
mail.send.logger	Whether to send emails to log file, default false	boolean
mail.threadcount	How many email sender threads to run	Positive integer

Table 63 - Email Alerter options

2.3.5 Legacy File Alerter

`icg.engine.alerter.file.LegacyFileAlerter`

Alerter that stores alerts in flat files.

Property	Description	Data Type or Valid Values
file.alert.filename.format	Optional, file name format string for alerts. Default "alerts/alert%d.txt"	Format string, one %d for alert number
file.alert.xslt	Optional, XSLT file to process alert XML with	File path
rule.name.regex	Regex that the alert’s rule name must match in order to process	Regex string
file.alert.threadcount	How many threads to run	Positive integer

Table 64 - Legacy File Alerter options

2.3.6 IRC Alerter

`icg.engine.alerter.irc.IRCAlerter`

Alerter that sends messages to a configurable IRC server.

Property	Description	Data Type or Valid Values
irc.alert.xslt	XSLT file to transform alerts, default `xslt/irc.xslt`	File path
irc.channel	IRC channel to connect to, default “engineAlerts”	string
irc.connection.host	IRC host to connect to	URL
irc.connection.name	IRC “real name”	string
irc.connection.nick	IRC nickname	string
irc.connection.password	IRC password	string
irc.connection.port	IRC connection port	Positive integer
irc.connection.ssl	Whether or not to use SSL	boolean
irc.connection.user	IRC connection username	string

Table 65 - IRC Alerter options

2.3.7 JMS Alerter

`icg.engine.alerter.jms.JMSAlerter`

Alerts to a JMS queue or topic. Spring loaded.

Property	Description	Data Type or Valid Values
jms.spring.path	Path to spring file to load JMS beans	File path
jms.connection.factory.bean.name	Spring bean for	string
jms.broker.username	Username for JMS broker	string
jms.broker.password	Password for JMS broker	string
jms.alertqueue.name	Queue name to send alerts to. Specify either jms.alertqueue.name or jms.alerttopic.name.	string
jms.alerttopic.name	Topic name to send alerts to. Specify either jms.alertqueue.name or jms.alerttopic.name.	string
jms.producer.threads	How many JMS threads to run, if value <= 0, default of 1-per-core will be used	Positive integer
jms.producer.use.message.id	Optional, whether to use the message’s ID, default true	boolean
jms.producer.use.message.timestamp	Optional, whether to use the message’s timestamp, default true	boolean
jms.producer.delivery.mode	Optional, JMS delivery mode, default `PERSISTENT`	`PERSISTENT`, `NON_PERSISTENT`, or `RELIABLE`
jms.producer.max.transaction.size	Optional, JMS transaction side, default 0	Integer string
jms.alert.xslt	Optional XSLT file to transform alerts	File path

Table 66 - JMS Alerter options

2.3.8 Kafka Alerter

`icg.engine.alerter.kafka.KafkaAlerterV9`

Sends alerts to Kafka topics, works with Kafka v0.9.x.

Property	Description	Data Type or Valid Values
bootstrap.servers	Kafka connection URL	URL
num.threads	Number of threads to start	Positive integer
topic	Kafka topic to send to	string

Table 67 - Kafka Alerter options

2.3.9 List Alerter

`icg.engine.alerter.list.ListAlerter`

Alerter that writes an attribute from alert's events to a `.list` file for use in Enrichments and Analytics.

Property	Description	Data Type or Valid Values
list.name	Name of the list (and list file) to send attributes to	string
threadcount	Number of threads to start	Positive integer
value.xpath	XPath to attribute in the alert's events to write to list	XPath string

Table 68 - List Alerter options

2.3.10 LUX Alerter

`icg.engine.alerter.jms.lux.LUXJsonAlerter`

Alerter to send JSON alerts to the LUX UI.

Property	Description	Data Type or Valid Values
Same properties as JMS Alerter		
lux.alert.format.title	Optional, format string for alert title, default `${alert.classification} ${alert.title}`	Alert title format string

Table 69 - LUX Alerter options

2.3.11 LUX Email Alerter

`icg.engine.alerter.email.lux.LUXJsonEmailAlerter`

Sends email alerts in LUX JSON format

Property	Description	Data Type or Valid Values
mail.debug.kml	Optinal, whether to save KML to a file for debug, default false	boolean
mail.subject	Optional, subject line for email alerts, default “LUX Alert!”	string
mail.send.email	Whether to send emails, default true	boolean
mail.send.logger	Whether to send emails to log file, default false	boolean
mail.threadcount	How many email sender threads to run	Positive integer
mail.alert.json.template.dir	Directory path to load Freemarker templates from, default “templates”	Directory path
mail.alert.json.template	Optional, Freemarker template to use to transform alerts if no `mail.alert.json.template` property is found in the alert params	File name
mail.alert.kml.content_type	Content-type for attached KML, default “text/kml”	Content type string
mail.alert.format.title	Title format for alert titles (prop wording is incorrect), default `${alert.classification} ${alert.title}`	Alert title format string
mail.alert.kml.default.icon.url	KML icon URL, needs to be specified for KML attachment to work. Example `https://dev3.icgsolutions.com/lux/googleearth/images/red-circle.png`	URL
mail.alert.kml.lux.webapp.url	URL to LUX UI webapp, required for attached KML. Example `https://dev3.icgsolutions.com/lux`	URL

Table 70 - LUX Email Alerter options

2.3.12 Legacy Socket Alerter

`icg.engine.alerter.socket.LegacySocketAlerter`

Send alerts over a socket.

Property	Description	Data Type or Valid Values
socket.alerter.host	Host to send alerts to, default localhost	Host string
socket.alerter.port	Port to send alerts on	Positive integer
socket.alerter.protocol	Only supports TCP at this time, this parameter can be ignored.	string
socket.alerter.xslt	Optional, XSLT file to transform alerts	File path

Table 71 - Socket Alerter options

2.3.13 SQS Alerter

`icg.engine.alerter.sqs.SQSAlerter`

SQSAlerter outputs to the AWS SQS service. SQS credentials are read from `~/.aws/credentials` file.

Property	Description	Data Type or Valid Values
lux.alert.format.title	Title format for alert titles (prop wording is incorrect), default `${alert.classification} ${alert.title}`	Alert title format string
freemarker.template	Optional, path to Freemarker template file	File path
num.threads	Number of threads to run	Positive integer
queue.url	SQS queue URL	URL
sqs.region	Supports: `us-east-1`, `us-west-1`, `us-west-2`, `current`, and `default`	SQS region string

Table 72 - SQS Alerter options

2.3.14 LUX JSON File Alerter

`icg.engine.alerter.file.lux.LUXJsonFileAlerter`

Alerter that stores alerts in flat files. Works with LUX JSON alerts and has an option to put them through a freemarker template.

Property	Description	Data Type or Valid Values
file.alert.filename.format	Optional, file name format string for alerts. Default "alerts/alert%d.txt"	Format string, one %d for alert number
rule.name.regex	Regex that the alert's rule name must match in order to process	Regex string
file.alert.threadcount	How many threads to run	Positive integer
lux.alert.format.title	Title format for alert titles, default `${alert.classification} ${alert.title}`	Alert title format string
lux.alert.json.template.dir	Directory path to load Freemarker templates from, default "templates"	Directory path
lux.alert.json.template	(optional) Freemarker template to use to transform alerts	File name

2.3.15 LUX JSON Socket Alerter

`icg.engine.alerter.socket.lux.LUXJsonSocketAlerter`

Send alerts over a socket. Works with LUX JSON alerts and has an option to put them through a freemarker template before passing them to an AlertEncoder.

Property	Description	Data Type or Valid Values
socket.alerter.host	Host to send alerts to, default localhost	Host string
socket.alerter.port	Port to send alerts on	Positive integer
socket.alerter.protocol	Only supports TCP at this time, this parameter can be ignored.	string
socket.alerter.xslt	Optional, XSLT file to transform alerts	File path
lux.alert.format.title	Title format for alert titles, default `${alert.classification} ${alert.title}`	Alert title format string
lux.alert.json.template.dir	Directory path to load Freemarker templates from, default “templates”	Directory path
lux.alert.json.template	(optional) Freemarker template to use to transform alerts	File name

2.4 Event Ingest Plugins

2.4.1 Bright Planet Ingest

`icg.engine.ingest.brightplanet.BrightPlanetIngest`

Polls the Bright Planet web service for the latest events.

Property	Description	Data Type or Valid Values
enable_translator	Whether to translate articles with a RESTful translation service, default true	boolean
base.url	Base URL for BP REST service	URL
poll.interval	How often to poll service, in milliseconds	integer
api.key	API key for BP service	string
initial.poll.offset	Initial polling delay, in milliseconds	integer
poll.timeout	Socket/connection timeout for poll call, in milliseconds	integer
data.feed	BP data feed to ingest, default “bits”	string

Table 72 - Bright Planet Ingest options

2.4.2 Cloudant Ingest

`icg.engine.ingest.cloudant.CloudantIngest`

Ingests events from a Cloudant database.

Property	Description	Data Type or Valid Values
batch.size	How many rows to request per batch, default 20	integer

cloudant.database	Cloudant database name	string
cloudant.key	key to use to connect to database	string
cloudant.password	password that goes with that key	string
cloudant.uri	URI to cloudant database	URI
cycle.time.seconds	How long to wait after a batch returns nothing before trying to read from Cloudant again, in seconds. Default 60	long
encoding	Data encoding, default "UTF-8"	string

Table 73 - Cloudant Ingest options

2.4.3 Email Ingest

`icg.engine.ingest.email.EmailIngest`

Ingests email from a mail account. After the email is read, the email can be marked unread so that it is read again, or left marked read, so that only newly received email is read.

Property	Description	Data Type or Valid Values
email_poll_secs	The number of seconds after reading email to wait before again trying to read email. If, when a new cycle is started, the queue has a size that is greater than QUEUE_REFILL_SIZE, then the producer skips that cycle and waits EMAIL_POOL_SECS before trying again to fill the queue.	Positive integer
queue_max_size	The max number of email messages that can be in the queue, default 600	Positive integer
queue_refill_size	The point the queue is allowed to drain to before refilling, default 60	Integer >= 0
encoding	Data encoding, default "UTF-8"	string
event.parser	Name of Java class containing EventParser implementation to use on event contents	Fully qualified class name
email_host	Email server host URL	URL
username	Username to log into email server	string
password	Password for email server	string
retrieve_unread_only	Whether to only retrieve unread emails, default true	boolean
mark_as_read	Whether to mark retrieved emails as read, default true	boolean
subject_filter	Optional, Search term for subject line	string
body_filter	Optional, Search term for email body	string

Table 74 - Email Ingest options

2.4.4 Facebook Ingest

`icg.engine.ingest.facebook.FacebookIngest`

Follows selected accounts and reads public posts and comments from their walls.

Property	Description	Data Type or Valid Values
cycle.time.seconds	How often to poll facebook for user feed, in seconds	Positive integer
history.file	Holds the state, written after every cycle, read on startup	File path
dot.encode	Replace control characters with dots in the message, necessary because they don't encode in xml. Default: false	boolean
rescan.old.posts	If true, run in exhaustive mode, searching entire graph for changes. if false, use since to look only at changes and go deep if updatedTime indicates something changed. This doesn't work since updateTime isn't updated on graph nodes that literally don't change when graph nodes 2+ deep below them do. Default: true	boolean
oauth.access.token	Access token in the form <application-id>|<application-secret> see http://developers.facebook.com/apps	Access token string
users	A comma separated list of user "NODE" names. A node name is a name you can enter as a URL to facebook and get a username and user id. e.g., http://developers.facebook.com/tools/explorer and enter JebBush in the Graph API "GET" field and click "Submit"	CSV string
since.age.seconds	The number of seconds of data to retain. The ingester will maintain a list of a users posts and salient information about those posts covering from "now minus since.age.seconds" until "now". Default: 1 week('s worth of seconds)	Positive long
include.likes	Include likes in the event stream. Default: true	boolean
include.comments	Include comments in the event stream. Default: true	boolean

include.profiles	Include profiles in the event stream. . Default: true	boolean
thread.count	How many worker threads to spin up. Deafult: 5	Positive integer

Table 75 - Facebook Ingest options

2.4.5 File Ingest

`icg.engine.ingest.file.FileIngest`

Ingest plugin for ingesting files from a directory. After the files are read, optionally recursively, they can be kept, deleted, or moved.

Property	Description	Data Type or Valid Values
dest.folder	Optional, destination folder for files after they are processed, goes with process mode MOVE	Folder path
exclude.pattern	Optional, regex pattern to use to filter files	Regex string
event.folder	Folder containing events to monitor	Folder path
process.mode	What to do with files after they are processed `DELETE`, `MOVE`, or `KEEP` `DELETE`: delete files after processing `MOVE`: move files to dest.folder after processing `KEEP`: keep files in folder (will be re-processed with scan.mode of `POLL`)	`DELETE`, `MOVE`, or `KEEP`
recurse	Whether or not to recursively process directories under event.folder. Deafult false	boolean
scan.interval	How often to scan event.folder for new files, in milliseconds. Values less than 1 mean to only scan once.	Integer
scan.mode	`POLL` or `MONITOR`. `POLL`: process all files in directory every n milliseconds. `MONITOR`: process new files as they are added	`POLL` or `MONITOR`
thread.count	How many threads to run, default 1	Positive integer

Table 76 - File Ingest options

2.4.6 FTP File Ingest

`icg.engine.ingest.ftp.FTPFileIngest`

Ingest plugin for ingesting files from a remote directory via FTP. After the files are read, optionally recursively, they can be kept or deleted.

Property	Description	Data Type or Valid Values
ftp.event.directory	Remote event directory to process	File path
ftp.event.local.copy.directory	Optional, local directory to copy remote files to	File path
exclude.pattern	Optional, files matching this pattern will not be processed	Regex string
ftp.host	Hostname for FTP connection	URL
ftp.password	Password for FTP connection	string
ftp.port	Port number for FTP connection	Positive integer
ftp.user	Username for FTP connection	string
include.pattern	Optional, files matching this pattern will be processed	Regex string
ftp.poll.time.seconds	How often to poll FTP directory for new files, in seconds. Default: 300	Positive integer
process.mode	What to do with files after they are processed **`DELETE`** or **`KEEP`** **`DELETE`**: delete files after processing **`KEEP`**: keep files in folder	**`DELETE`** or **`KEEP`**
ftp.reconnect.attempts	How many times to attempt to connect to FTP server, default 3	Positive integer
ftp.reconnect.time.seconds	How many seconds between connection attempts, default 60	Positive integer
ftp.client.class	FTP or SFTP class	**`icg.engine.ingest.ftp.FtpClient`** or **`icg.engine.ingest.ftp.SFtpClient`**

Table 77 - FTP File Ingest options

2.4.7 JMS Ingest

`icg.engine.jms`

Ingest events from a JMS queue or topic.

Property	Description	Data Type or Valid Values
encoding	Data encoding, default UTF-8	string
jms.spring.path	Path to spring file to load JMS beans	File path
jms.connection.factory.bean.name	Spring bean for	string
jms.broker.username	Username for JMS broker	string
jms.broker.password	Password for JMS broker	string
jms.queue.name	Queue name to read events from. Specify either jms.queue.name or jms.topic.name.	string
jms.topic.name	Topic name to read events from.	string

	Specify either jms.queue.name or jms.topic.name.	
jms.consumer.threads	How many JMS threads to run, if value <= 0, default of 1-per-core will be used	Positive integer
jms.consumer.max.transaction.size	Max JMS transaction size, default 0	Positive integer
jms.client.id	JMS client ID	string
jms.topic.subscriber.name	Durable subscriber ID. If this is specified, must also specify jms.client.id and jms.topic.name	string

Table 78 - JMS Ingest options

2.4.8 Kafka Ingest V8

`icg.engine.ingest.kafka.KafkaIngestV8`

Use with Kafka v0.8.2.x, Ingest plugin for reading from a Kafka topic.

Property	Description	Data Type or Valid Values
avro.mode	Whether or not to use Avro serialization, default false	boolean
avro.schema.url	Optional is Avro is being used, URL to schema	URL
num.threads	Number of threads to run	Positive integer
topic	Kafka topic to read from	string
zookeeper.node.port	Kafka connection URL	URL

Table 79 - Kafka Ingest V8 options

2.4.9 Kafka Ingest V9

`icg.engine.ingest.kafka.KafkaIngestV9`

Use with Kafka v0.9.x, Ingest plugin for reading from a Kafka topic

Property	Description	Data Type or Valid Values
bootstrap.servers	Kafka connection URL	URL
client.id	Client ID for Kafka Consumer, default “client-01”	string
group.id	Group ID for Kafka Consumer, default “lux-group”	string
start.from.beginning	Whether to start ingesting from the beginning of the topic, or only read new data. Default false	boolean
topic	Kafka topic to read from	string
zookeeper.connect	Zookeeper connect URL, only needed if start.from.beginning is true	URL

Table 80 - Kafka Ingest V9 options

2.4.10 Pastebin Ingest

`icg.engine.ingest.pastebin.PastebinIngest`

Scrapes pastebin API for all new pastes. Machine that this runs from will need to be in the Pastebin Scraping API whitelist. No configuration options.

2.4.11 Postgres Ingest

`icg.engine.ingest.postgres.PostgresIngest`

Ingests events from a Postgresql database. This ingester maintains a small amount of state - the sequence of the row of the last batch of rows read so that if the ingester is restarted, it restarts from more or less where it left off.

Property	Description	Data Type or Valid Values
cycle.time.seconds	How long to wait, in seconds, after a batch returns nothing before trying to read from Postgres again.	Positive integer
max.batch.size	How many rows to request per batch	Positive integer
postgres.password	Password for database connection	string
postgres.uri	URI of the Postgres database in form jdbc:postgresql://host:port/database	URL
postgres.user	User for database connection	string
state.file.name	Filename of the file that contains persistent state for the ingester	File path

Table 81 - Postgres Ingest options

2.4.12 RSS Ingest

`icg.engine.ingest.rss.RSSIngest`

Polls an RSS feed for updates. For each item in the feed that's new, it creates a copy of the feed as if it contained only that item and sends it as an event. Previously seen items are only stored in memory, so restarting the ingest will cause the old items to be repeated if they're still listed in the feed. Uses very loose parsing, so it can handle different versions of RSS as well as Atom.

Property	Description	Data Type or Valid Values
rss.event.xslt	Optional, XSLT file to transform RSS events	File path
rss.feed.poll.interval.seconds	How often to poll RSS feeds, in seconds. Default 600	Positive long
rss.feed.url	URL to RSS feed	URL

Table 82 - RSS Ingest options

2.4.13 RSS Link Fetcher

`icg.engine.ingest.rss.linkfetcher.RssLinkFetcher`

Extends RSSIngest to fetch articles linked to in RSS feed.

Property	Description	Data Type or Valid Values
rss.event.xslt	Optional, XSLT file to transform RSS events	File path
rss.feed.poll.interval.seconds	How often to poll RSS feeds, in seconds. Default 600	Positive long
rss.feed.url	URL to RSS feed	URL
article.max.chars	Maximum number of characters to fetch for an article, default 16384	Positive integer
poll.timeout	Timeout for HTTP connection to fetch articles	Positive integer
queue.size	Size of queue for HTTP requests, default 16	Positive integer
rss.language	Optional, adds rss.language property to events	string
poll.host.interval	How long to sleep between polling hosts, in milliseconds. Default 1	Positive long

Table 83 - RSS Link Fetcher Ingest options

2.4.14 Socket Ingest

`icg.engine.ingest.socket.SocketIngest`

Ingest plugin for reading from a socket

Property	Description	Data Type or Valid Values
socket.ingest.port.base	Base port to try opening, default 6456	Positive integer
socket.ingest.port.max	Max port to try opening, default 6466	Positive interger
socket.ingest.protocol	Socket protocol. `TCP` or `UDP`, default `UDP`	`TCP` or `UDP`

Table 84 - Socket Ingest options

2.4.15 Twitter Ingest

`icg.engine.ingest.twitter.TwitterIngest`

Uses Twitter API to ingest samples from specified topics and/or users.

Property	Description	Data Type or Valid Values
capture.end.hhmm	Optional, end time to write data to capture.folder, format HH:mm	HH:mm string

capture.folder	Optional, the name of a folder in which to put 1 file for every message received	File path
capture.start.hhmm	Optional, start time to write data to capture.folder, format HH:mm	HH:mm string
cKey	OAuth key for Twitter API authentication	string
cSec	OAuth secret key for Twitter API authentication	string
followings	Comma separated list of users to follow	CSV string
languages	Optional, comma separated list of 2-character language codes to filter on	CSV string
locations	A JSON string of up to 25 locations. Each location is a lat-lon rectangle, expressed as either an array of four numbers (swLon, swLat, neLon, neLat) or an array of two arrays of two numbers [[swLon, swLat], [neLon, neLat]]. Multiple locations can be wrapped in an array, or as values in a JSON object node (e.g. { "location1": [0,0,1,1], "location2": [1,2,3,4], ...}) (the field names don't matter; the names only serve to make the list more readable by humans.)). These wrapping methods can be nested; the parser will drill down recursively until it finds arrays of numbers.	JSON string
tok	OAuth token for Twitter API authentication	string
tokSec	OAuth token secret for Twitter API authentication	string
track	Comma separated list of subjects to follow	CSV string
ocr.enabled	Whether to run the tesseract-ocr library to attempt OCR against any images in the incoming tweets. Requires tesseract-orc and deps to be installed on the machine.	boolean
tesseract.data.path	Required if ocr.enabled is true, path to folder containing tessdata folder	File path

Table 85 - Twitter Ingest options

2.4.16 Blockchain Ingest

`icg.engine.ingest.websocket.blockchain.BlockchainIngest`

Ingests bitcoin transactions from blockchain.info. No configuration options.

2.4.17 YouTube Ingest

`icg.engine.ingest.youtube.YouTubeIngest`

Uses YouTube API to ingest video metadata from specified search parameters

Property	Description	Data Type or Valid Values
api_key	YouTube API Key	String
search_query	Search query string, you can use | for OR, - for NOT example: foo|bar -baz	String
location	Optional, lat/lon coordinates for geo search. Must be specified with location_radius example: 37.42307,-122.08427	String
location_radius	Optional, circular search radius from location point. Must be a float followed by unit designation [m,km,ft,mi] max 1000km. Must be specified with location. examples: 100m, 500km, 3.6mi	String

Table 86 - YouTube Ingest options

2.4.18 Reddit Ingest

`icg.engine.ingest.reddit.RedditIngest`

Uses Reddit API to scrape the main page periodically and ingest new submissions and comments.

Property	Description	Data Type or Valid Values
api_id	Reddit API application ID	String
api_secret	Reddit API application secret	String
username	Reddit username	String
password	Reddit password	String
submissions_stream_name	Stream to send sumissions on, default: reddit_submissions_stream	String
comments_stream_name	Stream to send comments on, default: reddit_comments_stream	String
subreddits	CSV string of subreddits to monitor, e.g. news,worldnews,all	CSV string

Table 87 - RedditIngest options

2.4.19 WMATA Ingest

`icg.engine.ingest.wmata.WMATAIngest`

Washington Metropolitan Area Transit Authority bus position ingest.

Property	Description	Data Type or Valid Values
api_key	WMATA API key	String

Table 88 - RedditIngest options

2.4.20 Postgres Custom SQL Ingest

`icg.engine.ingest.postgres.PostgresCustomSQLIngest`

Ingests data from a Postgresql database using a dynamic file on disk to load SQL commands. Each time the file is changed, the SQL is executed to fetch data. Designed for forensic (non-realtime) mode.

Property	Description	Data Type or Valid Values
postgres.password	Password for database connection	string
postgres.uri	URI of the Postgres database in form jdbc:postgresql://host:port/database	URL
postgres.user	User for database connection	string
sql.file	Path to file that will contain sql	File path

Table 89 - Postgres Custom SQL Ingest options

2.5 Event Parsers

Many ingest plugins have the ability to load EventParsers to parse events into LUX format after they are retrieved from a data source. Below are some of the generically-applicable EventParser implementations.

2.5.1 CSV File Parser

`icg.engine.event.ingest.data.parsers.csv.CsvFileParser`

Parses CSV rows from an InputStream. Supports custom parsing through the IFileEvent interface specified in the event.handler property, below are the configuration options using GenericFileEvent as the event handler.

Property	Description	Data Type or Valid Values
event.handler	IFileEvent class to load to parse rows, default GenericFileEvent	Fully qualified IFileEvent class name
lax.parsing	Whether or not to strictly parse the file, throwing errors on things like unterminated quotations. Default false.	boolean
no.headers	Whether the file contains headers, default false	boolean
field.names	Comma separated list of field names, must be equal to number of rows	CSV string
id.field	Field containing event ID, must be in field.names	string

geo.lat.field	Field containing latitude, must be in field.names	string
geo.lon.field	Field containing longitude, must be in field.names	string
publisher	Publisher for event	string
type	Type for event	string

Table 86 - CSV File Parser options

2.5.2 Generic JSON Parser

`cg.engine.event.ingest.data.parsers.genericJson.GenericJsonEventParser`
Makes an event from arbitrary JSON.

Property	Description	Data Type or Valid Values
date.found.name	Dotted notation path to the name of the event date found.	Dot notation path
enable.indexing.mode	If true, adds . to array element keys.	boolean
event.attribute.x.name	event.attribute.#.path, .name, .type specify the path to an attribute that, if found, will be decoded into the event with the given type. It is not an error for an event attribute to be missing.	string
event.attribute.x.path	Dot notation path to attribute	Dot notation path
event.attribute.x.type	Type of attribute x, **`TEXT`** or **`NUMBER`**	**`TEXT`** or **`NUMBER`**
geom.name	geom.name or alternately lat.name,lon.name are used to define a geo-location for the event. If both geom and lat/lon are present, geom.name is used. It is not an error if these are specified but not found in the message. If not specified (or specified but not found), event geo-location is null.	Dot notation path
lat.name	Dot notation path to latitude	Dot notation path
lon.name	Dot notation path to longitude	Dot notation path
hit.x.json_path	Points to the dotted notation path of a repetitive structure in the message.	Dot notation path
hit.x.type	Specifies what to name it, so that it doesn't wind up with a deeply nested name.	string
id.name	Dotted notation path to the name of the event id.	Dot notation path
payload.name	The dotted notation path to the payload in the message.	Dot notation path
publisher.name	Dotted notation path to the name of	Dot notation path

	the event publisher. Specify publisher.name or publisher.value.	
publisher.value	Constant value for the publisher. Specify publisher.name or publisher.value.	string
required.attribute.x.path	Specifies the path of an attribute that must be present for the message to be parsed. All required attributes must be present.	Dot notation path
rights.name	Dotted notation path to the name of the rights. Specify rights.name or rights.value.	Dot notation path
rights.value	Constant value of the rights. Specify rights.name or rights.value.	string
source.name	Dotted notation path to the name of the event source. Specify source.name or source.value.	Dot notation path
source.value	Constant value of the source. Specify source.name or source.value.	string
title.name	Dotted notation path to the name of the event title if not specified, title is null.	Dot notation path
type.name	Dotted notation path to the name of the event type. Specify type.name or type.value.	Dot notation path
type.value	Constant value of the type. Specify type.name or type.value.	string
data.groups	Comma separated string of data groups to add	CSV string

Table 87 - Generic JSON Parser options

2.5.3 Simple LUX Event Parser

`icg.engine.util.event.parsers.SimpleLUXEventParser`

Parsed an input XML or JSON String as a LUXEvent. No configuration options.

2.5.4 Streaming HTML Parser

`icg.engine.event.ingest.data.parsers.html.StreamingHtmlParser`

Parses an HTML stream into a title and an article.

Property	Description	Data Type or Valid Values
publisher	Constant value for publisher	string
type	Constant value for type	string

Table 88 - Streaming HTML Parser options

2.6 Event Output Plugins

All Event Output Plugins have a GeoAnalyticFilter that can be used for filtering events, so in addition to the below properties you can use GeoAnalyticFilter properties found in Table 4.

2.6.1 JMS Event Output

`icg.engine.event.output.jms.JMSEventOutput`

Sends events to JMS.

Property	Description	Data Type or Valid Values
jms.spring.path	Path to spring file to load JMS beans	File path
jms.connection.factory.bean.name	Spring bean for	string
jms.broker.username	Username for JMS broker	string
jms.broker.password	Password for JMS broker	string
jms.queue.name	Queue name to send alerts to. Specify either jms.queue.name or jms.topic.name.	string
jms.topic.name	Topic name to send alerts to. Specify either jms.queue.name or jms.topic.name.	string
jms.producer.threads	How many JMS threads to run, if value <= 0, default of 1-per-core will be used	Positive integer
jms.producer.use.message.id	Optional, whether to use the message's ID, default true	boolean
jms.producer.use.message.timestamp	Optional, whether to use the message's timestamp, default true	boolean
jms.producer.delivery.mode	Optional, JMS delivery mode, default `PERSISTENT`	`PERSISTENT`, `NON_PERSISTENT`, or `RELIABLE`
jms.producer.max.transaction.size	Optional, JMS transaction side, default 0	Integer string
<AbstractLUXEventConverter props>		

Table 89 - JMS Event Output options

2.6.2 Kafka Event Output Avro

`icg.engine.event.output.kafka.KafkaEventOutputAvro`

EventOutput to send events to Kafka using Avro serialization.

Property	Description	Data Type or Valid Values
bootstrap.servers	Connection URL for Kafka	URL

	hostname1:port1,hostname2:port2,hostname3:port3[/chroot/path]	
client.id	Client ID for Kafka connection, default “lux-client”	string
enriched.attributes	Optional CSV string of property names. To accommodate AVRO, glom all properties with the same name into an object node that has that name and an array of those values.	CSV string
group.id	Group ID for Kafka connection, default “lux”	string
num.threads	Number of threads to run	Positive integer
schema.registry.url	URL of Avro registry	URL
topic	Kafka topic to send to	string

Table 90 - Kafka Event Output Avro options

2.6.3 Kafka Event Output V8

`icg.engine.event.output.kafka.KafkaEventOutputV8`

EventOutput to send events to Kafka, use with Kafka 0.8.2.x

Property	Description	Data Type or Valid Values
kafka.node.port	Kafka URL connection string	URL
num.threads	Number of threads to run	Positive integer
topic	Kafka topic to send to	string

Table 91 - Kafka Event Output V8 options

2.6.4 Kafka Event Output V9

`icg.engine.event.output.kafka.KafkaEventOutputV9`

EventOutput to send events to Kafka, use with Kafka v0.9.x

Property	Description	Data Type or Valid Values
kafka.node.port	Kafka URL connection string	URL
num.threads	Number of threads to run	Positive integer
topic	Kafka topic to send to	string

Table 92 - Kafka Event Output V9 options

2.6.5 JSON Mapping Event Converter

`icg.engine.util.event.lux.JSONMappingEventConverter`

AbstractEventConverter to transform events to a new JSON format

Property	Description	Data Type or Valid Values

converter.name	The class name to enable this event converter	**`icg.engine.util.event.lux.JSONMappingEventConverter`**
event.query.x	Query into the LUX Event to what you'd like to map, e.g. **`/attributes/foo`**	XPath
json.mapping.x	Path to where the attribute should appear in the resulting JSON, e.g. **`/myjson/properties/foo`**	XPath
entry.type.x	Type of value	**`TEXT`**, **`INT`**, or **`FLOAT`**
fixed.value.x	Maps a fixed value to **`json.mapping.x`** rather than a value from **`event.query.x`**	String or number

Table 93 - JSON Mapping Event Converter options